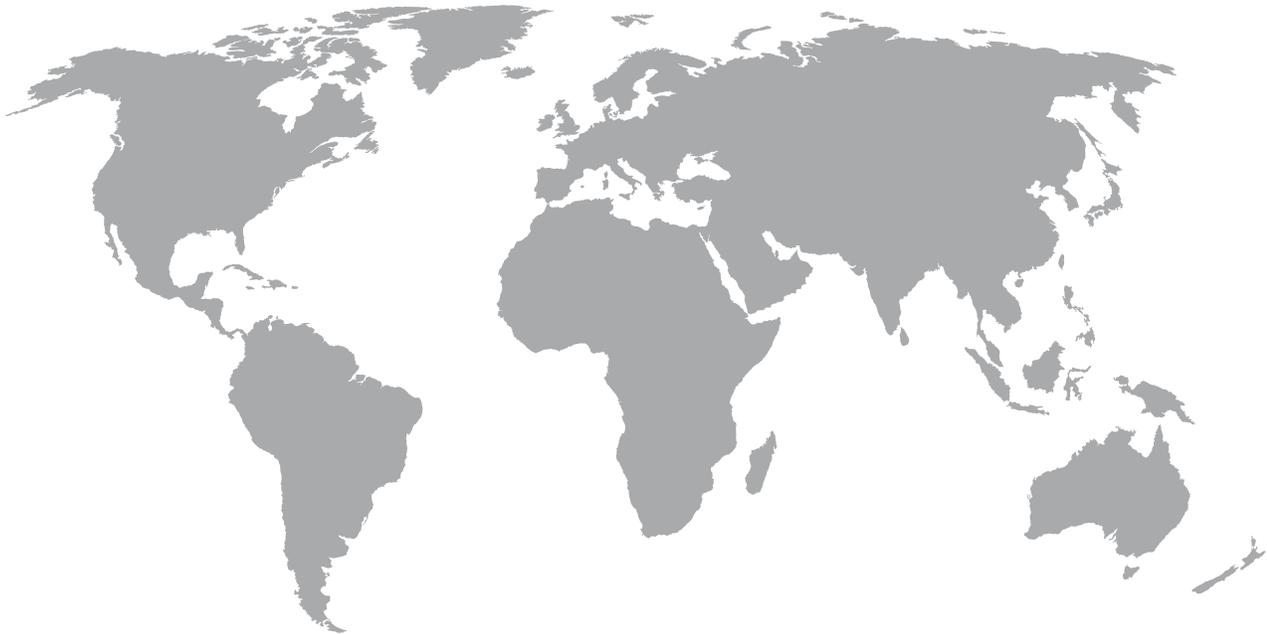




# Global Adult Tobacco Survey (GATS)



**Garantía de la calidad: directrices y documentación**



**Encuesta Mundial Sobre Tabaco en Adultos  
(GATS)  
Garantía de la calidad: directrices  
y documentación**

Septiembre de 2020

## **Encuesta Mundial Sobre Tabaco en Adultos (GATS)** Protocolo integral normalizado

.....

### **Cuestionario de la encuesta GATS**

Cuestionario básico con preguntas opcionales

Especificaciones pregunta por pregunta

### **Diseño de la muestra de la encuesta GATS**

Manual de diseño de la muestra

Manual de ponderación de la muestra

### **Ejecución de la encuesta GATS sobre el terreno**

Manual del encuestador sobre el terreno

Manual del supervisor sobre el terreno

Manual de cartografía y establecimiento de listas

### **Gestión de datos de la encuesta GATS**

Guía del programador sobre el sistema general de encuestas

Especificaciones de programación del cuestionario básico

Plan de ejecución de la gestión de datos

Guía de capacitación para la gestión de datos

### **Garantía de la calidad de la encuesta GATS: directrices y documentación**

#### **Paquete de análisis y presentación del informe de la encuesta GATS**

Plantillas de notas descriptivas

Informe del país: plan y directrices de tabulación

Definiciones de los indicadores

#### **Publicación y difusión de los datos de la encuesta GATS**

Política de publicación de datos

Difusión de datos: directrices para la publicación inicial de los datos

## **Cita propuesta**

Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Garantía de la calidad: directrices y documentación*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2020.

## **Agradecimientos**

### ***Organizaciones colaboradoras de la encuesta GATS***

- Centros para el Control y la Prevención de Enfermedades de los Estados Unidos (CDC)
- Fundación de los CDC
- Escuela de Salud Pública Johns Hopkins Bloomberg (JHSPH)
- RTI International
- Organización Mundial de la Salud (OMS)

### ***Apoyo financiero***

La *Iniciativa Bloomberg para Reducir el Consumo de Tabaco* ha prestado apoyo financiero a través de la Fundación de los CDC con una donación de Bloomberg Philanthropies.

---

Descargo de responsabilidad: Las opiniones expresadas en este manual no se corresponden necesariamente con la posición de las organizaciones colaboradoras de la encuesta GATS.



# Índice

---

Capítulo	Página
<b>1. Introducción.....</b>	<b>1-1</b>
1.1 Sinopsis de la Encuesta Mundial Sobre Tabaco en Adultos .....	1-1
1.2 Uso del presente manual .....	1-2
1.3 Garantía de la calidad en el marco de la encuesta GATS.....	1-3
<b>2. Plan del proceso de garantía de la calidad.....</b>	<b>2-1</b>
<b>3. Garantía de la calidad: fase previa a la recogida de datos .....</b>	<b>3-1</b>
3.1 Proceso de diseño y revisión del cuestionario .....	3-1
3.2 Programación y control de versiones.....	3-2
3.3 Prueba preliminar .....	3-5
3.4 Finalización del cuestionario .....	3-5
3.5 Transferencia y gestión de datos .....	3-6
3.6 Proceso de diseño y revisión de la muestra .....	3-6
3.7 Preparación del archivo de selección de la muestra maestra .....	3-7
3.8 Talleres de capacitación .....	3-7
<b>4. Garantía de la calidad: recogida y gestión de datos .....</b>	<b>4-1</b>
4.1 Recogida de datos sobre el terreno: configuración y mantenimiento.....	4-1
4.2 Gestión de datos del centro de datos nacional.....	4-7
<b>5. Garantía de la calidad: fase posterior a la recogida de datos .....</b>	<b>5-1</b>
5.1 Limpieza y preparación de los datos para los cálculos de ponderación de la muestra.....	5-2
5.2 Medidas de calidad: muestreo, error de muestreo y ponderaciones.....	5-6
5.3 Medidas de calidad: cobertura, falta de respuesta y otros errores ajenos al muestreo .....	5-12
5.4 Revisión formal de la calidad estadística.....	5-18
5.5 Creación de un archivo de datos de análisis .....	5-20
5.6 Creación de archivos de datos de uso público .....	5-20
<b>6. Requisitos relativos a los datos y la documentación .....</b>	<b>6-1</b>
<b>7. Bibliografía.....</b>	<b>7-1</b>
<b>Apéndice A: Glosario.....</b>	<b>A-1</b>
<b>Apéndice B: Fase previa a la recogida de datos .....</b>	<b>B-1</b>
B.1 Proceso de adaptación y revisión del cuestionario de la encuesta GATS .....	B-1
B.2 Proceso de programación del cuestionario de la encuesta GATS .....	B-2
B.3 Proceso de configuración de los dispositivos portátiles con el programa del sistema general de encuestas.....	B-2

<b>Apéndice C: Fase posterior a la recogida de datos .....</b>	<b>C-1</b>
C.1 Códigos de disposición finales y cálculos de la tasa de respuesta .....	C-1
C.2 Patrón de ajustes de calibración de las ponderaciones con posterioridad a la estratificación entre celdas de ajuste .....	C-5
C.3 Efecto de las ponderaciones variables de la muestra en la precisión de las estimaciones de la encuesta .....	C-7
C.4 Efecto general del diseño en la precisión de las estimaciones de la encuesta y la homogeneidad entre agrupamientos dentro de la unidad primaria de muestreo de las correspondientes estimaciones principales de la encuesta.....	C-8
C.5 Margen de error de las estimaciones principales de la encuesta.....	C-10
C.6 Estimaciones de errores de muestreo .....	C-13
C.7 Tasa de cobertura del marco de hogares .....	C-20

## Cuadros

---

<b>Número</b>	<b>Página</b>
Cuadro 2-1. Diagrama de flujo de la garantía de la calidad .....	2-1

# 1. Introducción

---

El consumo de tabaco es una de las principales causas prevenibles de muerte prematura y enfermedad en todo el mundo, ya que hay unos 1400 millones de personas de 15 años o más que consumen tabaco<sup>1</sup>. Asimismo, cada año más de 8 millones de personas pierden la vida por enfermedades relacionadas con el tabaco<sup>2</sup>. De mantenerse las tendencias actuales, el consumo de tabaco podría matar a 1000 millones de personas de aquí a finales del presente siglo, y se calcula que más de tres cuartas partes de esas muertes se producirán en países de ingresos bajos y medianos<sup>3</sup>. Para monitorear y gestionar la epidemia, es fundamental contar con un mecanismo de vigilancia eficaz y sistemático.

La **Encuesta Mundial Sobre Tabaco en Adultos** (GATS), un componente del Sistema Mundial de Vigilancia del Tabaco, es una referencia mundial para el monitoreo sistemático del consumo de tabaco en adultos y el seguimiento de los principales indicadores de control del tabaco. La encuesta GATS es una encuesta de hogares con representatividad nacional dirigida a personas adultas de 15 años o más que utiliza un cuestionario básico normalizado, un diseño de muestra y unos procedimientos de recogida y gestión de datos que fueron revisados y aprobados por expertos internacionales. La encuesta GATS tiene por objeto mejorar la capacidad de los países para diseñar, ejecutar y evaluar las intervenciones de control del tabaco.

A fin de maximizar la eficiencia de los datos recabados en la encuesta GATS, se ha creado una serie de manuales. Se trata de manuales diseñados para proporcionar a los países unos requisitos uniformes, así como varias recomendaciones sobre el diseño y la realización de la encuesta en cada etapa del proceso de la encuesta GATS. Los manuales también se han concebido para ofrecer orientación sobre la manera en que un país en particular podría ajustar las características del protocolo de la encuesta GATS para aprovechar al máximo la utilidad de los datos dentro del país. Se recomienda encarecidamente seguir el protocolo normalizado con el fin de mantener la coherencia y la comparabilidad entre países.

---

Los manuales de la encuesta GATS ofrecen orientación sistemática sobre el diseño y la realización de la encuesta.

---

## 1.1 Sinopsis de la Encuesta Mundial Sobre Tabaco en Adultos

La encuesta GATS está diseñada para producir estimaciones nacionales y subnacionales entre personas adultas de distintos países. La población a la que está dirigida la encuesta incluye a todos los hombres y mujeres no institucionalizados de 15 años o más que consideran el país como su lugar de residencia habitual. Todos los miembros de la población a la que está dirigida la encuesta serán objeto de muestreo en el hogar que sea su lugar de residencia habitual.

La encuesta GATS utiliza una metodología de muestreo por zonas geográficas en múltiples etapas para identificar los hogares

---

La entrevista de la encuesta GATS consta de dos partes: el **Cuestionario para hogares** y el **Cuestionario individual**. Ambos cuestionarios se administran mediante un dispositivo electrónico de recogida de datos.

---

<sup>1</sup> Organización Mundial de la Salud. *WHO report on the global tobacco epidemic, 2019: *Orer neip to qui tobacco use**. Ginebra (Suiza): Organización Mundial de la Salud; 2019. <https://apps.who.int/iris/bitstream/handle/10665/326043/9789241516204-eng.pdf?ua=1>

<sup>2</sup> 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 (Washington): Institute for Health Metrics and Evaluation; 2018.

<sup>3</sup> Mathers, C. D. y Loncar, D. «Projections of Global Mortality and Burden of Disease from 2002 to 2030». *PLoS Medicine*, 2006, 3(11):e442.

específicos con los que los encuestadores sobre el terreno se pondrán en contacto. Un país se divide en primer lugar en unidades primarias de muestreo y, posteriormente, en segmentos dentro de esas unidades primarias de muestreo y en hogares dentro de esos segmentos. A continuación, se selecciona una muestra aleatoria de hogares que participarán en la encuesta GATS.

La entrevista de la encuesta GATS consta de dos partes: el *Cuestionario para hogares* y el *Cuestionario individual*. El *Cuestionario para hogares* (evaluación de hogares) y el *Cuestionario individual* (entrevista individual) se llevarán a cabo utilizando un dispositivo electrónico de recogida de datos.

En cada dirección que figure en la muestra, los encuestadores sobre el terreno entregarán el *Cuestionario para hogares* a un adulto que resida en el hogar. El objetivo del *Cuestionario para hogares* es determinar si el hogar seleccionado cumple con los requisitos para llevar a cabo la encuesta GATS y confeccionar un listado, o lista de preselección, de todos los miembros del hogar que reúnen las condiciones para contestar a la encuesta. Una vez elaborada la lista de preselección de residentes del hogar que reúnen las condiciones, se seleccionará un individuo al azar para que rellene el *Cuestionario individual*. En este cuestionario se formulan preguntas sobre las características personales, el tabaquismo, los cigarrillos electrónicos, el tabaco sin humo, el abandono del tabaco, el tabaquismo pasivo, los aspectos económicos, los medios de comunicación y los conocimientos, actitudes y percepciones sobre el tabaco.

## 1.2 Uso del presente manual

En el presente manual se ofrece orientación para evaluar y garantizar la calidad de los datos recogidos en la encuesta GATS. Dicho manual complementa los manuales de la encuesta GATS existentes, en la medida en que no incluye descripciones de procedimientos documentados en otros lugares. En vez de ello, el manual describe los elementos del proceso de garantía de la calidad que deberían darse a lo largo de la ejecución de la encuesta GATS. El cumplimiento de estas directrices de garantía de la calidad es sumamente importante para el éxito de esta encuesta. El manual está destinado a todas las personas que participan en el proceso de ejecución y garantía de la calidad de la encuesta GATS, incluido el personal de tecnología de la información (TI) y los estadísticos responsables del muestreo y la ponderación. En el *apéndice A* se facilita un glosario al respecto.

En el presente documento, el proceso de garantía de la calidad de la encuesta GATS está organizado cronológicamente en tres fases: la fase previa a la recogida de datos, la recogida y gestión de datos y la fase posterior a la recogida de datos. Las actividades de garantía de la calidad varían en función de las distintas fases de dicho proceso, ofreciéndose directrices aplicables en cada una de ellas. El *capítulo 2* comienza con la exposición de un plan del proceso en el que se describe el alcance general de las actividades de garantía de la calidad.

En el *capítulo 3: fase previa a la recogida de datos* figuran distintas secciones en las que se describen las actividades de garantía de la calidad que deben llevarse a cabo durante el diseño y la revisión del cuestionario específico para el país, la programación del cuestionario para el dispositivo portátil utilizado durante la recogida de datos, el diseño de la muestra y la creación del archivo de selección de la muestra maestra.

El *capítulo 4: recogida y gestión de datos* se divide en tres secciones. En la primera sección se habla de las actividades de garantía de la calidad realizadas sobre el terreno, entre ellas, la configuración y el

mantenimiento del sistema de gestión de casos y el cuestionario. En dicha sección se incluyen también diversas subsecciones relativas a la gestión de datos y se establecen las normas para la presentación periódica de información durante el trabajo sobre el terreno. En la segunda sección se describen las actividades de gestión de datos y garantía de la calidad que tienen lugar una vez realizadas las entrevistas y subidos los datos al sitio central. La tercera sección hace referencia a las cuestiones de agregación y control de la calidad que se dan a nivel nacional. En esta sección se describe asimismo la preparación del archivo de datos brutos para su limpieza.

En el *capítulo 5: fase posterior a la recogida de datos* se describen las actividades de garantía de la calidad que han de llevarse a cabo una vez concluido el trabajo sobre el terreno y preparado el archivo de datos brutos para su limpieza. En la primera sección se exponen los procedimientos de limpieza y preparación de los datos para los cálculos de ponderación de la muestra. En la segunda sección se habla de la evaluación de la calidad del muestreo, el error de muestreo y las ponderaciones de la muestra, mientras que la tercera sección se centra en la evaluación de la calidad de la cobertura, la falta de respuesta y otros errores ajenos al muestreo. Por último, en la cuarta sección se resume el procedimiento de creación del archivo final de datos de análisis.

En tales capítulos se pretende ofrecer directrices y recomendaciones de garantía de la calidad que constituyen ejemplos de mejores prácticas. En los apéndices se proporciona información detallada sobre las metodologías o la documentación de cada sección de los capítulos, según proceda.

### **1.3 Garantía de la calidad en el marco de la encuesta GATS**

La garantía de la calidad es un proceso que consiste en actividades sistemáticas diseñadas con el objetivo de asegurar, evaluar y confirmar la calidad de los datos recogidos durante una encuesta (Biemer y Lyberg, 2003). Anteriormente, la calidad de los datos era sinónimo de exactitud de los datos, pero el término se ha ido transformando en un concepto más general. Los datos de alta calidad se definen simplemente como datos adecuados para su finalidad prevista. Esto no solo engloba la exactitud, sino también características como la puntualidad, la accesibilidad y la comparabilidad, lo que hace que la calidad sea un concepto multidimensional. Un conjunto de datos resulta exacto en la medida en que esté libre de errores. Asimismo, puede calificarse de puntual si está disponible en el momento oportuno. La accesibilidad viene determinada por la relativa facilidad o dificultad de uso. Los datos de calidad son comparables cuando son los mismos de una unidad a otra, al margen de que esa comparación de unidades se realice entre personas, encuestadores o unidades primarias de muestreo, o bien de un país a otro. La normalización aumenta la comparabilidad de los datos de la encuesta; por ejemplo, el uso de preguntas normalizadas garantiza que diferentes encuestadores formulen las mismas preguntas de la misma manera. De igual modo, la normalización de los procesos de garantía de la calidad asegura que los errores de datos se corrijan de la misma manera.

Las múltiples dimensiones de la calidad suelen conllevar una serie de restricciones mutuas, por lo que la maximización de la calidad es una cuestión de conciliación y equilibrio entre todas ellas. Por ejemplo, los datos de alta calidad deben ser exactos, pero también presentarse puntualmente. Si se tarda mucho en garantizar un alto grado de exactitud, los datos pueden resultar irrelevantes.

Las directrices expuestas en el presente documento constituyen procedimientos normalizados para la garantía de la calidad de los datos recogidos mediante la ejecución de la encuesta GATS. Se alienta a

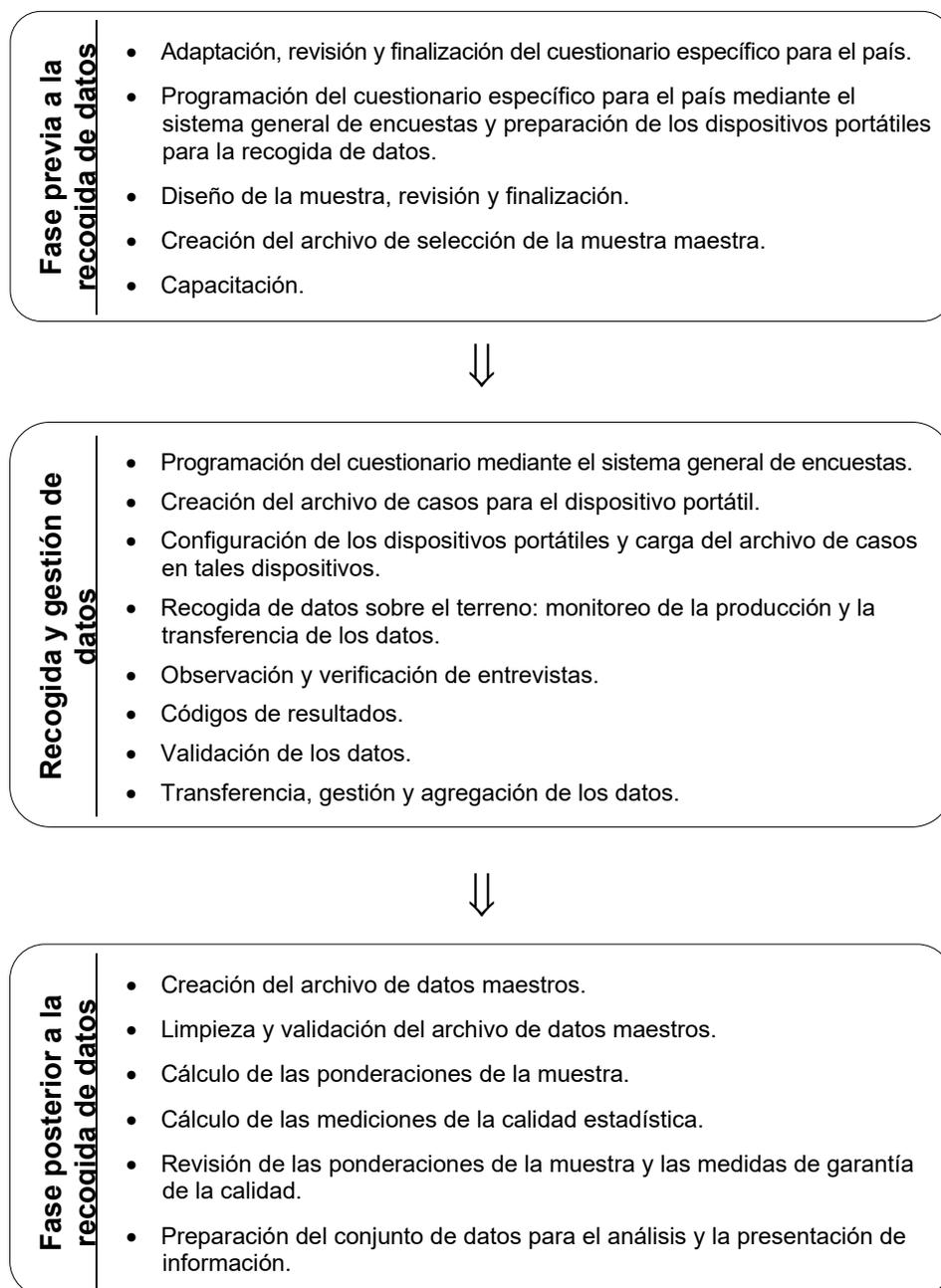
los países a que incorporen actividades adicionales de garantía de la calidad más allá de las descritas en este documento.

## 2. Plan del proceso de garantía de la calidad

---

En el cuadro que aparece a continuación figura un diagrama de las actividades de garantía de la calidad en el que se expone la secuencia temporal general y el flujo de pasos sin solapamiento. Dichos pasos se organizan en tres fases: la fase previa a la recogida de datos, la recogida de datos y la fase posterior a la recogida de datos.

**Cuadro 2-1. Diagrama de flujo de la garantía de la calidad**





### 3. Garantía de la calidad: fase previa a la recogida de datos

---

La fase previa a la recogida de datos engloba una serie de tareas que deben realizarse a modo de preparación para la recogida de datos de la encuesta completa. Entre estas tareas figuran el diseño del cuestionario, la programación, la realización de la prueba preliminar, el diseño de la muestra, la preparación del archivo de selección de la muestra maestra y la capacitación. En el presente capítulo se describe el mecanismo de garantía de la calidad y las directrices normalizadas que deben seguirse en estas diversas áreas. Concretamente, en él se abordan los siguientes aspectos:

- 3.1 Proceso de diseño y revisión del cuestionario
  - 3.1.1 Directrices para la adaptación del cuestionario
- 3.2 Programación y control de versiones
  - 3.2.1 Procesos de programación
  - 3.2.2 Comprobación del cuestionario
  - 3.2.3 Configuración del dispositivo portátil
- 3.3 Prueba preliminar
- 3.4 Finalización del cuestionario
- 3.5 Transferencia y gestión de datos
- 3.6 Proceso de diseño y revisión de la muestra
- 3.7 Preparación del archivo de selección de la muestra maestra
- 3.8 Talleres de capacitación

#### 3.1 Proceso de diseño y revisión del cuestionario

La encuesta GATS establece un proceso normalizado para que los países participantes diseñen sus cuestionarios específicos. El Comité de Revisión del Cuestionario de la encuesta GATS es un grupo integrado por expertos en control del tabaco y diseño de cuestionarios, procedentes de países tanto desarrollados como en desarrollo, que se encarga de revisar y aprobar todos los cuestionarios de la encuesta GATS. Dicho Comité colabora estrechamente con todos los países para adaptar el cuestionario de la encuesta GATS a la situación de cada uno de ellos, manteniendo al mismo tiempo las preguntas básicas normalizadas de la encuesta GATS (consúltese el *Cuestionario básico con preguntas opcionales de la encuesta GATS* para obtener información más detallada al respecto) a fin de garantizar la comparabilidad entre países. En el *apéndice B.1* puede encontrarse información detallada específica sobre la adaptación del cuestionario y el proceso de revisión.

---

**El Comité de Revisión del Cuestionario de la encuesta GATS revisa y aprueba el cuestionario para garantizar la calidad y la comparabilidad entre países.**

---

##### 3.1.1 Directrices para la adaptación del cuestionario

El Comité de Revisión del Cuestionario recomienda a los países que sigan las directrices concretas que se indican a continuación al adaptar el *Cuestionario básico de la encuesta GATS* para facilitar el proceso de revisión y garantizar el cumplimiento de las normas de calidad:

- Resaltar todas las adaptaciones al país que se hayan realizado en el *Cuestionario básico de la encuesta GATS* (a fin de facilitar la consulta). Esto incluye las listas de elementos de las preguntas, las categorías de respuesta, las preguntas opcionales y las preguntas añadidas para el país.
- Utilizar el formato de texto tachado para señalar cualquier pregunta básica que el país desee eliminar (a fin de facilitar la consulta).
- Mantener la numeración y el orden de las preguntas básicas y opcionales para garantizar la coherencia a efectos de comparación entre países. (Esto no siempre es una solución viable u óptima, pero debe llevarse a cabo en la medida de lo posible).
- En el caso de las preguntas nuevas añadidas específicamente en relación con el país, usar letras dobles según la sección (por ejemplo, AA10, BB17, EE4). De este modo, no se alterará la numeración vigente al añadir nuevas preguntas relativas al país.
- Modificar las instrucciones de salto únicamente si es necesario para incorporar las preguntas opcionales añadidas o las preguntas añadidas para el país.
- No revisar las preguntas básicas normalizadas (a excepción de las categorías específicas para el país) con el fin de mantener la coherencia en la comparación entre países. (Cualquier excepción al respecto deberá justificarse ante el Comité de Revisión del Cuestionario).
- Limitar el número de preguntas adicionales para que el cuestionario tenga una extensión razonable.
- Limitar la complejidad de las preguntas adicionales para facilitar la programación del cuestionario de cada país.
- Justificar las adaptaciones. La aportación de la correspondiente justificación ayuda al Comité de Revisión del Cuestionario en el proceso de revisión y aprobación de los cuestionarios de los países.

### 3.2 Programación y control de versiones

La encuesta GATS aplica un método de recogida de datos basado en la entrevista personal asistida por computadora mediante la utilización de dispositivos portátiles como instrumentos electrónicos. Los dispositivos portátiles están equipados con el software del sistema general de encuestas<sup>1</sup>. Siempre hay características específicas de cada país, ya que cada uno de ellos adapta el cuestionario a su medida, cuenta con diferentes idiomas que requieren fuentes especiales y puede que utilice un hardware distinto (por ejemplo, se pueden usar diversos modelos de dispositivos portátiles). Esta complejidad conlleva la necesidad de atenerse a una serie de pasos comunes y mejores prácticas.

---

**Es necesario atenerse a las comprobaciones y los procesos normalizados para garantizar que la programación de los cuestionarios de la encuesta GATS se realice con un alto grado de calidad y eficiencia.**

---

Se ha establecido una secuencia de pasos (expuesta en el *apéndice B.2*) que requieren que el país organizador realice aportaciones cooperativas y energéticas con antelación suficiente (preferiblemente entre 6 y 8 semanas antes del inicio de la capacitación para la realización de una prueba preliminar). Tales pasos

---

<sup>1</sup> El software del sistema general de encuestas está diseñado para su ejecución en una plataforma Android y se ha probado e implementado utilizando tabletas con sistema operativo Android.

exigen que el personal del país organizador (como el personal de TI/los programadores, el personal de la encuesta y los gestores de la encuesta) estén identificados y disponibles para trabajar ya sea durante el taller de orientación o inmediatamente después. Dicho personal recibirá capacitación presencial o a través de seminarios web para comenzar la labor preparatoria y se espera que dedique un número considerable de horas de trabajo a prepararse para la prueba preliminar de la encuesta GATS y la ejecución de la encuesta completa.

En las subsecciones que figuran a continuación se exponen diversas cuestiones específicamente relacionadas con el proceso de preparación del cuestionario y la encuesta que deben ponerse de relieve para aumentar la calidad.

### 3.2.1 Procesos de programación

A fin de mantener la calidad y la eficiencia de los preparativos de la encuesta GATS, es necesario llevar a cabo los siguientes procesos:

- El personal de la encuesta o sobre el terreno o bien los puntos de contacto del organismo de ejecución, y no el personal de TI, son los responsables de verificar la exactitud del idioma específico del país (las traducciones) en función del cuestionario de su país. El personal de TI colaborará con el personal de la encuesta o sobre el terreno o bien con los puntos de contacto del organismo de ejecución en la inserción del texto mediante el software de diseño del cuestionario del sistema general de encuestas.
- El punto de contacto del organismo de ejecución debe aprobar el *Cuestionario para hogares* y el *Cuestionario individual* (por escrito) antes de la capacitación del personal sobre el terreno (con los correspondientes números de versión).
- Modificaciones del texto: cualquier modificación en la redacción del cuestionario de la encuesta deberá conllevar la correspondiente actualización en todos los idiomas en los que dicho cuestionario vaya a administrarse.
- La modificación de la lógica de programación (intervalos, controles de validez, etcétera) de un cuestionario aprobado exige una nueva revisión por parte del Comité de Revisión del Cuestionario.
- La modificación del contenido o de la redacción de una pregunta o de las respuestas requiere una nueva revisión por parte del Comité de Revisión del Cuestionario.
- Control de las versiones de los archivos DB3 (archivos de base de datos de los cuestionarios de la encuesta): el país organizador es propietario de los archivos DB3 y el personal de TI y de la encuesta del país es responsable de mantener un registro de auditoría de las modificaciones de los cuestionarios y las especificaciones de programación del sistema general de encuestas.
- Tras la reunión de información *a posteriori* sobre la prueba preliminar y el comienzo de la ejecución completa, el país debe enviar los archivos DB3 finales de CMSDB, Survey0 y Survey1.

### 3.2.2 Comprobación del programa del cuestionario

Para comprobar el programa del cuestionario deben seguirse y documentarse varios pasos orientados a la garantía de la calidad:

- Los programadores deben actualizar la versión del programa del cuestionario cada vez que se realice una revisión. Asimismo, este aspecto ha de confirmarse antes de empezar a revisar el

programa del cuestionario. El software del sistema general de encuestas alberga un archivo de archivos DB3 anteriores a modo de historial y copia de seguridad.

- Para cada idioma, se revisan las especificaciones completas de programación del sistema general de encuestas examinando cada pregunta en el dispositivo portátil y se comprueba, para cada pregunta, que el texto de esta, las respuestas y los caracteres especiales se correspondan exactamente con lo contemplado en el documento de *Especificaciones de programación del cuestionario de la encuesta GATS*. Si se detecta algún problema, se documenta claramente en las especificaciones; tras una iteración se solucionan los problemas y luego se procede a la iteración de comprobación completa. Este proceso iterativo continúa hasta que se aprueban todas las preguntas en el dispositivo portátil. No es posible, bajo ningún concepto, pasar por alto este proceso. El procedimiento debe llevarse a cabo con la ayuda de un experto en encuestas en el idioma local. Dicho procedimiento no debe dejarse en manos de los programadores ni de personal que carezca de los conocimientos lingüísticos necesarios. El punto de contacto en el país debe aprobar la versión final tras revisar las modificaciones o actualizaciones realizadas. Antes de dar el visto bueno es preciso revisar las especificaciones para verificar que todas y cada una de las preguntas se hayan comprobado repasando reiteradamente el cuestionario (consúltese la *Guía de capacitación para la gestión de datos de la encuesta GATS, módulo 2.1*, para obtener información detallada al respecto).
- Los patrones de salto y el texto en inglés siempre se someten a pruebas rigurosas con anterioridad a la visita sobre el terreno. El software del cuestionario no debe modificarse durante la recogida de datos sobre el terreno a menos que sea absolutamente imprescindible.
- Para cada idioma, el recuento de las etiquetas HTML de apertura y cierre y de los símbolos ( , ), [ , ], { , y } debe ser el mismo. De lo contrario, será necesario realizar una nueva comprobación y corregir este aspecto.

### 3.2.3 Configuración del dispositivo portátil

Para cargar el programa del sistema general de encuestas en los dispositivos portátiles hay que atenerse a una serie de procedimientos formales de garantía de la calidad:

- Antes de realizar la carga, tendrá que haberse concluido la configuración inicial de todos los dispositivos portátiles y, si procede, tendrán que haberse desinstalado las versiones anteriores del software de la encuesta GATS y haberse eliminado las correspondientes carpetas.
- Se imprime una serie de etiquetas con casillas de verificación de los pasos de control de la calidad (a la hora de instalar la encuesta GATS en los dispositivos portátiles) que se adhieren a tales dispositivos. La persona encargada de controlar la calidad de la carga en el dispositivo debe confirmar que se ha llevado a cabo cada uno de los pasos.

En el *apéndice B.3* se especifican con detalle los pasos del proceso de carga en los dispositivos portátiles (consúltese la *Guía de capacitación para la gestión de datos de la encuesta GATS, módulo 2.2*, para obtener información detallada al respecto).

### 3.3 Prueba preliminar

La realización de la prueba preliminar en el país, con anterioridad a la ejecución de la encuesta completa, tiene seis objetivos principales. El logro de tales objetivos, los cuales se exponen a continuación, ayudará a garantizar que la recogida de datos de la encuesta completa de la encuesta GATS se lleve a cabo con un alto grado de calidad:

1. Brindar asistencia al personal en el país para programar, probar y mantener una encuesta basada en el uso de dispositivos portátiles, en función de la necesidad.
2. Revisar el cuestionario de la prueba preliminar de la encuesta GATS, en particular la formulación de las preguntas y respuestas, las instrucciones para el sondeo, la programación de la lógica de salto y la extensión y el tiempo de duración de la entrevista.
3. Brindar asistencia al personal en el país en la utilización del hardware y el software de los dispositivos portátiles, en función de la necesidad.
4. Observar la realización de la prueba preliminar en un entorno rural y en otro urbano, siempre que sea viable y proceda.
5. Someter a prueba el plan de recogida, transferencia y gestión de datos durante la prueba preliminar que vaya a adaptarse para la ejecución de la encuesta completa.
6. Informar a los asociados de la encuesta GATS sobre los puntos fuertes y débiles y las recomendaciones en relación con los preparativos y la realización de la prueba preliminar.

---

**La prueba preliminar de la encuesta GATS permite probar el cuestionario y los procedimientos de recogida y gestión de datos para garantizar que la encuesta completa se lleve a cabo con un alto grado de calidad.**

---

### 3.4 Finalización del cuestionario

Para finalizar el cuestionario de la encuesta GATS con miras a la ejecución de la encuesta completa, deberán seguirse los pasos que figuran a continuación:

- **Reunión de información a posteriori sobre la prueba preliminar.** Todo el personal (encuestadores sobre el terreno, supervisores sobre el terreno, personal del proyecto) debe aportar su opinión sobre el funcionamiento del cuestionario en su conjunto, así como señalar las preguntas concretas de la encuesta en las que debe centrarse la revisión.
- **Análisis de los datos de la prueba preliminar.** El análisis de los datos de la prueba preliminar puede proporcionar información valiosa para las revisiones posteriores a dicha prueba. Estas son algunas de las cuestiones concretas que examinar:
  - Falta de respuestas a los ítems: detectar las preguntas con alta tasa de respuestas del tipo «No sabe» o «No contesta».
  - Frecuencias: examinar las distribuciones de frecuencia de las preguntas clave para determinar si los datos parecen adecuados. La detección de alguna anomalía podría indicar un problema de diseño del cuestionario o de algún otro tipo (por ejemplo, una cuestión de capacitación).

- Intervalos permitidos: examinar la idoneidad de los intervalos permitidos programados en los dispositivos portátiles y realizar los ajustes necesarios para la ejecución de la encuesta completa.
- Categorías de respuesta: determinar si las adaptaciones del cuestionario específico para el país (por ejemplo, la modificación de las listas de elementos de las preguntas y las categorías de respuesta) han resultado adecuadas. Medir, por ejemplo, las frecuencias de las marcas de cigarrillos a fin de modificar la lista de marcas en el cuestionario de la encuesta completa.
- **Revisión para la ejecución de la encuesta completa.** La versión revisada con posterioridad a la prueba preliminar debe someterse al examen del Comité de Revisión del Cuestionario, incluida la justificación (basada en los resultados de la prueba preliminar) de las correspondientes correcciones. Es necesario seguir las mismas directrices de garantía de la calidad descritas en las *secciones 4.1 y 4.2* (según proceda) con el fin de preparar el cuestionario para la ejecución de la encuesta completa (incluida la programación revisada, etcétera).

### 3.5 Transferencia y gestión de datos

La prueba preliminar constituye un simulacro perfecto de la ejecución de la encuesta completa en lo referente al proceso de gestión de datos. El proceso de transferencia y gestión de datos se finalizará para la ejecución de la encuesta completa basándose en la experiencia y las recomendaciones de la prueba preliminar. Por tanto, cada fase del proceso de gestión de datos debe desarrollarse de conformidad con el plan de gestión de datos teniendo en cuenta la encuesta completa. Por ejemplo:

- Durante la prueba preliminar debe ponerse en práctica el mecanismo de retroalimentación para la difusión de informes.
- El método de transmisión de datos, por ejemplo, a través de WiFi, SIM, la nube o un sitio FTP, ha de probarse durante la prueba preliminar, debiendo efectuarse los ajustes necesarios, si procede, para la encuesta completa en función de las enseñanzas extraídas de dicha prueba.

### 3.6 Proceso de diseño y revisión de la muestra

La encuesta GATS mantiene un proceso normalizado para que los países participantes elaboren su diseño de la muestra. El Comité de Revisión de la Muestra de la encuesta GATS es un grupo integrado por expertos en metodología de muestreo, procedentes de países tanto desarrollados como en desarrollo, que se encarga de revisar y finalizar los diseños de la muestra de los países participantes. Los puntos de contacto de los CDC y dicho Comité colaboran estrechamente con cada país para adaptar el diseño de la muestra de la encuesta GATS, respetando al mismo tiempo las normas de la encuesta GATS para garantizar la comparabilidad en el análisis entre países. Los detalles específicos del diseño de la muestra de la encuesta GATS pueden consultarse en el *Manual de diseño de la muestra de la encuesta GATS*.

Como parte del proceso de muestreo de la encuesta GATS, tal vez sea necesario realizar actividades de cartografía y establecimiento de listas para crear un marco de muestreo probabilístico de zonas apropiado. Las recomendaciones y los procedimientos completos de garantía de la calidad en lo que respecta a la cartografía y el establecimiento de listas figuran en el *Manual de cartografía y establecimiento de listas de la encuesta GATS*.

### 3.7 Preparación del archivo de selección de la muestra maestra

Una vez extraída la muestra de la encuesta GATS, es preciso preparar un archivo de selección de la muestra maestra. El archivo de selección de la muestra maestra es un conjunto de datos que contiene un identificador de caso para cada hogar de la lista de muestreo junto con la información necesaria para calcular la ponderación de la muestra y analizar los datos complejos de la encuesta.

El archivo de selección de la muestra maestra incluye identificadores de muestra e identificadores geográficos. Los identificadores de muestra son, por ejemplo, el estrato, la unidad primaria de muestreo, la unidad secundaria de muestreo (segmento), el tipo de lugar de residencia (entorno urbano o rural) y la asignación de género preseleccionada del hogar, si se utiliza la aleatorización por género. La identificación geográfica consiste, por ejemplo, en la información de la dirección, incluidos los códigos de país, región, provincia o estado, circunscripción o distrito y ciudad, pueblo o aldea.

---

**El archivo de selección de la muestra maestra incluye un único identificador de caso, identificadores de muestra e identificadores geográficos del hogar.**

---

Una vez finalizado el archivo de selección de la muestra maestra, este debe organizarse y convertirse en archivo de texto mediante Excel y/o MS Access. Dicho archivo de texto servirá para cargar los casos en los dispositivos portátiles con miras a la recogida de datos, según proceda. Asimismo, al utilizar el software del sistema general de encuestas, las pantallas de Tabla de casos e Información de casos deberán finalizarse para reflejar el archivo de casos revisado. En el siguiente capítulo se ofrecen más detalles sobre este proceso.

### 3.8 Talleres de capacitación

La capacitación es un importante mecanismo de garantía de la calidad para la ejecución de la encuesta GATS. Los talleres de capacitación impartidos a los países abarcan la orientación sobre la encuesta GATS, la prueba preliminar (analizada en secciones anteriores) y la capacitación para la encuesta completa. Los cursos de capacitación forman parte de las normas y el protocolo de la encuesta GATS destinados a garantizar la recogida de datos de alta calidad en cada país.

Se imparten cursos de capacitación de la encuesta GATS en todos los niveles del equipo de encuestas del país involucrados en la ejecución de la encuesta, desde los encuestadores hasta los supervisores, pasando por el personal de TI y el equipo central encargado de supervisar todo el proceso de principio a fin. Los asociados nacionales e internacionales facilitan la capacitación de la encuesta GATS en todos los ámbitos. La capacitación de todo el personal posibilita que todos los agentes involucrados en el proceso de encuesta tengan clara su función en la garantía de la buena calidad de los datos. La capacitación es un proceso continuo que se lleva a cabo con anterioridad a la ejecución de la encuesta y durante dicha ejecución, con una sesión informativa exhaustiva al final de cada capacitación.

Los objetivos de la capacitación de la encuesta GATS son los siguientes:

- Garantizar la aplicación normalizada del protocolo, las normas y los materiales de la encuesta.
- Asegurar una comprensión clara de los fundamentos de la encuesta y su protocolo.
- Garantizar la aplicación normalizada del sistema electrónico de recogida, gestión, transferencia y agregación de datos de la encuesta GATS.

- Involucrar y motivar a los encuestadores.
- Realizar sugerencias prácticas.
- Mejorar la calidad general de los datos producidos.

### 3.8.1 Taller de orientación sobre la encuesta GATS

El taller de orientación constituye el primer paso del proceso de garantía de la calidad en el desarrollo de la encuesta GATS. Ello conlleva un examen en profundidad del protocolo normalizado de la encuesta GATS, lo que incluye el cuestionario, el diseño de la muestra, la recogida y gestión electrónica de datos y la difusión de información. Los países comparten sus experiencias en la realización de encuestas, lo que constituye un reflejo de su capacidad técnica para ejecutar la encuesta GATS en el país correspondiente pero también sirve para determinar la asistencia técnica necesaria por parte de los asociados al organismo en cuestión a fin de lograr los resultados deseados. Los talleres de orientación repasan las directrices de elaboración de propuestas y permiten a los países adaptar el protocolo a su situación concreta. Por último, la orientación ofrece una guía paso a paso sobre la elaboración, la presentación, la revisión y la financiación de las propuestas, que es fundamental para el inicio y la ejecución del proyecto de la encuesta.

### 3.8.2 Prueba preliminar

Los objetivos de la realización de una prueba preliminar se han detallado en la *sección 3.3*. La capacitación para la prueba preliminar abarca múltiples componentes, entre ellos, la capacitación del personal de TI, la capacitación del encuestador/supervisor sobre el terreno, la práctica sobre el terreno y la reunión de información *a posteriori*. La prueba preliminar permite al personal del proyecto adquirir experiencia en la capacitación del personal sobre el terreno. Durante la prueba preliminar, el personal de TI y sobre el terreno recibe capacitación en el uso del software y el hardware necesarios para ayudar a la ejecución de la encuesta GATS. La capacitación garantiza la preparación adecuada del personal y su conocimiento del funcionamiento del software y el hardware en cuestión.

### 3.8.3 Capacitación para la encuesta completa

La capacitación para la encuesta completa de los encuestadores y supervisores constituye un mecanismo clave de garantía de la calidad de la encuesta GATS. Su objetivo es capacitar a los encuestadores y supervisores para generar datos útiles y fiables a partir de la encuesta sobre el terreno. La capacitación abarca los fundamentos de la encuesta, enraizados en el contexto del control del tabaco y el Sistema Mundial de Vigilancia del Tabaco, así como las técnicas de entrevista, la localización de los hogares objeto de muestreo, los consentimientos, el contenido del cuestionario y los componentes prácticos, lo que incluye la escenificación de roles y la observación sobre el terreno. Los formadores revisan los cuestionarios en papel y el cuestionario en tabletas electrónicas portátiles con los participantes, a fin de mejorar su conocimiento de la encuesta. Además, gracias a las escenificaciones de roles y la práctica sobre el terreno, los destinatarios de la capacitación se familiarizan con el cuestionario y la naturaleza de los problemas previsibles durante una entrevista con encuestados de trato difícil, así como la manera de afrontar tales problemas. También se proporcionan materiales de capacitación a los alumnos a modo de referencia tanto en el transcurso de la capacitación como durante el trabajo sobre el terreno. Entre dichos materiales se incluyen el cuestionario, las especificaciones pregunta por pregunta y el manual para encuestadores.

## 4. Garantía de la calidad: recogida y gestión de datos

---

En este capítulo se describen los procesos y procedimientos de garantía de la calidad de los datos aplicables en la recogida y gestión de datos de la encuesta GATS. La recogida y gestión de datos abarca la configuración del software del sistema general de encuestas de la encuesta GATS, el uso del software durante la recogida de datos sobre el terreno, la transferencia de los datos recogidos, la agregación y la creación de un archivo maestro de la base de datos agregada. Este proceso engloba asimismo el monitoreo, la presentación de información y el apoyo técnico durante la recogida de datos sobre el terreno.

Para cualquier país determinado, la recogida de datos, la agregación y la garantía de la calidad deben llevarse a cabo en dos ámbitos: 1) sobre el terreno y 2) en el centro de datos nacional del país en cuestión. En el presente capítulo se describen los procesos y procedimientos de garantía de la calidad de los datos aplicables a cada uno de esos dos ámbitos, abordando de manera específica los siguientes aspectos:

- 4.1 Recogida de datos sobre el terreno: configuración y mantenimiento
  - 4.1.1 Archivo de casos y gestión de casos
  - 4.1.2 Dispositivos portátiles de recogida electrónica de datos y software del cuestionario de la encuesta
  - 4.1.3 Procedimientos de garantía de calidad en la recogida de datos sobre el terreno
- 4.2 Gestión de datos del centro de datos nacional
  - 4.2.1 Agregación
  - 4.2.2 Control de la calidad de los datos recogidos
  - 4.2.3 Comunicación y apoyo técnico

### 4.1 Recogida de datos sobre el terreno: configuración y mantenimiento

En el protocolo normalizado de la encuesta GATS se recomienda el establecimiento de determinados procedimientos sistemáticos y su aplicación por parte del personal de TI y sobre el terreno para mantener la exactitud e integridad de los datos durante todo el período de recogida de estos. Dichos procedimientos se describen a continuación.

#### 4.1.1 Archivo de casos y gestión de casos

1. Se recomienda que el archivo de casos se prepare y conserve en el centro de datos nacional mediante el software del sistema general de encuestas, según proceda, en función del plan de ejecución de la gestión de datos del país. (Téngase en cuenta que el archivo de casos es un subconjunto del archivo de selección de la muestra maestra).
2. Durante la creación del archivo de casos y antes de configurar el dispositivo portátil:
  - Todos los encuestadores sobre el terreno deben estar identificados y tener asignado un dispositivo portátil. Las asignaciones de identificadores de caso deberán realizarse de manera única y específica para cada encuestador sobre el terreno, utilizando la variable de identificación del encuestador sobre el terreno en el archivo de casos. En el momento de la configuración, no deberá asignarse un mismo identificador de caso a más de un identificador de encuestador sobre el terreno. Ello garantizará que no se produzcan

- duplicados, que la carga de asignación de casos del encuestador sobre el terreno sea razonable y que los casos sean fácilmente identificables.
- Todos los identificadores de encuestadores sobre el terreno deben ajustarse a la convención de la encuesta GATS de asignación de un identificador numérico de seis dígitos.
  - Todos los identificadores de caso deben ajustarse a la convención de la encuesta GATS de asignación de un identificador numérico de seis dígitos.
3. En el archivo de casos ha de incluirse una asignación por cada dispositivo portátil/encuestador sobre el terreno que vaya a recoger los datos.
- Asegúrese de que todos los encuestadores sobre el terreno que figuren en la hoja de cálculo electrónico de identificadores de encuestadores sobre el terreno tengan su correspondiente asignación en el archivo de casos. El archivo debe revisarse antes de configurar los dispositivos portátiles. Cuando se cargue en los dispositivos portátiles, el número previsto de casos tiene que aparecer en el sistema de gestión de casos. De no ser así, compruebe el identificador del encuestador sobre el terreno asignado en el software del dispositivo portátil y el identificador del encuestador sobre el terreno que figura en el archivo de casos. Cabe señalar que, para cada identificador de caso, se crearán dos registros que se cargarán en el sistema de gestión de casos (CH = 00 y CI = 01).
4. Cada tabla de casos del dispositivo portátil debe contener únicamente los casos que se asignen al encuestador sobre el terreno y en los que se pretenda que este trabaje. Normalmente, no deben asignarse y cargarse en el dispositivo portátil más de 300 identificadores de caso por encuestador sobre el terreno. Prestaciones como el tiempo de carga de la aplicación y la capacidad de localizar fácilmente un caso durante el trabajo sobre el terreno podrían verse afectadas si se carga un número de casos superior al recomendado.
5. Antes de asignar los casos en el dispositivo portátil, revise la frecuencia de los campos del archivo de casos mientras este aún se encuentre en la computadora del centro de datos nacional. Consúltese el documento relativo al archivo de casos del sistema de gestión de casos de la *Guía del programador sobre el sistema general de encuestas de la encuesta GATS* para conocer los requisitos de los campos:
- Debe comprobarse el campo «IDENTIFICADOR DE CASO» en busca de duplicados y aplicarse la norma de la encuesta GATS de utilizar identificadores numéricos de seis dígitos.
  - Compruebe que en el campo «NOMBRE DEL PROYECTO» figure el nombre del país para asegurarse de que exista un campo de identificador de país único en el archivo de casos (puede utilizarse la búsqueda por nombre del proyecto para actualización en la plantilla del archivo de casos de MS Access).
  - Compruebe que el valor de la variable «fecha de creación» sea la fecha en que se creó el archivo.
  - Compare el archivo de selección de la muestra maestra con el archivo de casos del dispositivo portátil para garantizar la integridad y la corrección del contenido. (Téngase en cuenta que el archivo de casos del dispositivo portátil es un subconjunto del archivo de selección de la muestra maestra).

En la *Guía del programador sobre el sistema general de encuestas de la encuesta GATS* figura información detallada sobre la creación del archivo de casos, las descripciones de los campos y la hoja de cálculo electrónico del inventario de encuestadores sobre el terreno de los dispositivos portátiles.

#### **4.1.2 Dispositivos portátiles de recogida electrónica de datos y software del cuestionario de la encuesta**

1. **Parámetros de configuración del encuestador sobre el terreno:** Cada dispositivo portátil debe estar etiquetado con un identificador de encuestador sobre el terreno único, ajustado a la convención de la encuesta GATS y coincidente con el identificador de encuestador sobre el terreno establecido y mostrado dentro del menú de configuración del sistema de gestión de casos. Esta información debe registrarse en el centro de datos nacional en una hoja de cálculo electrónico, junto con el nombre del encuestador sobre el terreno y el número de serie del dispositivo portátil en cuestión.
  - Coteje los parámetros de configuración del identificador del encuestador sobre el terreno: compruebe que la asignación de la hoja de cálculo electrónico, el campo del identificador del encuestador sobre el terreno en el menú de configuración y la etiqueta del dispositivo portátil coinciden con anterioridad a la recogida de datos.
2. **Versiones de software: sistema de gestión de casos y cuestionarios para hogares e individuales:** Los dispositivos portátiles deben tener cargadas las versiones finales aprobadas de los cuestionarios.
  - Verifique en cada dispositivo portátil que la versión de cada uno de los cuestionarios esté cargada y operativa comprobando la opción «Información» en la parte superior de la pantalla y verificando que las versiones de la sección «Información de la encuesta» sean correctas.
  - En la medida de lo posible, todos los dispositivos portátiles deben configurarse al mismo tiempo para garantizar la coherencia de la versión, la configuración y el proceso de control de la calidad.
3. **Carga de casos:** Al iniciar la recogida de datos o el trabajo sobre el terreno, en el sistema de gestión de casos de cada dispositivo portátil solo deberían aparecer los casos que figuren en el archivo de casos con el correspondiente identificador de encuestador sobre el terreno asociado. Los casos para la capacitación deberán borrarse antes de que comience el trabajo sobre el terreno a fin de evitar la posibilidad de trabajar con este tipo de casos y exportarlos. Los analistas tienen que estar preparados para filtrar los casos para la capacitación a partir de los datos operacionales en los supuestos en que los encuestadores sobre el terreno incumplan esta regla.

En la *Guía del programador sobre el sistema general de encuestas de la encuesta GATS* figuran instrucciones detalladas sobre la configuración de los dispositivos portátiles y la creación e instalación de los archivos conexos.

#### **4.1.3 Procedimientos de garantía de calidad en la recogida de datos sobre el terreno**

En esta sección se ofrece una sinopsis de los procedimientos de garantía de calidad recomendados que pueden aplicarse durante la recogida de datos sobre el terreno. Es preciso consultar tanto el *Manual del encuestador sobre el terreno de la encuesta GATS* como el *Manual del supervisor sobre el terreno de la encuesta GATS* para obtener información más detallada al respecto.

### ***Función del supervisor sobre el terreno***

Los supervisores sobre el terreno actúan como enlace esencial entre los encuestadores sobre el terreno y el personal de gestión de encuestas de la encuesta GATS. Se espera que tales supervisores monitoreen la producción y el desempeño, y comuniquen cualquier problema sobre el terreno que pueda repercutir en la finalización oportuna de la encuesta GATS. A ellos corresponde la responsabilidad de garantizar que los datos de su equipo se recopilen de conformidad con el protocolo de recogida de datos de la encuesta GATS y sean de la máxima calidad. Con el fin de ayudar a que los encuestadores sobre el terreno entiendan la importancia que el equipo de la encuesta GATS otorga a la calidad de los datos, los supervisores sobre el terreno deben incorporar el control de la calidad como parte integrante de sus actividades semanales.

El primer paso del supervisor sobre el terreno para garantizar la calidad de los datos es comunicar las expectativas de calidad a sus encuestadores sobre el terreno. Se espera que los supervisores proporcionen retroalimentación constante a los encuestadores sobre el terreno acerca de cuestiones de calidad, tanto positivas como negativas, y hagan hincapié en la importancia que el equipo de la encuesta GATS otorga a la *calidad*.

### ***Monitoreo de la producción***

La producción abarca todas las actividades necesarias para alcanzar con éxito los objetivos de tasa de respuesta de la encuesta GATS. Entre estos objetivos figuran el inicio de las actividades relativas a cada uno de los hogares asignados; los esfuerzos de contacto, localización y convencimiento de los encuestados que inicialmente se niegan a responder para participar en la encuesta; y la culminación satisfactoria de estas actividades, de conformidad con todas las especificaciones de la encuesta GATS. El supervisor sobre el terreno y el personal de la encuesta en el país han de colaborar entre sí en el monitoreo periódico del tiempo de realización de las entrevistas. Además, el monitoreo de la producción durante la recogida de datos resulta esencial para garantizar que la encuesta GATS se concluya en la fecha prevista y dentro del presupuesto asignado.

Los tipos de informes que se enumeran a continuación pueden generarse mediante el software del sistema general de encuestas, y deben ser utilizados por el personal de la encuesta en el país o el personal de TI para monitorear diversos aspectos de la producción, siendo lo ideal que se difundan entre los supervisores sobre el terreno periódicamente. Cuándo y dónde estén disponibles estos informes dependerá del modelo de gestión de datos establecido para cada país.

- *Estado de casos del Cuestionario para hogares y Estado de casos del Cuestionario individual:* En estos informes figuran los códigos de resultados asignados más recientemente tanto a los casos relativos a hogares como a los casos individuales gestionados por el encuestador. Este informe se visualiza a nivel nacional y puede difundirse entre los supervisores sobre el terreno y/o los encuestadores sobre el terreno.
- *Informe sobre las excepciones:* Este informe ayudaría al personal de la encuesta GATS a monitorear el número de repeticiones de visitas en los casos en que no se hubiese completado el cuestionario y a comprender los motivos de tales repeticiones.
- *Informe sobre las distribuciones de frecuencias de los Cuestionarios individuales:* Este informe ayudaría al personal de la encuesta de la encuesta GATS y al personal de TI a monitorear las respuestas al cuestionario por encuestador o unidad primaria de muestreo, con el fin de evaluar

la calidad de las entrevistas de la encuesta mediante el análisis de medidas de calidad como las siguientes:

- Problemas potenciales con el encaminamiento de la entrevista o los datos «faltantes».
- Tasas de respuesta del tipo «no sabe» o «no contesta» más altas de lo previsto.
- Tasas de respuesta del tipo «otros» (en lugar de alguna de las opciones de respuesta precodificadas) más altas de lo previsto.
- Respuestas del tipo «otros (especificar)» poco claras e imposibles de codificar con facilidad o que sean duplicados de alguna de las opciones precodificadas.

### ***Monitoreo de la transferencia de datos***

La transmisión periódica de datos por parte de los supervisores sobre el terreno al centro de datos nacional resulta fundamental para la gestión de la encuesta GATS y para la detección temprana y solución de posibles problemas. Los supervisores sobre el terreno son responsables de que cada encuestador sobre el terreno transmita los datos de forma periódica (normalmente, a diario). Tales supervisores pueden colaborar con el personal de TI del país para determinar si hay algún problema con las transmisiones.

### ***Observación***

La observación de los encuestadores sobre el terreno es una parte importante de las actividades de control de la calidad por parte de los supervisores sobre el terreno y puede servir de ayuda a dichos encuestadores para mejorar el desempeño de su trabajo. La observación de la labor de los encuestadores sobre el terreno también ayudará a evitar cualquier tentación por su parte de tomar atajos en la administración del cuestionario o de saltarse el protocolo de entrevistas de la encuesta GATS. Como mínimo, se espera que los supervisores sobre el terreno observen a cada uno de sus encuestadores sobre el terreno durante los primeros días de su período de operaciones y posteriormente con menos frecuencia, sobre la base de las directrices de los organismos de ejecución. En el marco de dicha observación, el supervisor sobre el terreno acompañará al encuestador sobre el terreno a los hogares objeto de muestreo para verificar que el encuestador en cuestión registre correctamente el resultado de la visita en el sistema de gestión de casos del sistema general de encuestas y realice adecuadamente la evaluación de hogares y la entrevista posterior. Estas visitas iniciales contribuirán en gran medida a asegurarse de detectar rápidamente los errores en la administración del cuestionario o en el uso de los códigos de resultados y de poder facilitar al encuestador sobre el terreno la capacitación u orientación adicionales necesarias.

Además, los supervisores sobre el terreno pueden decidir acompañar a los encuestadores sobre el terreno en visitas particularmente difíciles (por ejemplo, a hogares que se hayan negado a participar en la encuesta anteriormente).

### ***Verificación***

Una forma de comprobar la calidad de los datos recogidos por el personal encargado de las entrevistas es realizar entrevistas breves de verificación con los hogares que ya hayan sido evaluados y entrevistados. La realización de una breve entrevista de verificación permitirá a los supervisores sobre el terreno confirmar que el encuestador sobre el terreno haya hecho lo siguiente:

- Que haya localizado y evaluado el hogar correcto. A veces, un encuestador sobre el terreno puede seleccionar un hogar diferente del que era objeto de muestreo. Si esto sucede, el supervisor sobre el terreno debe ordenar al encuestador en cuestión que acuda al hogar correcto y realice la evaluación y la entrevista de ese hogar.
- Que haya registrado correctamente la edad, el género y la situación con respecto al tabaquismo de los miembros del hogar. Si la edad no se registra correctamente (por ejemplo, el encuestador sobre el terreno incluye a residentes menores de 15 años en la lista de preselección, además de a los residentes de 15 años o más), este error afectará a la identificación de las personas que reúnen los requisitos y podría ocasionar que el encuestador sobre el terreno seleccionase a un miembro incorrecto del hogar para la entrevista.
- Que haya administrado el *Cuestionario individual* al miembro del hogar seleccionado.

El número exacto de entrevistas de verificación se determinará conjuntamente con el personal de gestión del organismo de ejecución; sin embargo, generalmente se espera que los supervisores sobre el terreno realicen entrevistas de verificación de manera aleatoria en relación con aproximadamente el 10% de los casos asignados a cada encuestador sobre el terreno. Las entrevistas de verificación pueden consistir en 1) una breve verificación en la que se formulen al encuestado seleccionado unas pocas preguntas encaminadas a comprobar que haya realizado recientemente una encuesta sobre temas relacionados con el tabaquismo y a evaluar el desarrollo de la entrevista, o bien, 2) si el país lo desea, una repetición de la entrevista completa con el Cuestionario para hogares y, siempre que sea factible, con el Cuestionario individual. Dado que el proceso de repetición de la entrevista puede suponer una molestia para los encuestados, no es necesario administrar el Cuestionario individual en su totalidad. Además, las entrevistas de verificación pueden llevarse a cabo en soporte impreso o a través de un dispositivo portátil que tenga cargados los hogares objeto de verificación. Nunca debe usarse el dispositivo portátil del encuestador sobre el terreno para la repetición de entrevistas, sino que será preciso utilizar otro dispositivo portátil.

Posteriormente, el supervisor sobre el terreno comparará las respuestas obtenidas en la repetición de la entrevista con las registradas por el encuestador sobre el terreno. Si existen discrepancias y se sospecha que el encuestador sobre el terreno ha entrevistado a la persona incorrecta del hogar, el supervisor sobre el terreno volverá a enviar a dicho encuestador al hogar en cuestión para entrevistar a la persona correcta, y las respuestas deberán compararse de nuevo posteriormente. En todos los casos, al margen de que se haya detectado o no algún error en el proceso de verificación, la información de la entrevista de verificación habrá de transmitirse al centro de datos nacional junto con el resto del material.

### ***Asignación de códigos de resultado final de entrevistas no realizadas***

Los supervisores sobre el terreno deben aprobar el código de resultado final de entrevista no realizada antes de finalizar un cuestionario. (En el Manual del supervisor sobre el terreno de la encuesta GATS figura información adicional para intentar completar los casos pendientes correspondientes a entrevistas no realizadas). Si no es necesario adoptar más medidas con respecto a un caso, el encuestador sobre el terreno debe introducir en el sistema de gestión de casos el correspondiente código de resultado final de entrevista no realizada (véase el Manual del encuestador sobre el terreno de la encuesta GATS para consultar la lista y la descripción de los códigos de resultados sobre el terreno).

A medida que vaya concluyendo la recogida de datos en una zona determinada (por ejemplo, una unidad primaria de muestreo, una región, un estado, etcétera), deberá ir asignándose un código de resultado final a todos los casos en los que se haya trabajado. El programa asignará automáticamente a los ejemplares cumplimentados tanto del *Cuestionario para hogares* como del *Cuestionario individual* el código de resultado final «completado».

En el caso de aquellos ejemplares de tales cuestionarios que no hayan culminado en una entrevista completa (por ejemplo, en caso de negativa a responder o de que no hubiera nadie en casa), los encuestadores sobre el terreno, con la aprobación de sus supervisores, deberán asignar los códigos de resultado final de entrevista no realizada.

---

**A medida que vaya concluyendo la recogida de datos, deberá ir asignándose un código de resultado final a todos los casos en los que se haya trabajado.**

---

## 4.2 Gestión de datos del centro de datos nacional

En la presente sección se describen los procedimientos de recogida y gestión de datos por parte del centro de datos nacional, a saber: combinación/agregación de los archivos, aspectos del control de la calidad como el monitoreo y la presentación de informes, comunicación y apoyo técnico, y creación de archivos de datos brutos.

### 4.2.1 Agregación

A medida que los datos vayan llegando al punto de agregación de datos (servidor en la nube, servidor FTP, etcétera), los archivos deberán revisarse y procesarse y tendrán que realizarse las correspondientes actividades de control de la calidad. Asimismo, será necesario llevar a cabo ciertas actividades de documentación y archivo de los archivos de datos.

Durante el procesamiento, el software de agregación ordena y procesa los archivos dentro de la carpeta seleccionada (normalmente C:\GATSAndroid2\IncommingDB3s\...) y en orden alfabético e independientemente de la fecha. Es fundamental que los archivos se agreguen en el orden exportado. Al agregar todos los archivos exportados, la convención aplicable a la denominación de tales archivos (FI#\_YYYY\_MM\_DD.db3) garantiza una clasificación alfabética automática y el correcto procesamiento de los datos. Los archivos de exportación son acumulativos, por lo que el archivo más reciente contendrá todos los datos, incluidas las actualizaciones o ediciones existentes en archivos anteriores.

El proceso de agregación no permite la duplicación de identificadores de casos.

#### **Proceso de agregación**

- En la carpeta C:\GATSAndroid2\DATA existe un archivo maestro de agregación denominado master.db3zero que constituye el punto de partida para la recogida de datos. Este archivo es necesario al inicio de la recogida de datos para que la agregación comience con un archivo maestro que no contenga ningún dato. El archivo master.db3zero deberá copiarse y dicha copia pasará a denominarse master.db3 cuando sea necesaria la reinicialización.
- En la carpeta C:\GATSAndroid2\ debe encontrarse un archivo con el nombre Node.id.
  - Antes de que la agregación se ejecute por primera vez en la máquina, el archivo de nodo deberá editarse para que tenga un nombre distintivo y significativo.
  - Una vez ejecutada la agregación, el archivo «Node.ID» no deberá modificarse.

- En el caso de los países que apliquen el cifrado de datos, los últimos archivos db3 recopilados deberán descifrarse mediante la opción Agregación de datos > Descifrador de base de datos del software del sistema general de encuestas. Esta operación tendrá que realizarse antes de la agregación si se está utilizando el cifrado de datos.
- Agregue los últimos archivos db3 recopilados a los agregados anteriormente para garantizar la eficiencia, siempre y cuando los archivos se reciban y procesen en orden (por fecha de exportación). Si bien la agregación a un archivo maestro vacío cada vez que se ejecute la operación de agregar resulta admisible, hay que tener en cuenta que este procedimiento llevará más tiempo, ya que será necesario procesar todos los archivos db3 exportados.
- Guarde todos los archivos exportados recibidos archivándolos en una subcarpeta dentro de la carpeta C:\GATSAndroid2\l. No será necesario volver a agregar tales archivos.

Tanto durante la agregación como con posterioridad a esta, será preciso comprobar varios aspectos:

- Que todos los archivos recibidos tengan nombres de archivo que incluyan la fecha de exportación del archivo en cuestión, lo que evita el procesamiento de archivos no operativos. Si se reciben archivos no operativos (es decir, archivos de datos anteriores recibidos posteriormente), los datos más recientes referentes a un identificador de caso en particular se sobrescribirán con datos anteriores quizás incompletos. Para asegurarse de haber procesado los datos más recientes al realizar la agregación a un archivo master.db3 no vacío, consulte los archivos db3 archivados. (Agregue siempre los archivos por orden de fecha o simplemente utilice los archivos más recientes).
- La exactitud y exhaustividad de la agregación de datos: ¿se contabilizaron e incluyeron todos los archivos db3 previstos al ejecutar la agregación?
- ¿Se agregó el archivo db3 más reciente para cada encuestador sobre el terreno?
- El examen de los datos: deberá realizarse la correspondiente comprobación y validación. El analista deberá examinar los datos de eventos, así como los cuestionarios cumplimentados para detectar posibles problemas, patrones o datos incoherentes. También generará frecuencias por pregunta para su correspondiente análisis. Puede encontrarse información sobre la visualización de archivos DB3 en el *Plan de ejecución de la gestión de datos de la encuesta GATS*.

Revise el informe resumido y el informe de estado de la agregación generados a partir de esta para confirmar los recuentos.

El informe resumido de agregación de datos proporciona información sobre los resultados de la agregación. Cada fila muestra los resultados de una unidad de agregación (normalmente un encuestador sobre el terreno) y una tabla de base de datos específica. La tabla se indica en la columna AggTable, el número de filas de la tabla de entradas se muestra en la columna AggRowsTotal y el número de filas insertadas desde esa tabla se muestra en la última columna AggRowsInserted.

El informe de estado de la evaluación de hogares es un resumen del estado de todos los casos en el archivo de datos agregados. El número total de casos para cada unidad de agregación (normalmente un encuestador sobre el terreno) se indica en la columna de casos. A continuación, el total se desglosa entre las diversas categorías de estado. Las columnas de estado (entrevista concluida, sin realizar, pendiente y final) dependen del código de resultado más reciente asignado al caso en cuestión.

- A partir de los datos agregados o cualquier archivo db3, se pueden crear datos con formato de archivo plano o archivos CSV mediante las herramientas de agregación de datos incluidas en el conjunto de herramientas.
- Se recomienda revisar los datos de eventos (que figuran en los archivos DU y DuEvt) para conocer el contenido y entender las operaciones sobre el terreno.
- El archivo DU contiene solo un registro por cada identificador de caso (el *Cuestionario para hogares* y el *Cuestionario individual* tienen identificadores de caso diferenciados, -00 y -01 respectivamente) y el código de resultado más reciente para el identificador del caso en cuestión. El archivo DuEvt contiene el historial de los eventos en relación con un identificador de caso determinado introducido por los encuestadores sobre el terreno y puede contener más de un registro por identificador de caso.

### **Generación de archivos de datos para su revisión**

- Los datos del cuestionario (tabla de respuestas) se almacenan en una tabla en el caso de los formularios tanto del *Cuestionario para hogares* como del *Cuestionario individual*. Hay una fila por pregunta (identificador de pregunta), por lo que estos datos deben reestructurarse antes del análisis. Los datos de las respuestas deben transponerse para facilitar su revisión y utilización inmediata en el software estadístico. Con ello se creará un registro basado en el identificador de caso para los datos tanto del *Cuestionario para hogares* como del *Cuestionario individual*. El software del sistema general de encuestas transpone los datos de las respuestas basándose en los archivos db3 de la encuesta especificados mediante archivos de identificadores de casos diferenciados (archivos planos) para el *Cuestionario para hogares* y el *Cuestionario individual*.
- Los datos de eventos de la encuesta GATS se almacenan en la tabla DUEVT asignando una fila a cada resultado por cada identificador de caso. Los datos de la encuesta GATS relativos a la unidad de vivienda se almacenan en la tabla DU asignando una fila por identificador de caso. Los datos de DU también contienen una copia del código de resultado más reciente asignado a cada caso.

En la tabla que figura a continuación se resumen estas fuentes de datos del sistema del dispositivo portátil.

<b>Fuente de los datos de la tabla</b>	<b>Tabla</b>	<b>Filas</b>
Datos del cuestionario	Respuestas	Una fila por pregunta por cada identificador de caso y formulario
Datos de eventos	DUEVT	Una fila por resultado por cada identificador de caso y formulario
Unidad de vivienda	DU	Una fila por unidad de vivienda
Ediciones o modificaciones de dirección	Registro de direcciones	Una fila por edición de dirección
Notas del encuestador sobre el terreno	Notas	Una fila por tipo de nota (ya sea por identificador de caso o por pregunta (identificador de pregunta))

## 4.2.2 Control de la calidad de los datos recogidos

En el centro de datos nacional, deberán haberse establecido procedimientos de control de la calidad rutinarios para revisar y monitorear los datos a medida que vayan llegando y procesándose, así como para elaborar informes y copias de seguridad de estos. El centro de datos nacional ha de establecer un calendario de procesamiento diario como mínimo para la presentación de información y el monitoreo y otro que también abarque el proceso rutinario de copia de seguridad.

El protocolo normalizado de la encuesta GATS consta de una lista de variables indicadoras de la encuesta GATS y de recomendaciones de presentación de información a nivel nacional. La información y los detalles relativos a las variables indicadoras clave pueden consultarse en las *Definiciones de los indicadores de la encuesta GATS*. Los países podrán realizar las correspondientes adiciones o adaptaciones de las variables indicadoras en el cuestionario final aprobado.

---

**Se recomienda que las variables indicadoras clave se revisen a nivel de datos brutos durante la recogida de datos.**

---

Informes elaborados por el país: Se recomienda que el país elabore informes de estado de eventos por caso. El monitoreo de la encuesta (por ejemplo, los avances sobre el terreno, la recogida y transmisión de datos) constituye una actividad clave, coordinada y continua.

- Durante el período de recogida de datos, el centro de datos nacional debe presentar un informe semanal de la tasa de respuesta generado a partir del software del sistema general de encuestas. En dicho informe deberán indicarse las entrevistas concluidas y la tasa de respuesta general. En la *Guía del programador sobre el sistema general de encuestas de la encuesta GATS* se pueden encontrar ejemplos de trabajo sobre el terreno e informes de estado sobre la tasa de respuesta.
- Se alienta al centro de datos nacional a utilizar los informes generados a partir del software del sistema general de encuestas y, en función de la necesidad, a elaborar informes en Access<sup>1, 2</sup>, SPSS<sup>1, 3</sup>, SAS<sup>1, 4</sup>, STATA<sup>1, 5</sup> u otro software que ayude a revisar los datos en busca de incoherencias, anomalías o información incompleta. Los datos agregados transpuestos han de utilizarse para generar informes.

Las incoherencias en los datos se pueden identificar mediante la realización de controles de validación de las variables incluidas en el *Cuestionario para hogares* y el *Cuestionario individual*, tanto internamente en cada uno de ellos como entre ambos. Consúltense el *capítulo 5* sobre la garantía de la calidad en la fase posterior a la recogida de datos para obtener información sobre los controles de validación.

Procedimientos de copia de seguridad: Se deben realizar copias de seguridad de los datos de forma rutinaria, por ejemplo, semanalmente, dependiendo de la frecuencia de agregación.

- Los archivos db3 individuales (nivel de encuestador sobre el terreno) tendrán que archivarse.

---

<sup>1</sup> Los nombres comerciales se utilizan únicamente con fines de identificación y no implican el respaldo del Departamento de Salud y Servicios Humanos de los Estados Unidos o de los Centros para el Control y la Prevención de Enfermedades.

<sup>2</sup> Microsoft Office Access© (Microsoft Corporation, Redmond, Washington).

<sup>3</sup> SPSS© (SPSS Inc., Chicago, Illinois).

<sup>4</sup> SAS© (SAS Institute Inc., Cary, Carolina del Norte).

<sup>5</sup> STATA© (Stata Corp., College Station, Texas).

- También deben realizarse copias de seguridad de los datos archivados en un punto secundario, como una red o una memoria USB, y habrá de conservarse una copia en dos ubicaciones físicas independientes.
- Una vez finalizada la recogida de datos, deben cotejarse los datos recibidos con el archivo de selección de la muestra maestra, así como los archivos exportados desde el terreno.

#### **4.2.3 Comunicación y apoyo técnico**

La comunicación desde el centro de datos nacional durante la recogida de datos ha de conllevar el envío semanal de informes de estado a la OMS y los CDC y la notificación de todos y cada uno de los problemas técnicos detectados. Todos los problemas técnicos relacionados con los dispositivos portátiles y el software, estén o no resueltos, deben notificarse por correo electrónico lo antes posible en cuanto se produzcan. Ello ayudará a facilitar el apoyo técnico oportuno, la resolución de problemas y el intercambio de información.



## 5. Garantía de la calidad: fase posterior a la recogida de datos

---

La fase posterior a la recogida de datos abarca una serie de tareas necesarias para preparar un archivo de datos de análisis con miras al análisis de datos y tiene lugar en el momento en que ya se han recogido y agregado todos los datos de la encuesta. Esto engloba la preparación de los datos para los cálculos de ponderación de la muestra; la evaluación de la calidad del muestreo, el error de muestreo y las ponderaciones; y la medición de la calidad de la cobertura del marco, la falta de respuesta y otros errores ajenos al muestreo. Asimismo, este proceso incluye la preparación de un archivo de datos de análisis. En el presente capítulo se describen las directrices y procedimientos de garantía de la calidad aplicables y recomendables en cada una de estas actividades, abordando concretamente los siguientes aspectos:

- 5.1 Limpieza y preparación de los datos para los cálculos de ponderación de la muestra
  - 5.1.1 Creación del archivo de la base de datos maestra
  - 5.1.2 Eliminación de las variables identificativas confidenciales
  - 5.1.3 Limpieza y validación del archivo de datos fusionado
  - 5.1.4 Asignación de códigos de disposición finales
- 5.2 Medidas de calidad: muestreo, error de muestreo y ponderaciones
  - 5.2.1 Patrón de ajustes de calibración de las ponderaciones con posterioridad a la estratificación entre celdas de ajuste
  - 5.2.2 Efecto multiplicativo de las ponderaciones variables de la muestra en la precisión de las estimaciones de la encuesta
  - 5.2.3 Efecto general del diseño en la precisión de las estimaciones de la encuesta y la homogeneidad entre agrupamientos dentro de la unidad primaria de muestreo de las correspondientes estimaciones principales de la encuesta
  - 5.2.4 Margen de error de las estimaciones principales de la encuesta
- 5.3 Medidas de calidad: cobertura, falta de respuesta y otros errores ajenos al muestreo
  - 5.3.1 Tasa de cobertura del marco de hogares
  - 5.3.2 Patrones de las tasas de corte de los encuestados
  - 5.3.3 Patrones de las tasas de respuesta de los hogares por estratos de muestreo de la primera etapa
  - 5.3.4 Patrones de las tasas de respuesta de las personas entre las variables utilizadas para los ajustes por falta de respuesta
  - 5.3.5 Patrones de las tasas de rechazo de las personas entre las variables utilizadas para los ajustes por falta de respuesta
  - 5.3.6 Tasas de falta de respuesta a los ítems para las variables indicadoras de las notas descriptivas
- 5.4 Revisión formal de la calidad estadística
  - 5.4.1 Documentación de las ponderaciones de la muestra
- 5.5 Creación de un archivo de datos de análisis
- 5.6 Creación de archivos de datos de uso público

## 5.1 Limpieza y preparación de los datos para los cálculos de ponderación de la muestra

En esta sección se proporcionan directrices para la fusión de archivos de datos, la validación de las variables y los patrones de salto, así como la creación de los códigos de disposición utilizando un paquete de software estadístico<sup>1</sup>.

### 5.1.1 Creación del archivo de la base de datos maestra

La opción de generación de archivos de análisis combina los datos de la encuesta en un único archivo delimitado por comas y crea programas de entrada de SAS, SPSS y STATA capaces de leer el archivo de base de datos maestra fusionado. Para dicha creación puede utilizarse el software del sistema general de encuestas.

Entre los archivos utilizados por la herramienta de generación de archivos de análisis se incluyen el archivo db3 maestro agregado final creado mediante la agregación de los archivos db3 de todos los encuestadores y las dos bases de datos de los cuestionarios (Survey0.db3 y Survey1.db3).

Asegúrese de utilizar las ubicaciones de archivo actualizadas del archivo agregado y las bases de datos del cuestionario de la encuesta al hacer uso de la herramienta de generación de archivos de análisis. Para realizar modificaciones, actualice la ubicación o el nombre del archivo en el cuadro de texto o seleccione la opción «...» y el sistema le permitirá seleccionar un archivo existente. El archivo de la base de datos maestra será un archivo delimitado por comas en formato Unicode (UTF-8). La herramienta también generará un código de entrada para leer los datos en SAS, SPSS y STATA.

Para ver el archivo CSV, se requiere un editor de texto habilitado para Unicode, como WordPad. Los archivos SAS, SPSS y STATA también están en Unicode para garantizar que las etiquetas y los formatos se visualicen correctamente. Puede que haya que actualizar los programas SAS, SPSS y STATA antes de su ejecución para garantizar que se establezcan opciones específicas para cada país. El proceso daría lugar a la creación de un conjunto de datos maestro. El archivo de base de datos maestra generado de este modo debe utilizarse para eliminar cualquier variable identificativa confidencial y debe hacerse llegar al punto de contacto de TI en el país.

### 5.1.2 Eliminación de las variables identificativas confidenciales

En el conjunto de datos maestro generado, hay que eliminar los campos identificativos confidenciales de la sección de campos de selección de la muestra maestra.

### 5.1.3 Limpieza y validación del archivo de datos fusionado

Verifique que las variables tengan los valores adecuados y que los patrones de salto hayan funcionado correctamente y compruebe los campos en blanco. Si bien muchas de estas comprobaciones de la calidad de los datos se llevan a cabo durante la administración del cuestionario mediante el software del dispositivo portátil, es importante realizar nuevas comprobaciones por si hubiera errores no detectados en la programación del software de dicho dispositivo.

1. Compruebe que los patrones de salto de los datos hayan funcionado según lo especificado en el cuestionario. Los patrones originales de salto de los datos derivados del cuestionario final

---

<sup>1</sup> El centro coordinador de datos de la encuesta GATS en los CDC brinda asistencia técnica en el uso de los siguientes paquetes de software estadístico: SAS, SPSS y STATA.

aprobado para el país deben verificarse para todas las variables básicas, recomendándose comprobar también los patrones para el resto de las variables.

- Un encuestado que no haya respondido a un ítem debido a un patrón de salto tendrá un campo en blanco para ese elemento omitido.
  - Por ejemplo, si un encuestado se ha negado a responder a la pregunta A02a (el mes de nacimiento del encuestado) o A02b (el año de nacimiento del encuestado), deberá haber respondido a la pregunta A03 (la edad del encuestado), pero si ha respondido a las preguntas A02a y A02b, entonces A03 deberá haberse omitido.
  - Por ejemplo, si un encuestado responde a la pregunta B01 que fuma «todos los días» (valor 1), avanzaría a la pregunta B04 y no deberían figurar respuestas a las preguntas B02 y B03.
- En SAS, un campo en blanco para una variable numérica es igual a '.' (para las variables de caracteres es ' '). En dicho software, este patrón de salto podría verificarse utilizando el código que figura a continuación. Cualquier registro que no siga el patrón de salto se volcará a un conjunto de datos de SAS.

```
SI B01 = 1 & (B02 NE . O B03 NE .) ENTONCES BANDERA=1;  
O SI B01 = 2 & (B02 = . O B03 = .) ENTONCES BANDERA =1;  
O SI B01 = 3& (B02 NE . O B03 = .) ENTONCES BANDERA =1;  
EN CUALQUIER OTRO CASO, BANDERA=0;  
SI BANDERA=1 ENTONCES SALIDA; *Un valor de 1 para BANDERA indica que el patrón  
de salto no ha funcionado y un valor de 0 indica que sí lo ha hecho;
```

2. Compruebe que los únicos campos en blanco sean los derivados de un patrón de salto (por ejemplo, no son aplicables para el encuestado y deben estar en blanco). Cualquier otro valor en blanco no será válido.

- Los campos que están en blanco debido a los patrones de salto deberán convertirse en un nuevo valor. Esto los diferenciará de los campos en blanco que no sean atribuibles a patrones de salto (campos en blanco no válidos). Dado que los valores de las variables originales no deben sobrescribirse, será preciso crear un nuevo conjunto de datos para concluir este paso.
- En SAS, utilice el código «.S» para los valores faltantes cuando las correspondientes preguntas no fueran aplicables (omisión válida). El código de SAS para el patrón de salto utilizado anteriormente podría modificarse para indicar lo siguiente:

```
SI B01 = 1 & (B02 = . O B03 = .) ENTONCES B02 = .S & B03 =.S;  
EN CUALQUIER OTRO CASO, SI B01 =3 & B02 = . ENTONCES B02 = .S;
```

- Una vez sustituidos todos los valores de los patrones de salto, no debería haber otros campos en blanco. Si existen otros campos en blanco, vuelque esos registros en el archivo de error.
3. Compruebe cada variable para asegurarse de que los valores no válidos no estén presentes. Utilice el diccionario de datos del país (libro de códigos) para consultar los valores válidos de cada variable del conjunto de datos.
- Compruebe que los valores de las respuestas a cada pregunta sean válidos.

- Ejemplo: Las únicas respuestas válidas para B01 son: 1 («todos los días»), 2 («no todos los días»), 3 («nunca en absoluto»), -9 («no contesta»), o el código de valor faltante específico del software utilizado (por ejemplo, .S cuando se utiliza SAS).
- Ejemplo: Algunas variables llevan intervalos asociados, por lo que no debería existir un valor por debajo o por encima de los límites del intervalo en cuestión. Por ello, para la variable B04, el valor mínimo posible debería ser 0 y el valor máximo posible debería ser 98. También existe la opción de que el valor sea -9 si la respuesta ha sido «No sabe» o «No contesta». El código de valor faltante específico del software utilizado (por ejemplo, .S cuando se usa SAS) también sería un valor válido.
  - *En SAS, los registros con respuestas no válidas para B01 se podrían seleccionar y volcar en el archivo de error mediante el siguiente código:*  
 SI B01 NO ESTÁ EN(1,2,3,-7,-9) ENTONCES SALIDA;
  - *Si el analista lo prefiere, como alternativa, se podrían ejecutar frecuencias simples en cada una de las variables para ver si los valores están fuera del intervalo. Entonces cualquier registro con valores no válidos podría identificarse y volcarse al archivo de error.*
- Verifique la edad del encuestado
  - Antes de verificar que la edad del encuestado sea la correcta, primero hay que calcular e incluir la edad en el conjunto de datos. Para hacer el cálculo, utilice la fecha de la entrevista del *Cuestionario individual* (IQ\_EVENTDATE), así como el mes (A02a) y el año (A02b) de nacimiento de los encuestados, suponiendo que el día sea el día 15 del mes. Si los encuestados responden «No sabe» o «No contesta» a A02a o A02b, entonces la encuesta avanza a la pregunta A03 y se les pide que indiquen su edad (sin embargo, A03 solo debe responderse si no se facilita información en A02a o A02b). Si no se indica el mes o año de nacimiento, use el valor de A03 como edad del encuestado. Si se indica el mes y el año, calcule la edad restando la fecha de nacimiento de la fecha de la entrevista. Verifique que la edad del encuestado sea válida (el intervalo válido es de 15 a 125).
    - *En SAS, la edad se puede calcular y verificar posteriormente mediante el siguiente código:*  
 SI A02A <=12 Y A02B NE -7 Y A02B NE -9 ENTONCES REALIZAR LA OPERACIÓN;  
 FECHA DE NACIMIENTO = MDY(A02A, 15, A02B);  
 EDAD = MÍNIMO(YRDIF(FECHA DE NACIMIENTO, IQ\_EVENTDATE, 'ACTUAL'));  
 FIN;  
 EN CUALQUIER OTRO CASO, EDAD = A03;

#### 4. Otras comprobaciones de validación de datos:

- Realice comprobaciones para cerciorarse de que, en caso de no haberse omitido ninguna de las variables indicadas a continuación, sus valores sean inferiores a la edad del encuestado: B04, B05, B08, B09, B11, B12, B13, C04, C05, C08, C09, C11, C12 y C13 (dependiendo de la escala métrica de la variable, podría ser necesario convertir primero en años estos valores de variable).

#### 5.1.4 Asignación de códigos de disposición finales

En el archivo de datos fusionado se incluirán los códigos de resultado final asignados por los encuestadores sobre el terreno a todos los casos de la muestra en los que se haya trabajado durante la recogida de datos (véase el *Manual del encuestador sobre el terreno de la encuesta GATS* para consultar la descripción de los códigos de resultados sobre el terreno). Cada caso de entrevista a hogares y cada caso de entrevista individual deben contar con un código de resultado final (véase la *sección 4.1.3* para la asignación de códigos de resultado final de entrevista no realizada durante la recogida de datos).

---

**Deben asignarse códigos de disposición finales a cada caso. Los códigos de disposición finales se utilizan para calcular e indicar las tasas de respuesta y las medidas de garantía de la calidad.**

---

La utilización de los códigos de resultado final conlleva la necesidad de asignar códigos de disposición finales. Posteriormente, los códigos de disposición finales se utilizan para calcular e indicar las tasas de respuesta y las medidas de garantía de la calidad. Consúltese el *apéndice C.1* para la asignación de los códigos de disposición finales a partir de los códigos de resultado final, así como para el cálculo de las tasas de respuesta y la utilización de una plantilla para informar de estos resultados de la recogida de datos.

A continuación se ofrecen unas directrices importantes de necesaria aplicación para este proceso:

1. A cada caso se le debe asignar un código de disposición final referente al *Cuestionario para hogares* en función del código de resultado final sobre el terreno de dicho cuestionario. Si se ha concluido la entrevista del *Cuestionario para hogares* y se ha seleccionado a una persona para la entrevista del *Cuestionario individual* (código de disposición final 1), también se deberá asignar un código de disposición final referente al *Cuestionario individual*. Si el código de disposición final del *Cuestionario para hogares* es distinto de 1 (por ejemplo, no se ha seleccionado a nadie, negativa a responder, cuestionario incompleto), entonces no existirá un código de disposición final del *Cuestionario individual*, ya que la entrevista con este no se habrá realizado.
2. Existe una serie de preguntas clave en el *Cuestionario individual* que determinan la legitimidad de la entrevista. Si la respuesta registrada a cualquiera de estas preguntas clave es «No sabe» o «No contesta», la entrevista no se considera legítima. De ahí que, en todos los casos del *Cuestionario individual* codificados con un código de resultado final de 400 (*Cuestionario individual* cumplimentado), se deban seguir las normas indicadas a continuación para determinar qué código de disposición ha de asignarse.
  - Si las respuestas a B01, B02 o B03 (preguntas clave sobre la prevalencia del tabaquismo) se han registrado como «No sabe» o «No contesta», es preciso asignar un código de disposición final de 16 (el encuestado seleccionado no es competente). Esto no debería ser algo habitual. (Esta norma también será aplicable a C01, C02 y C03 (preguntas clave sobre la prevalencia del consumo de tabaco sin humo) únicamente si el consumo de tabaco sin humo constituye una práctica habitual en el país).
  - En cuanto al resto de los casos del *Cuestionario individual* codificados con un código de resultado final de 400 (*Cuestionario individual* cumplimentado), es preciso asignar un código de disposición final de 11 (*Cuestionario individual* cumplimentado). Esto será lo habitual en la mayoría de los casos.

3. En todos los casos del *Cuestionario individual* codificados con un código de resultado final de 402 (*Cuestionario individual* parcialmente cumplimentado), se deben seguir las normas indicadas a continuación para determinar qué código de disposición debe asignarse.
  - Si el *Cuestionario individual* se cumplimentó al menos hasta la pregunta E01 Y en ninguna de las preguntas clave (indicadas anteriormente en el punto #2) se respondió «No sabe» o «No contesta», es necesario asignar un código de disposición final de 11 (*Cuestionario individual* cumplimentado).
  - Si el *Cuestionario individual* se interrumpió antes de llegar a la pregunta E01, se debe asignar un código de disposición final de 12 (incompleto). (Los casos con un código de disposición final de 12 (incompleto) se consideran referentes a personas que no respondieron a la encuesta GATS y los correspondientes datos no se incluyen en el análisis).
  - Si el *Cuestionario individual* se cumplimentó al menos hasta la pregunta E01 pero se respondió a cualquiera de las preguntas clave (indicadas anteriormente en el punto #2) con «No sabe» o «No contesta», es preciso asignar un código de disposición final de 16 (el encuestado seleccionado no es competente). (Como se ha señalado anteriormente en el punto #2, se determina que este tipo de encuestados no son competentes si no pueden proporcionar respuestas legítimas a las preguntas clave y, por consiguiente, los datos de su entrevista no se incluyen en el análisis).
4. Solo los casos con un código de disposición de 11 deben incluirse en el conjunto final de datos de análisis. (De ahí la importancia de asignar los códigos de disposición correctamente).
5. Utilice tabulaciones cruzadas para cotejar todos los códigos de resultados finales con los códigos de disposición a fin de detectar posibles errores de clasificación. Si no se diera la debida correspondencia entre ambos códigos, sería un indicio de la existencia de un problema con el código del software utilizado para crear los códigos de disposición.

## 5.2 Medidas de calidad: muestreo, error de muestreo y ponderaciones

En la presente sección se describen los cálculos necesarios para evaluar directamente la calidad de las estimaciones realizadas a partir de las muestras de la encuesta GATS e indicar los efectos de los agrupamientos de muestras y las ponderaciones desiguales en esas estimaciones. También se incluyen directrices encaminadas a comprobar la exactitud de las ponderaciones calculadas.

### 5.2.1 Patrón de los ajustes de calibración de las ponderaciones con posterioridad a la estratificación entre celdas de ajuste

El último paso en la generación de las ponderaciones de la muestra conlleva la calibración de tales ponderaciones con los recuentos de la población mediante correlaciones conocidas de mediciones clave de los resultados del estudio, denominadas variables de calibración (por ejemplo, género, educación, edad, entorno urbano o rural y región, como se sugiere en el *Manual de ponderación de la muestra de la encuesta GATS*).

El contexto general y las instrucciones para calcular los ajustes con posterioridad a la estratificación se facilitan más adelante en esta sección. En el *apéndice C.2* figura información más detallada sobre estos temas.

1. **Contexto:** La calibración se ajusta para el desequilibrio de la muestra no contemplado en el ajuste por falta de respuesta. Se aplican ajustes por separado a todos los miembros de las celdas de ajuste creadas estratégicamente. El objetivo es otorgar más peso a los subgrupos de población que permanecen infrarrepresentados y disminuir el de aquellos que siguen estando sobrerrepresentados. Cuanto más se alejen los valores de los ajustes de calibración de 1,00 (ya sea por encima o por debajo), mayor será la influencia potencial del desequilibrio de la muestra (más allá de lo contemplado en el ajuste por falta de respuesta) en el sesgo de las estimaciones de la encuesta.
2. **Generación de ajustes con posterioridad a la estratificación:** Desde el punto de vista procedimental, la calibración con posterioridad a la estratificación implica la creación de «celdas de ajuste» mediante la clasificación cruzada de las mediciones de correlación. El «ajuste con posterioridad a la estratificación» en cada una de estas celdas de ajuste será  $<1$  si la submuestra de esa categoría está sobrerrepresentada en la muestra una vez tenida en cuenta la falta de respuesta y la selección de la muestra, y el «ajuste con posterioridad a la estratificación» será  $>1$  en aquellas categorías en las que la muestra esté infrarrepresentada.
3. **Presentación de información sobre los ajustes con posterioridad a la estratificación:** Cree una tabla en la que figuren todas las celdas de ajuste, indicando en cada caso la definición de las celdas mediante las variables categóricas utilizadas para la calibración. Para cada celda, registre el valor reconstruido del ajuste con posterioridad a la estratificación y observe su tamaño relativo con respecto a 1,00. El resultado óptimo es que todos los ajustes con posterioridad a la estratificación estén cerca de 1, con valores ligeramente superiores a 1 en algunos casos y situándose el resto levemente por debajo de 1,00.

### 5.2.2 Efecto multiplicativo de las ponderaciones variables de la muestra en la precisión de las estimaciones de la encuesta

En el *Manual de diseño de la muestra de la encuesta GATS* se exige aplicar un diseño en el que las probabilidades de selección (y, por consiguiente, las ponderaciones de la muestra) variarán ligeramente debido al uso de mediciones estimadas de los tamaños de los agrupamientos, los ajustes a las ponderaciones de la muestra y la asignación equitativa de los tamaños de muestra entre las regiones cuando se requieran estimaciones regionales que cumplan con las normas de precisión de la encuesta GATS. En el *Manual de ponderación de la muestra de la encuesta GATS* se describe el modo de calcular dichas ponderaciones. Una vez se hayan limpiado los datos del cuestionario y se hayan adjuntado las ponderaciones finales de la muestra, los datos de la muestra de la encuesta GATS estarán listos para el análisis y la presente tarea.

El contexto general y las instrucciones para el cálculo de este efecto se facilitan más adelante en esta sección. En el *apéndice C.3* figura información más detallada sobre estos temas.

1. **Contexto:** La variación de las ponderaciones de la muestra puede aumentar la cantidad de error de muestreo en las estimaciones de la encuesta y, por lo tanto, dar lugar a mayores varianzas y errores cuadráticos medios de tales estimaciones. Más concretamente, el aumento multiplicativo de la varianza de las estimaciones de la encuesta, denominado  $Meff_w$ , depende de la variabilidad de las ponderaciones relativas al conjunto de observaciones de la muestra que se utilizan para generar la estimación.

---

**Cuanto más ponderaciones variables haya, mayor será el valor de  $Meff_w$ .**

---

2. **Estimación de  $Meff_w$** : Para calcular  $Meff_w$ , se necesitan el promedio simple y la varianza de las ponderaciones relativas al conjunto de observaciones de la muestra utilizadas para generar las estimaciones de la encuesta.
  - El valor de  $Meff_w$  para las estimaciones derivadas de este conjunto de observaciones de la muestra se obtiene calculando primero la proporción entre la varianza y el cuadrado de la media y sumando después uno a dicha proporción. La fórmula de cálculo específica es la ecuación (1) que figura en el *apéndice C.3*.
  
3. **Presentación de información sobre  $Meff_w$** : Dado que  $Meff_w$  se aplica a todas las estimaciones derivadas del conjunto específico de observaciones de la muestra que definen un ámbito de presentación de información sobre las estimaciones de la encuesta (por ejemplo,  $Meff_w$  en relación con la población general, la población rural, la educación, la edad, el género, etcétera),
  - debe calcularse e indicarse para todos los subgrupos principales de población en relación con los cuales se calculen e indiquen las estimaciones de la encuesta GATS. Esto podría hacerse en una tabla en la que se incluyera una lista de los subgrupos informadores con sus correspondientes valores de  $Meff_w$ .
  
4. **Interpretación de  $Meff_w$** : Un ejemplo de interpretación adecuada del valor de  $Meff_w$  calculado para un ámbito de presentación de información es el siguiente:
 

---

**Es preferible que  $Meff_w < 2,00$  .**

---

  - «La variación de las ponderaciones de la muestra multiplicó la varianza de todas las estimaciones (del ámbito de presentación de información) por ( $Meff_w$ )».
  - Si  $Meff_w < 2,00$  , no será necesario hacer nada por reducir la variación de las ponderaciones.
  
5. **Ejemplo**: Supongamos que, para las estimaciones relativas al entorno rural,  $Meff_w = 1,62$ 

«La variación de las ponderaciones de la muestra multiplicó la varianza de todas las estimaciones de los encuestados residentes en zonas rurales por 1,62».

  - $Meff_w = 1,62$  se considera un valor aceptablemente bajo para estas estimaciones.

### 5.2.3 Efecto general del diseño en la precisión de las estimaciones de la encuesta y la homogeneidad entre agrupamientos dentro de la unidad primaria de muestreo de las correspondientes estimaciones principales de la encuesta

En el *Manual de diseño de la muestra de la encuesta GATS* se exige aplicar un diseño en el que se seleccionen agrupamientos de encuestados y en el *Manual de ponderación de la muestra de la encuesta GATS* se describe el modo de calcular las ponderaciones de la muestra. Una vez que se hayan limpiado los datos del cuestionario y se hayan calculado las ponderaciones finales de la muestra, los datos de la muestra de la encuesta GATS estarán listos para su revisión con anterioridad a la notificación de las conclusiones del análisis.

El contexto general y las instrucciones para calcular el efecto general del diseño en una estimación de la encuesta y la homogeneidad entre agrupamientos dentro de la unidad primaria de muestreo se facilitan más adelante en esta sección. En el *apéndice C.4* figura información más detallada sobre estos temas.

1. **Contexto:** La varianza de una estimación de encuesta ( $\hat{\theta}$ ) calculada a partir de un diseño de muestra complejo dividida entre la varianza de una estimación comparable basada en una muestra aleatoria simple del mismo tamaño se denomina «efecto del diseño» general para ( $\hat{\theta}$ ), o  $Deff(\hat{\theta})$ . Hay dos componentes multiplicativos en  $Deff(\hat{\theta})$ : el efecto del diseño debido al muestreo por conglomerados,  $Deff_{CS}(\hat{\theta})$ , y el efecto multiplicativo de las ponderaciones variables de la muestra,  $Meff_w$  (véase la sección 5.2.2). Además,  $Deff_{CS}(\hat{\theta})$  guarda relación directa con la cantidad de homogeneidad dentro de la unidad primaria de muestreo,  $\hat{\rho}_\theta$ , y el tamaño promedio del agrupamiento de la muestra.
2. **Estimación de  $Deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ :** Se puede solicitar una estimación de  $Deff(\hat{\theta})$  directamente con algunos paquetes de software de análisis de encuestas (por ejemplo, SUDAAN), o bien se puede calcular este parámetro a partir de la estimación y su varianza cuando la estimación sea una proporción o tasa. Dado que  $Deff(\hat{\theta})$  depende de  $Deff_{CS}(\hat{\theta})$ ,  $\hat{\rho}_\theta$ , y del tamaño promedio del agrupamiento de la muestra, se puede obtener una estimación de  $\hat{\rho}_\theta$  resolviendo la correspondiente ecuación y utilizando las estimaciones proporcionadas de  $Deff(\hat{\theta})$  y estas otras medidas.
  - Las fórmulas detalladas para estimar  $Deff(\hat{\theta})$ , cuando los paquetes de software de análisis de encuestas no faciliten este parámetro, así como para estimar  $\hat{\rho}_\theta$  en cualquier circunstancia, pueden consultarse en las ecuaciones (1) y (2) que figuran en el *apéndice C.4*. Las estimaciones resultantes se indican mediante  $deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ , respectivamente. Usamos  $deff(\hat{\theta})$  para indicar la estimación de  $Deff_{CS}(\hat{\theta})$ .
  - Dado que habrá múltiples valores de  $deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ , será necesario resumirlos indicando la mediana y los valores mínimo y máximo.
3. **Presentación de información sobre  $Deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ :** Las estimaciones de  $Deff(\hat{\theta})$  y  $\hat{\rho}_\theta$  deberán indicarse para todas las mediciones clave del estudio (véase el *apéndice C.4* para consultar las correspondientes definiciones).
4. **Interpretación de  $Deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ :** Un ejemplo de interpretación adecuada de los valores estimados,  $deff(\hat{\theta})$  y  $\hat{\rho}_\theta$ , es el siguiente:
  - «La varianza de la estimación de la encuesta (con respecto a la característica de la población,  $\theta$ ), dado el diseño de la muestra de la encuesta GATS, ( $\hat{\theta}$ ) veces mayor que si el muestreo aleatorio simple hubiera generado el mismo número de encuestados».
  - Es preferible que el efecto general del diseño sobre una estimación, es decir,  $Deff(\hat{\theta})$ , sea inferior a 2,50.
  - La cantidad de variación total en la medición del estudio (vinculada a  $\theta$ ) que no se tiene en cuenta dentro de las unidades primarias de muestreo es  $\hat{\rho}_\theta$ .
  - Rara vez ocurre que  $\hat{\rho}_\theta > 0,15$  o sea un valor negativo; es más habitual que  $\hat{\rho}_\theta < 0,05$  o incluso que  $\hat{\rho}_\theta < 0,01$ .
5. **Ejemplo:** En el análisis para estimar la tasa actual de prevalencia del tabaquismo,  $P$ , se han obtenido los

---

Rara vez ocurre que  $\hat{\rho}_\theta > 0,15$ ; es más habitual que  $\hat{\rho}_\theta < 0,05$  o incluso que  $\hat{\rho}_\theta < 0,01$ .

siguientes valores a partir de una muestra en la que el tamaño promedio de los agrupamientos era de 20 elementos:

$$deff(\hat{\theta}) = 1.92, \quad Meff_w = 1.36.$$

«La varianza de la estimación de la encuesta de la tasa actual de prevalencia del tabaquismo, dado el diseño de la muestra de la encuesta GATS, es 1,92 veces mayor que si el muestreo aleatorio simple hubiera producido el mismo número de encuestados».

- El valor calculado de  $\hat{\rho}_\theta = 0,022$  constituye un valor típico.

#### 5.2.4 Margen de error de las estimaciones principales de la encuesta

Existen varias maneras de indicar la precisión estadística de las estimaciones de la encuesta. El margen de error de las estimaciones constituye una de ellas. La encuesta GATS recomienda que se indique el margen de error estimado ( MOE ) junto con las estimaciones de las principales mediciones de la encuesta. En el *Manual de diseño de la muestra de la encuesta GATS* se describen las tres características principales de las muestras de la encuesta GATS que influyen en la calidad estadística de las estimaciones y las conclusiones extraídas de los datos. Estas características son el uso de la estratificación, la selección de agrupamientos de miembros de la población a la que está dirigida la encuesta y la selección de miembros de la población con probabilidades desiguales (lo que requiere el uso de ponderaciones de la muestra en el análisis). Se pueden utilizar varios paquetes de software de análisis de encuestas para incorporar estas características en el análisis de los datos de la encuesta GATS.

El contexto general y las instrucciones de cálculo se facilitan más adelante en esta sección. En el *apéndice C.5* figura información más detallada sobre estos temas.

1. **El margen de error es específico de cada medición:** Cada estimación obtenida a partir de los datos de la encuesta GATS tiene su propio margen de error. Por definición, el margen de error es la mitad de la amplitud prevista del intervalo de confianza de una estimación de alguna de las principales mediciones de la encuesta. Otra forma de interpretar el margen de error es plantearlo como la probabilidad de que la estimación se aproxime a la medición real de la encuesta en la población.
2. **Principales mediciones de la encuesta:** En el marco de la encuesta GATS, se elaborarán estimaciones de la variación de una ronda a otra, así como estimaciones específicas de cada ronda, en relación con diversas tasas o hábitos de consumo de tabaco. Es probable que estas estimaciones se indiquen con respecto a la población en su conjunto, además de atendiendo a los diferentes subgrupos informadores de dicha población, por ejemplo, por género, residencia en entorno urbano o rural y categoría de edad. Es posible presentar estimaciones regionales cuando el tamaño total de la muestra regional sea de al menos 8000 elementos.

Las tres tasas que se indican a continuación se han considerado las más importantes a efectos de planificación:

- *Tasa de prevalencia del consumo de tabaco:* Porcentaje de personas que consumen actualmente productos de tabaco;
- *Tasa de tabaquismo pasivo:* Porcentaje de personas expuestas al humo generado por otros fumadores; y

- **Tasa de abandono:** Porcentaje de fumadores actuales que han intentado dejar de fumar.
3. **Estimación del margen de error:** La estimación del margen de error no suele calcularse mediante un software de análisis de encuestas, sino que normalmente se dispone de la información necesaria para efectuar dicho cálculo. Para calcular e interpretar el  $\dot{M}OE$  hacen falta tres cosas: i) la estimación de la medición de la encuesta, ii) su error cuadrático medio estimado (o la varianza) y iii) una medición específica acompañada del correspondiente nivel deseado de confianza estadística en el valor del  $\dot{M}OE$

El nivel de confianza generalmente se basa en un valor (Z) de la distribución normal típica. Por ejemplo, para un nivel de confianza del 95%, podemos usar  $Z=1,96$ .

El margen de error se calcula como el producto de la medición de confianza deseada y el error cuadrático medio de la estimación. La fórmula de cálculo propiamente dicha es la ecuación (1) que figura en el *apéndice C.5*.

4. **Presentación de información sobre el  $\dot{M}OE$  :** Se recomienda que el  $\dot{M}OE$  se calcule para todas las estimaciones relacionadas con las principales mediciones de la encuesta. Ello incluye las estimaciones nacionales generales de tales mediciones, así como las estimaciones de las mediciones en cuestión relativas a todos los subgrupos informadores importantes (por ejemplo, por género, residencia en entorno urbano o rural, región y edad).
- Se debe calcular una estimación del margen de error para todas las estimaciones relacionadas con las principales mediciones de la encuesta.**
- Es preciso presentar conjuntamente las principales estimaciones de la encuesta y sus correspondientes valores de  $\dot{M}OE$  . En el *apéndice C.6* se ofrece información sobre la presentación de estimaciones de los errores de muestreo y se facilitan ejemplos de estructuración de las tablas con los indicadores clave de la encuesta GATS y los valores que se recomienda indicar.

5. **Interpretación:** Cuando se interpreta conjuntamente con el valor de una estimación de la encuesta, el  $\dot{M}OE$  indica la probabilidad de que la estimación se aproxime al valor real registrado entre la población.

Si se utiliza  $Z=1,96$  para calcular el  $\dot{M}OE$  , la estimación de la encuesta y su correspondiente margen de error se pueden interpretar de la siguiente manera:

«Estamos seguros al 95% de que la estimación, (VALOR DE LA ESTIMACIÓN), se sitúa dentro de un margen de error de (VALOR DE SU  $\dot{M}OE$  ) con respecto al correspondiente valor entre la población».

6. **Ejemplo:** Supongamos que el valor indicado de una estimación de encuesta de la encuesta GATS es del 43,4%, con un error cuadrático medio del 1,5611% calculado de conformidad con el diseño de la muestra real en el país en cuestión.

El valor del  $\dot{M}OE$  para esta estimación será  $3,1\% \left( 1,96 \right) (1,5611\%)$

Interpretación:

«Estamos seguros al 95% de que la estimación del 43,4% se sitúa dentro de un margen de error del 3,1% con respecto al correspondiente valor entre la población».

### 5.3 Medidas de calidad: cobertura, falta de respuesta y otros errores ajenos al muestreo

#### 5.3.1 Tasa de cobertura del marco de hogares

Como se describe en el *Manual de diseño de la muestra de la encuesta GATS*, la penúltima fase del muestreo de la encuesta GATS consiste en seleccionar una muestra de hogares entre los enumerados en las unidades primarias de muestreo o segmentos. Por consiguiente, la calidad de la muestra de hogares y, en última instancia, de las estimaciones de la encuesta GATS, depende de la exhaustividad de estas listas de hogares, que se confeccionan de diversas maneras.

El contexto general y las instrucciones para calcular la tasa de cobertura del marco de hogares se facilitan más adelante en esta sección. En el *apéndice C.7* figura información más detallada sobre estos temas.

1. **Contexto:** Una parte fundamental del proceso de muestreo en una encuesta de hogares es el conteo de los hogares en la fase de muestreo en la que se eligen las direcciones de estos. Se pueden utilizar diferentes metodologías para el conteo de hogares (por ejemplo, el uso de un listado tradicional de hogares elaborado por censistas sobre el terreno o de censos electorales modificados). Lo habitual es que una parte neta de los hogares quede excluida del conteo resultante y que, debido a ello, la cobertura de la muestra final de hogares sea incompleta. Dado que el sesgo provocado por la infracobertura depende tanto de i) la diferencia agregada entre las principales mediciones del estudio relativas a los hogares abarcados por el marco y los que quedan fuera de este como de ii) el alcance de la falta de cobertura (es decir, el 100% menos la tasa de cobertura de los hogares) en la muestra, resulta útil estimar la tasa de cobertura del marco aplicable a los hogares.
  - La tasa de cobertura del marco se define como el número de hogares contabilizados en el marco de hogares dividido entre el número de hogares que componen la población.
  - Pese a su utilidad para describir la cobertura del marco de hogares de la muestra, a menudo no se dispone de datos extraídos de un censo reciente sobre el número total de hogares que componen la población. Si se dispone de esta información, deberá calcularse la tasa de cobertura del marco, si bien no constituye un elemento de datos imprescindible para la revisión oficial de las medidas de calidad estadística de la encuesta GATS.
2. **Estimación de la tasa de cobertura del marco:** El numerador de la tasa de cobertura del marco se puede estimar mediante la suma de las ponderaciones de preajuste (básicas) en todos los hogares seleccionados. El denominador debe obtenerse de alguna fuente actual que goce de gran credibilidad (por ejemplo, un censo reciente o alguna otra encuesta reciente con un tamaño de muestra superior al de la encuesta GATS).
3. **Presentación de información sobre la tasa de cobertura del marco:** La tasa de cobertura del marco estimada debe indicarse en el *Informe del país de la encuesta GATS* junto con la descripción del diseño de la muestra o con una evaluación aparte de la calidad de los datos.
4. **Interpretación de la tasa de cobertura del marco:** La tasa de cobertura del marco indica el alcance de la cobertura del marco de hogares, pero no el tamaño del sesgo de cobertura, ya que

este último depende de la suma de aquel más la diferencia agregada (en las principales mediciones de la encuesta) entre los hogares abarcados y no abarcados. Una tasa de cobertura de hogares superior al 95% se consideraría aceptable.

5. **Ejemplo:** La mejor estimación externa disponible del número de hogares que componen la población a la que está dirigida la encuesta de un determinado país de la encuesta GATS es de 36 436 103, mientras que la suma ponderada de los 13 000 hogares de la muestra seleccionada para la encuesta GATS es de 34 788 543. La tasa de cobertura resultante de  $0,955 = \frac{34\,788\,543}{36\,436\,103}$  se consideraría aceptable.
- 
- Una tasa de cobertura de hogares superior al 95% se consideraría aceptable.**
- 

### 5.3.2 Patrones de las tasas de corte de los encuestados

Algunas entrevistas de la encuesta no se concluyen en su totalidad. Al encuestado puede resultarle pesada la participación en la entrevista y, por lo tanto, puede perder la paciencia con el encuestador, o bien el encuestado puede necesitar terminar la entrevista antes de tiempo debido a limitaciones legítimas de calendario y no poder o no estar dispuesto a reprogramar otra cita para finalizar la entrevista. En cualquier supuesto, una entrevista concluida parcialmente es un indicador de la desafección del encuestado, que puede considerarse un reflejo de la calidad de los esfuerzos del encuestador o las actitudes de los encuestados hacia el tema de la encuesta y, por consiguiente, en última instancia, de la calidad de los datos.

- **Fuente(s) de datos:** Para estos cálculos, debe utilizarse el archivo de datos con los códigos de disposición finales relativos a los hogares/las personas (véase el *apéndice C.1*). Un código de disposición de 11 en el *Cuestionario individual* indica que el encuestado finalizó al menos las secciones A a D de la entrevista individual de la encuesta GATS. Un código de disposición de 12 en el *Cuestionario individual* indica que la entrevista se interrumpió antes de que se pudieran terminar las secciones A a D.
- **Método/cálculo:** La definición de «encuestado» hace referencia a cualquier persona seleccionada a la que se le haya asignado un código de disposición final de 11 o 12 en el *Cuestionario individual*. Asimismo, la tasa de corte queda definida de la siguiente manera:

$$\text{Tasa de corte} = \frac{R_c}{R} = \frac{\text{núm. de encuestados de la encuesta GATS que solo completaron parcialmente la entrevista}}{\text{núm. de encuestados de la encuesta GATS}}$$

o bien

$$= \frac{\text{Código de disposición 12}}{\text{Códigos de disposición (11 + 12)}}$$

- **Usos:** Los valores de la tasa de corte deben calcularse directamente para la muestra en la fase final del proceso de garantía de la calidad, una vez finalizada la recogida de datos. Además, los supervisores sobre el terreno han de esforzarse constantemente por identificar a aquellos encuestadores sobre el terreno que resulten ineficaces en el transcurso de la recogida de datos y tomar medidas correctivas al respecto (utilizando los correspondientes códigos de resultados sobre

el terreno para los cálculos). Esto conllevaría un cálculo de la tasa de corte por cada encuestador o equipo de encuestadores, así como la detección de aquellos encuestadores o equipos de encuestadores cuya tasa sea excepcionalmente alta (por ejemplo, una tasa superior a la media de la tasa registrada entre todos los encuestadores o equipos más dos desviaciones típicas). Los valores de la tasa de corte también pueden calcularse en función de los siguientes criterios: 1) la semana de recogida de datos en la que tuvo lugar la entrevista, 2) la edad del encuestado, 3) el género del encuestado y 4) la ubicación del encuestado en un entorno urbano o rural.

- **Interpretación:** En general, cuanto menor sea el valor de la tasa de corte, mejor. Una disminución de la tasa de corte a lo largo de la recogida de datos indicaría una posible reducción de la motivación de los encuestados para proporcionar datos útiles en el marco de la encuesta GATS hacia la fase final de dicha recogida. Si bien las tasas de corte son mediciones útiles del desempeño en la recogida de datos y la calidad general de la encuesta, este cálculo no es esencial para la revisión oficial de las medidas de calidad estadística de la encuesta GATS.

### 5.3.3 Patrones de las tasas de respuesta de los hogares por estratos de muestreo de la primera etapa

La tasa de respuesta de los hogares se calcula como la proporción entre el número ponderado de hogares para los que se han facilitado datos completos de la lista de preselección y el número ponderado de hogares abarcados (véase la fórmula 19 en el *Manual de ponderación de la muestra de la encuesta GATS*). Los hogares que proporcionan solo datos parciales de la lista de preselección se tratan como negativas o interrupciones por lo que respecta a la tasa de respuesta de los hogares y no se incluyen en el numerador del cálculo de dicha tasa. El *Cuestionario para hogares* se considera cumplimentado si tiene un código de disposición final de hogar de 1 (véase el *apéndice C.1*). Utilizando solo los hogares abarcados, la tasa de respuesta de los hogares se calcula de la siguiente manera:

$$\text{Hogares – Tasa de respuesta} = \frac{[1]}{[1] + [3] + [4] + [5] + [6] + [9]} \times 100$$

donde

- 1 = *Cuestionario para hogares* cumplimentado, con una persona seleccionada
- 2 = *Cuestionario para hogares* cumplimentado, sin ninguna persona seleccionada
- 3 = *Cuestionario para hogares* parcialmente cumplimentado, sin que haya sido posible finalizar la lista de preselección (entrevista incompleta)
- 4 = *Cuestionario para hogares* no cumplimentado, sin que haya sido posible identificar a un encuestado para la evaluación
- 5 = Nadie en casa
- 6 = Negativa del hogar a responder a la encuesta
- 9 = Otros casos de falta de respuesta del hogar

Obsérvese que el código de disposición final 2 en el *Cuestionario para hogares* queda excluido tanto del numerador como del denominador de la tasa de respuesta de los hogares, ya que se considera que los hogares con dicha codificación no cumplen los requisitos para responder a la encuesta. Esta tasa de respuesta se ajusta a RR1 según lo definido por la American Association of Public Opinion Research (Asociación Estadounidense para la Investigación de la Opinión Pública)

(<https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>).

Los países deben facilitar la documentación completa sobre el modo de calcular las tasas de respuesta de los hogares. Dichas tasas tendrán que indicarse para todo el territorio del país por estratos de muestreo de la primera etapa (es decir, entorno urbano o rural y región) si los tamaños de muestra son suficientes para la estimación regional. Cuando proceda, deberá efectuarse un desglose adicional utilizando la clasificación cruzada de las variables entorno urbano o rural y región.

#### 5.3.4 Patrones de las tasas de respuesta de las personas entre las variables utilizadas para los ajustes por falta de respuesta

La tasa de respuesta de las personas se calcula como la proporción entre el número ponderado de entrevistas de la encuesta GATS parcial o totalmente concluidas y el número ponderado de personas de la muestra que cumplen los requisitos elegidos de la lista de preselección de todos los miembros del hogar. Concretamente, los códigos de disposición finales relativos a las personas son los siguientes:

- 11 = *Cuestionario individual* cumplimentado
- 12 = Entrevista incompleta
- 13 = Determinación *a posteriori* de que la persona seleccionada no cumplía los requisitos para participar en la encuesta GATS
- 14 = El encuestado seleccionado no se encontraba en casa
- 15 = El encuestado seleccionado se niega a responder a la encuesta
- 16 = El encuestado seleccionado no es competente
- 17 = Otros casos de falta de respuesta individual

Una forma de calcular la tasa de respuesta de las personas es la siguiente:

$$\text{Personas} - \text{Tasa de respuesta} = \frac{[11]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

La tasa de respuesta de las personas indicada anteriormente supone que todas las personas seleccionadas y cuya aptitud se desconoce (es decir, código de disposición final 14) cumplen los requisitos para participar en la encuesta GATS. Este criterio puede hacer que se subestimen las tasas de respuesta de las personas si los entrevistadores seleccionan a menudo a encuestados que, una vez iniciada la entrevista, se descubre que no cumplen los requisitos para realizar la encuesta. Se recomienda que los países estimen la proporción de los encuestados seleccionados de la lista de preselección que realmente reúnan las condiciones para responder a la encuesta de la encuesta GATS (e) utilizando frecuencias ponderadas de los códigos de disposición como sigue:

$$e_{\text{personas-tasa de respuesta}} = \frac{[11] + [12] + [15] + [16] + [17]}{[11] + [12] + [13] + [15] + [16] + [17]}$$

Si esta proporción para la muestra en su conjunto es inferior a 0,90, los países deben ajustar el componente desconocido de la tasa de respuesta de las personas multiplicando las personas cuya aptitud se desconoce (código de disposición final 14) por esta proporción (e). La siguiente fórmula se

ajusta a la tasa de respuesta RR3 de la Asociación Estadounidense para la Investigación de la Opinión Pública (2004):

$$\text{Personas} - \text{Tasa de respuesta} = \frac{[11]}{[11] + [12] + [e^{\text{personas-tasa de respuesta}}[14]] + [15] + [16] + [17]} \times 100$$

Esta tasa de respuesta se ajusta a RR3 según lo definido por Asociación Estadounidense para la Investigación de la Opinión Pública y documentado anteriormente.

Los países deben facilitar la documentación completa sobre el modo de calcular las tasas de respuesta.

Dichas tasas tendrán que indicarse para todo el territorio del país y por subzonas regionales si los tamaños de muestra son suficientes para la estimación regional. Se debe realizar un desglose adicional dentro de la región (si procede) para las variables utilizadas con el fin de definir el ajuste por falta de respuesta. Entre tales variables puede haberse incluido la ubicación en entorno urbano o rural y la edad indicada en la lista de preselección (15 a 24, 25 a 34, 35 a 44, 45 a 54, 55 o más), así como el género (hombre o mujer) y la situación con respecto al tabaquismo (fumador o no fumador).

En el *Manual del supervisor sobre el terreno de la encuesta GATS* se describen las responsabilidades de esta figura en lo que respecta al examen y la corrección de las bajas tasas de respuesta de los equipos de encuestadores.

### 5.3.5 Patrones de las tasas de rechazo de las personas entre las variables utilizadas para los ajustes por falta de respuesta

Las tasas de rechazo se calculan como la proporción entre el número ponderado de encuestados seleccionados que cumplan los requisitos y que se negaron a responder al *Cuestionario individual de la encuesta GATS* (código de disposición final 15 en el *Cuestionario individual*) y la suma del número ponderado total de cuestionarios individuales de la encuesta GATS cumplimentados (11) e incompletos (12) más los no encuestados de aptitud desconocida (14), los que se negaron a responder (15), los encuestados considerados no competentes (16) y otros casos de falta de respuesta individual (17). Partiendo de los códigos de disposición individuales, el cálculo se realiza de la siguiente manera:

$$\text{Personas} - \text{Tasa de rechazo} = \frac{[15]}{[11] + [12] + [14] + [15] + [16] + [17]} \times 100$$

Esta tasa de rechazo se ajusta a REF1 según lo definido por la Asociación Estadounidense para la Investigación de la Opinión Pública.

Las tasas de rechazo de las personas deben calcularse directamente para la muestra en la fase final del proceso de garantía de la calidad, una vez finalizada la recogida de datos. Además, los supervisores sobre el terreno han de esforzarse constantemente por identificar a aquellos encuestadores sobre el terreno que resulten ineficaces en el transcurso de la recogida de datos y tomar medidas correctivas al respecto. Esto conllevaría un cálculo de las tasas de rechazo por cada encuestador o equipo de encuestadores, así como la detección de aquellos encuestadores o equipos de encuestadores cuya tasa sea excepcionalmente alta (por ejemplo, una tasa superior a la media de la tasa registrada entre todos los encuestadores o equipos más dos desviaciones típicas). Estas tasas de rechazo también pueden

calcularse en función de los siguientes criterios: 1) la semana de recogida de datos en la que tuvo lugar la entrevista, 2) la edad del encuestado, 3) el género del encuestado y 4) la ubicación del encuestado en un entorno urbano o rural.

Los países deben facilitar la documentación completa sobre el modo de calcular las tasas de rechazo. Dichas tasas tendrán que indicarse para todo el territorio del país y por subzonas regionales si los tamaños de muestra son suficientes para la estimación regional. Se debe realizar un desglose adicional dentro de la región (si procede) para las variables utilizadas con el fin de definir el ajuste por falta de respuesta. Entre tales variables puede haberse incluido la ubicación en entorno urbano o rural y la edad indicada en la lista de preselección (15 a 24, 25 a 34, 35 a 44, 45 a 54, 55 o más), así como el género (hombre o mujer) y la situación con respecto al tabaquismo (fumador o no fumador).

### 5.3.6 Tasas de falta de respuesta a los ítems para las variables indicadoras de las notas descriptivas

La tasa de falta de respuesta a los ítems de la encuesta GATS (es decir, la tasa porcentual de todos los encuestados que no proporcionan datos útiles para un ítem específico del cuestionario de la entrevista entre todos los encuestados que deberían haber respondido a la pregunta) debe calcularse y documentarse una vez concluido el procesamiento de los datos de la encuesta para los ítems del cuestionario correspondientes a todos los indicadores que vayan a incluirse en las *Notas descriptivas de la encuesta GATS* específicas para cada país (véanse las *Definiciones de los indicadores de la encuesta GATS* para más información). Las tasas de falta de respuesta a los ítems se calculan como la proporción entre el número de encuestados para los que no se obtuvo una respuesta válida pertinente (*MISSING<sup>x</sup>* para el ítem x), y la resta del número total de encuestados a nivel de unidad (I) menos el número de encuestados con un salto válido para el ítem x (*V<sup>x</sup>*):

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

El número total de personas que no responden a la encuesta a nivel de unidad de x (*MISSING<sup>x</sup>*) se obtendrá de una frecuencia no ponderada de encuestados con datos faltantes para el ítem x tras realizar una limpieza de datos adecuada para garantizar que se hayan seguido los patrones de salto adecuados. El número total de encuestados a nivel de unidad se obtendrá a partir de la frecuencia total no ponderada de la respuesta hombre o mujer aportada a la pregunta A01 (una variable sin previsión de campos en blanco). El número total de encuestados con un salto válido para el ítem x se puede obtener como la frecuencia del ítem x con una respuesta de .S (véase la *sección 5.1.3*). Muchos de los indicadores de las *Notas descriptivas de la encuesta GATS* son mediciones compuestas obtenidas a partir de las respuestas a varias preguntas. La falta de respuesta a cualquier ítem utilizado para obtener una medición compuesta dará lugar a la falta de respuesta de los ítems de la medición compuesta. Las tasas de falta de respuesta a los ítems inferiores al 5% se consideran bajas. Al igual que ocurre con las medidas de calidad 5.3.2 y 5.3.5 expuestas con anterioridad, los esfuerzos de los supervisores sobre el terreno deben orientarse a subsanar la falta excesiva de respuesta a los ítems registrada entre los datos de cada encuestador o equipo de encuestadores como medio para mejorar la calidad de la recogida de datos a medida que esta avance sobre el terreno.

## 5.4 Revisión formal de la calidad estadística

Se ha elaborado un protocolo científico para que el Comité de Revisión de la Muestra lleve a cabo una revisión oficial de las medidas de calidad estadística en cada país de la encuesta GATS después de que se hayan procesado los datos de la encuesta y se hayan calculado las ponderaciones de la muestra, pero antes de que se realice el análisis de los datos y se generen los resultados. La finalidad de esta revisión es verificar que las ponderaciones de la muestra se hayan calculado de conformidad con las especificaciones del *Manual de ponderación de la muestra de la encuesta GATS*, examinar diversas medidas indicativas de un alto grado de calidad de los datos de la encuesta y garantizar que los aspectos clave del muestreo y el análisis estén plenamente documentados. Se dispondrá de asistencia técnica en la preparación de esta revisión.

A continuación se describen las medidas de calidad específicas y la documentación necesaria para llevar a cabo este tipo de revisiones.

### 5.4.1 Documentación de las ponderaciones de la muestra

Cada etapa del proceso de ponderación y ajuste de los datos debe describirse pormenorizadamente en un documento elaborado por los estadísticos de los países directamente responsables de calcular y ajustar las ponderaciones estadísticas de la muestra de la encuesta GATS. En dicho documento ha de exponerse con detalle el proceso de ponderación y ajuste, aportando ejemplos ilustrativos del modo exacto de calcular todos los componentes de las ponderaciones finales para un grupo de 5 a 10 encuestados de la muestra real. Más concretamente, quienes vayan a realizar la revisión deberán llevar a cabo los siguientes preparativos de cara al examen<sup>1</sup>:

- Redactar un resumen del diseño de la muestra de la encuesta por etapas y describir el modo de calcular las probabilidades de selección para cada etapa y la ponderación básica general.
- Describir con detalle la manera en que se hayan ultimado las estrategias de ajuste por falta de respuesta y calibración, indicando concretamente lo siguiente: i) la definición de las clases de ponderación para los ajustes por falta de respuesta de hogares y personas, ii) las celdas de ajuste creadas para calibrar las ponderaciones y iii) los recuentos de población con los que se calibraron las ponderaciones. Consúltese el *Manual de ponderación de la muestra de la encuesta GATS* como guía para esta descripción.
- Elegir un grupo de 5 a 10 encuestados reales para ilustrar el proceso de cálculo de ponderaciones que se haya seguido en su caso. Es preciso seleccionar a encuestados de diferentes estratos, unidades primarias de muestreo y grupos de género. Si se han utilizado componentes de ponderación basados en intervalos semiabiertos, debe incluirse un encuestado de uno de estos agrupamientos geográficos entre alguno de los 5 a 10 ejemplos. Al final de este paso, se debe mostrar el cálculo de la ponderación ajustada final para cada ejemplo. Más concretamente, para cada uno de estos encuestados deberán llevarse a cabo las siguientes acciones:
  - Determinar los componentes de cada probabilidad condicional e incondicional (de cada etapa de muestreo) que se haya utilizado para elaborar la PONDERACIÓN BÁSICA.

---

<sup>1</sup> Obsérvese que, dado que muchos de los elementos de esta lista también deben incluirse en la documentación de cada *Informe del país de la encuesta GATS*, pueden enviarse borradores del texto de esta documentación (pero que contengan las versiones finales de los resultados numéricos requeridos) a quienes realicen la revisión oficial de la calidad estadística del país.

- Indicar y describir en su totalidad cada componente del cálculo del AJUSTE POR FALTA DE RESPUESTA para los hogares y las personas. Al concluir este paso, las ponderaciones básicas deberán ajustarse para tener en cuenta la falta de respuesta de los hogares y las personas.
- Indicar y describir pormenorizadamente cada componente del método de AJUSTE DE PONDERACIÓN POR CALIBRACIÓN para forzar la distribución ponderada final de la muestra con el fin de que se corresponda con la de una población con una calibración determinada.
- Presentar los resultados de los cálculos señalados a continuación sobre la base de las fórmulas facilitadas en las *secciones 5.2 y 5.3* del presente manual:
  - Ajustes de calibración con posterioridad a la estratificación para cada celda de ajuste utilizada en la calibración (5.2.1).
  - Efecto multiplicativo de las ponderaciones variables de la muestra (*Meff*) por cada uno de los subgrupos de población para los que vayan a indicarse las estimaciones de la encuesta en el *Informe del país de la encuesta GATS* (5.2.2).
  - Mediciones individuales y mediciones sinópticas (es decir, mínima, máxima y mediana) del efecto general del diseño y la medida de la homogeneidad dentro de la unidad primaria de muestreo para todas las estimaciones de la encuesta presentadas en las *Notas descriptivas de la encuesta GATS* (5.2.3).
  - Mediciones individuales y mediciones sinópticas (es decir, mínima, máxima y mediana) del margen de error u otra medida comparable de la precisión, para todas las estimaciones de la encuesta presentadas en las *Notas descriptivas de la encuesta GATS* (5.2.4).
  - Mediciones individuales y mediciones sinópticas (es decir, mínima, máxima y mediana) de las tasas de respuesta de los hogares mediante las celdas de ajuste utilizadas para producir el ajuste por falta de respuesta de los hogares (5.3.3).
  - Tasas de respuesta de las personas entre las celdas de ajuste utilizadas para los ajustes por falta de respuesta de las personas (5.3.4).
  - Tasas de rechazo de las personas entre las celdas de ajuste utilizadas para el ajuste por falta de respuesta de las personas (5.3.5).
  - Mediciones individuales y mediciones sinópticas (es decir, mínima, máxima y mediana) de las tasas de falta de respuesta a los ítems para todos los ítems de la encuesta utilizados para producir las tasas de consumo de tabaco en las *Nota descriptivas de la encuesta GATS* (5.3.6).

En el *Manual de ponderación de la muestra de la encuesta GATS (capítulo 4)* se sugieren varios métodos para garantizar con más eficacia la calidad de las ponderaciones de la muestra. Se llevarán a cabo las comprobaciones que figuran a continuación, por lo que quienes realicen la revisión de la calidad estadística necesitarán los correspondientes datos para ello:

- El tamaño promedio de las ponderaciones básicas, dividido entre el tamaño promedio de las ponderaciones ajustadas por falta de respuesta, es aproximadamente igual a la tasa de respuesta general final referente a las personas.

- La distribución ponderada de la muestra (obtenida a partir de las ponderaciones ajustadas finales) entre todas las celdas de ajuste de calibración debe coincidir exactamente con la distribución correspondiente de los recuentos externos de población con los que se hayan calibrado las ponderaciones.
- La suma de las ponderaciones ajustadas finales sobre la muestra total debe ser el tamaño total del recuento de la población (de personas mayores de 15 años) para la fuente utilizada en la calibración. Si esa fuente fue el último censo, entonces esta suma de ponderaciones debe ser el recuento de población de dicho censo a la que está dirigida la encuesta.
- La mayoría de los ajustes con posterioridad a la estratificación (por celda de ajuste) tienen que ser ligeramente mayores o menores que 1.
- Se obtiene un valor de  $Meff \leq 2,00$  para cada subgrupo informador en el *Informe del país de la encuesta GATS*.

## 5.5 Creación de un archivo de datos de análisis

Una vez concluida la ponderación de la muestra y realizadas todas las comprobaciones de garantía de la calidad, es preciso crear un nuevo archivo que contenga solo los casos individuales con el código de disposición final 11 (*Cuestionario individual* cumplimentado). En lo que respecta al *Cuestionario individual de la encuesta GATS*, únicamente se considerarán «encuestados» los casos individuales con dicho código.

*Obsérvese que no es necesario crear el conjunto final de datos de análisis antes de realizar las comprobaciones de garantía de la calidad basadas en el margen de error, el efecto del diseño o las principales estimaciones de la encuesta calculados mediante procedimientos de SAS o SPSS para muestras complejas, porque los casos con códigos de disposición final del Cuestionario individual distintos de 11 no tendrán valores para la variable de ponderación (es decir, tendrán valores faltantes) y el software los omitirá automáticamente por defecto a menos que se especifique lo contrario.*

Este nuevo archivo se denomina archivo de datos de análisis y debe utilizarse al realizar análisis de datos para producir estimaciones de encuestas (es decir, indicadores), en particular las estimaciones que figuran en las *Notas descriptivas de la encuesta GATS* y el *Informe del país de la encuesta GATS*.

## 5.6 Creación de archivos de datos de uso público

El conjunto de datos de uso público es el archivo de datos que los CDC publican para la población general. En él figuran exclusivamente los casos de entrevistas concluidas con todas las variables de respuesta de la encuesta y las variables indicadoras básicas normalizadas de la encuesta GATS. Asimismo, en el archivo de datos pueden encontrarse las variables geográficas y las variables de muestreo de la encuesta, como el estrato, el agrupamiento y las ponderaciones finales de la encuesta.

La finalidad de la publicación del conjunto de datos de uso público es facilitar a los investigadores externos acceso a los datos de la encuesta GATS. Atendiendo a la *Política de publicación de datos de la encuesta GATS*, dicho conjunto se publicará un año después de que el país publique sus *Notas descriptivas de la encuesta GATS*.

## 6. Requisitos relativos a los datos y la documentación

---

En este capítulo se describen los archivos de datos y la documentación que se enviarán desde los países y se archivarán en el centro coordinador de datos de la encuesta GATS.

Archivos de datos y documentación solicitados al país:

1. Archivos DB3 finales programados: CMSDB.db3, Survey0.db3 y Survey1.db3.
2. Archivo DB3 maestro creado por el software de agregación de CMSDB.db3, Survey0.db3 y Survey1.db3.
3. Archivo de la base de datos maestra (con uso de la variable de idioma inglés) generado mediante la herramienta de generación de archivos de análisis que contenga los archivos planos transpuestos para la tabla de respuestas con respuestas válidas tanto para los *Cuestionarios para hogares* como para los *Cuestionarios individuales*, así como el archivo de selección de la muestra maestra (identificadores de caso e identificadores de muestra).
4. Archivos de sintaxis de SAS, SPSS y STATA.
5. Componentes y documentación de las ponderaciones de la muestra (consúltese la *sección 5.4*).
6. Conjunto de datos ponderados: el archivo o los archivos planos, finales y limpios con las ponderaciones y especificaciones de ponderación.



## 7. Bibliografía

---

- Biemer, Paul P. y Lyberg, Lars E. 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 Ángeles, California.
- Gabler, S., Haeder, S., y Lahiri, P. (1999). «A model based justification of Kish's formula for design effects for weighting and clustering», *Survey Methodology*, 25 (1): págs. 105 y 106.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Especificaciones de programación del cuestionario básico*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Cuestionario básico con preguntas opcionales*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Plan de ejecución de la gestión de datos*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Guía de capacitación para la gestión de datos*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Manual del encuestador sobre el terreno*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Manual del supervisor sobre el terreno*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Manual de cartografía y establecimiento de listas*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Guía del programador sobre el sistema general de encuestas*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Especificaciones pregunta por pregunta*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Manual de diseño de la muestra*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Grupo de Colaboración de la Encuesta Mundial Sobre Tabaco en Adultos. *Encuesta Mundial Sobre Tabaco en Adultos (GATS): Manual de ponderación de la muestra de la encuesta GATS*. Atlanta (Georgia): Centros para el Control y la Prevención de Enfermedades, 2019.
- Referencia interna sobre valores faltantes de carácter especial en SAS:  
<http://support.sas.com/documentation/cdl/en/lrcon/61722/HTML/default/a000992455.htm>
- Kish, L. (1965). *Survey Sampling*, John Wiley and Sons, Nueva York.
- Kalton, G. y Flores-Cervantes, I. (2003). «Weighting Methods». *Journal of Official Statistics*, 19 (2), págs. 81 a 97.
- Potter, F. J. (1988). «Survey of Procedures to Control Extreme Sampling Weights». *Proceedings of the Section on Survey Research Methods, American Statistical Association*, págs. 453 a 458.



## Apéndice A: Glosario

---

**Archivo de casos:** Lista de muestreo e información que se carga en los dispositivos portátiles para la recogida de datos.

**Cuestionario básico:** Conjunto de preguntas básicas normalizadas que se utilizan en la encuesta GATS. Los países participantes adaptan el cuestionario básico a sus necesidades específicas ateniéndose a una serie de directrices concretas.

**Archivos DB3:** Tipo de archivos (SQLite) utilizados para la programación del software de encuestas del sistema general de encuestas.

**Sistema general de encuestas:** Programa de software de encuestas que se carga en los dispositivos portátiles para llevar a cabo la encuesta GATS.

**Especificaciones de programación del sistema general de encuestas:** Especificaciones complementarias del cuestionario en papel utilizado por los países para crear sus correspondientes versiones adaptadas que se elaboran para cada una de tales versiones y sirven para programar el cuestionario en los dispositivos portátiles.

**Configuración de los dispositivos portátiles:** Proceso de copia de la versión del software en el dispositivo portátil que generalmente se realiza en todos los dispositivos portátiles a la vez antes de iniciar la capacitación del encuestador sobre el terreno.

**Archivo de selección de la muestra maestra:** Conjunto de datos que contiene un identificador de caso para cada hogar de la lista de muestreo junto con la información necesaria para calcular la ponderación de la muestra y analizar los datos complejos de la encuesta.

**Garantía de la calidad:** Proceso consistente en la realización de actividades sistemáticas diseñadas para garantizar, evaluar y confirmar la calidad de los datos recogidos durante una encuesta.

**Comité de Revisión del Cuestionario:** Grupo compuesto por expertos internacionales en control del tabaco y diseño de cuestionarios. El Comité de Revisión del Cuestionario revisa y aprueba todos los cuestionarios adaptados a los países de la encuesta GATS para garantizar la calidad, la normalización y la comparabilidad.

**Comité de Revisión de la Muestra:** Grupo compuesto por expertos internacionales en muestreo. El Comité de Revisión de la Muestra revisa y aprueba los diseños de muestra, la ponderación de la muestra y las medidas de garantía de la calidad de los datos correspondientes a los países participantes.



## Apéndice B: Fase previa a la recogida de datos

---

### B.1 Proceso de adaptación y revisión del cuestionario de la encuesta GATS

Adaptación del cuestionario: Los países de la encuesta GATS colaborarán con los CDC en la elaboración de un cuestionario de la encuesta GATS adaptado específicamente a cada país, lo que engloba lo siguiente:

- **Adaptación del cuestionario básico de la encuesta GATS.** Los países adaptarán el cuestionario básico a su situación específica. Ello incluye modificar las listas de elementos de las preguntas (por ejemplo, tipos de tabaco) y las opciones de respuesta (por ejemplo, la lista de marcas de cigarrillos) para adecuarlas al país.
- **Incorporación de preguntas opcionales.** La encuesta GATS dispone de una lista con sugerencias de preguntas de encuesta opcionales que los países pueden añadir al cuestionario si lo desean.
- **Inclusión de preguntas adicionales elaboradas por el país.** Los países pueden incluir preguntas de encuesta propias que no figuren en el cuestionario básico ni entre las preguntas opcionales.
- **Eliminación de preguntas irrelevantes.** Los países pueden eliminar preguntas básicas que no sean relevantes en su situación (por ejemplo, cuando no exista prevalencia del tabaco sin humo). Para ello, es necesario proporcionar una justificación adecuada.

La revisión del cuestionario de la encuesta GATS es un proceso consistente en los 10 pasos que figuran a continuación:

1. **Adaptación del borrador del cuestionario específico para el país.** Los países colaborarán con los CDC en la creación de un borrador del cuestionario específico para cada país que se presentará al Comité de Revisión del Cuestionario.
2. **Establecimiento de formato.** Los CDC revisarán el borrador del cuestionario enviado y establecerán el formato necesario (por ejemplo, editar las instrucciones de salto o resaltar elementos para su revisión por parte del Comité de Revisión del Cuestionario).
3. **Revisión especializada por el Comité de Revisión del Cuestionario.** Los miembros del Comité de Revisión del Cuestionario proceden a dicha revisión y formulan observaciones al respecto.
4. **Recopilación de observaciones de los revisores.** Los CDC recopilan las observaciones del Comité de Revisión del Cuestionario y las envían nuevamente al país para su revisión.
5. **Comunicación (resolución).** El Comité de Revisión del Cuestionario y el país colaboran en la resolución de cualquier cuestión pendiente (a través del correo electrónico y/o por teleconferencia).
6. **Conclusión del cuestionario adaptado.** Una vez resueltas todas las cuestiones, se procede a la conclusión del cuestionario adaptado al país y el Comité de Revisión del Cuestionario otorga su aprobación oficial al respecto.
7. **Notificación de aceptación por parte del país.** El país otorga su aceptación definitiva del cuestionario aprobado por el Comité de Revisión del Cuestionario.
8. **Traducción.** Los países traducirán la versión aprobada del cuestionario de la encuesta GATS al idioma o los idiomas que corresponda y proporcionarán asimismo una retrotraducción del cuestionario al inglés. (Nota: Este paso puede desarrollarse a la vez que el paso 1 en aquellos casos en que los países traduzcan primero el cuestionario básico y lo adapten en sus propios

idiomas y, posteriormente, lo retrotraduzcan al inglés para enviarlo al Comité de Revisión del Cuestionario).

9. **Verificación de las versiones traducidas y retrotraducidas finales.** El Comité de Revisión del Cuestionario revisará las versiones traducidas y retrotraducidas para verificar que la traducción sea adecuada (es decir, que el significado de las preguntas no haya cambiado). (Nota: Esto forma parte del proceso de revisión del Comité de Revisión del Cuestionario si la traducción se realiza simultáneamente al paso 1 y se envía inicialmente una versión retrotraducida para dicha revisión).
10. **Programación.** Se crean especificaciones de programación para el cuestionario adaptado al país y se programa la encuesta específica para cada país en el dispositivo portátil.

## B.2 Proceso de programación del cuestionario de la encuesta GATS

A continuación se indican los principales pasos y directrices sobre plazos de programación del cuestionario de la encuesta GATS:

1. Cuestionario de la encuesta GATS del país y propuesta de prueba preliminar aprobados entre 6 y 8 semanas antes de la capacitación para la prueba preliminar.
2. Comienzo de la programación del cuestionario en inglés 6 semanas antes de la capacitación para la prueba preliminar.
3. Hardware de los dispositivos portátiles enviado y operativo 4 semanas antes de la capacitación para la prueba preliminar.
4. Traducciones de los menús y los mensajes del sistema general de encuestas, en caso necesario, terminadas 1 semana antes de la capacitación para la prueba preliminar.
5. Traducción del texto de los cuestionarios (*Cuestionario para hogares* y *Cuestionario individual*) en todos los idiomas del país organizador terminadas e importadas en las bases de datos survey0 y survey1 con 2 semanas de antelación a la capacitación para la prueba preliminar.
6. Visto bueno del país organizador al *Cuestionario para hogares* y el *Cuestionario individual* otorgado 1 semana antes de la capacitación para la prueba preliminar.
7. Comienzo del control de versiones de los cuestionarios con la versión autorizada según lo mencionado anteriormente y mantenimiento de un control estricto en adelante.

## B.3 Proceso de configuración de los dispositivos portátiles con el programa del sistema general de encuestas

En esta sección se describen las medidas de calidad que deben aplicarse en las diferentes fases del proceso para garantizar que cada dispositivo portátil que vaya a utilizarse sobre el terreno tenga las mismas especificaciones de software y hardware y que estas sean correctas.

1. **Creación de la versión principal:** La creación de la versión principal debe llevarse a cabo en una sola máquina que esté libre de cualquier virus y que tenga instalado el software antivirus adecuado. Una vez creada, dicha versión deberá cargarse en dos dispositivos portátiles distintos y deberá realizarse una iteración completa de prueba en ambos dispositivos respondiendo a cada pregunta del cuestionario. Si se detecta algún error en cualquiera de los dispositivos portátiles, será necesario corregir la programación —como se menciona en la *sección 3.3.2—*, teniendo que repetirse este proceso hasta que se logre un 100% de corrección. Este proceso debe llevarse a cabo en presencia del coordinador o la persona de contacto, junto con el responsable de TI.

2. **Colocación de etiquetas adhesivas de comprobación para el control de la calidad (plantilla):** Todos los dispositivos portátiles deben tener colocada, en su parte posterior, una etiqueta adhesiva con el identificador de encuestador sobre el terreno, el número de serie y una lista de verificación para el control de la calidad en blanco. Véase el ejemplo a continuación:

<b>701010</b>		NS A10	
		<b><u>CC1</u></b>	<b><u>CC2</u></b>
<input type="checkbox"/> Configuración inicial	<input type="checkbox"/> Fecha	<input type="checkbox"/> Fecha	
<input type="checkbox"/> APK cargado	<input type="checkbox"/> Identificador de encuestador sobre el terreno	<input type="checkbox"/> Identificador de encuestador sobre el terreno	
<input type="checkbox"/> Transmisión	CH _____	<input type="checkbox"/> Eliminar casos para la capacitación	
	CI _____	<input type="checkbox"/> Carga _____	
	<input type="checkbox"/> Carga _____	Transmisión	
Ini _____	Transmisión	Casos _____	
	CC1 _____	CC2 _____	

3. **Preparación de los dispositivos portátiles:** Una vez copiada la versión final, todos los dispositivos portátiles deben configurarse mediante el proceso gradual adecuado (consúltese el proceso de preparación de dispositivos portátiles descrito en la *Guía del programador sobre el sistema general de encuestas de la encuesta GATS*). Será necesario repetir el proceso de separación de los dispositivos portátiles en diferentes cajas (antes y después de la configuración).
4. **Ejecución de la fase 1 de control de la calidad (CC1):** Los dispositivos portátiles han de repartirse equitativamente entre los miembros del personal de control de la calidad, que deberá formar parte del equipo que haya configurado los dispositivos portátiles o haya supervisado el proceso. Cada paso de CC1 debe llevarse a cabo en todos los dispositivos portátiles. La ejecución de CC1 ha de ser un proceso riguroso de control de la calidad en el que se compruebe la corrección de los elementos indicados a continuación en cada dispositivo portátil y se refleje posteriormente la correspondiente información en la lista de verificación de la etiqueta adhesiva:
- 1) Fecha y hora.
  - 2) Identificador de encuestador sobre el terreno (verificar que figure en la etiqueta adhesiva el identificador del *Cuestionario para hogares*).
  - 3) Versión x del *Cuestionario para hogares*.
  - 4) Versión x del *Cuestionario individual*.
  - 5) Carga (duración de la batería).

Se podrían añadir otras comprobaciones en función del entorno del país, como el número de casos, dispositivos portátiles para encuestadores o encuestadoras, etcétera. En caso de detectarse cualquier problema, el dispositivo portátil tendrá que colocarse en una caja aparte y reiniciarse nuevamente, debiendo repetirse todo el proceso para este tipo de dispositivos desde el paso 1.

5. **Ejecución de la fase 2 de control de la calidad (CC2):** CC2 se lleva a cabo una vez que se haya terminado de comprobar todos los dispositivos portátiles mediante CC1 y se deje constancia del visto bueno en la etiqueta adhesiva. Los dispositivos portátiles deberán reasignarse de modo que la persona que ejecute la fase 1 de control de la calidad sea distinta de la que ejecute la fase 2 de control de la calidad. Durante CC2 deberá llevarse a cabo el mismo proceso descrito anteriormente para CC1; el paso adicional en CC2 es comprobar que todas las etiquetas adhesivas estén

marcadas y se hayan verificado asimismo todos sus elementos. Es importante llevar a cabo CC2, ya que se ha observado que una pequeña parte de los errores se detecta en esta fase. Si se detectan problemas en cualquier dispositivo portátil durante CC2, deberá realizarse una nueva iteración con el dispositivo en cuestión empezando por un reinicio completo en el paso 1.

6. **Consignación del visto bueno en la hoja de aprobación definitiva por parte del responsable de TI y la persona de contacto del país:** Una vez que todos y cada uno de los dispositivos portátiles se hayan sometido a CC1 y CC2, el responsable de TI y la persona de contacto para la encuesta GATS en el país deben documentar y aprobar un rápido informe resumido en el que figuren el número de dispositivos portátiles, los números de versión del *Cuestionario para hogares* y el *Cuestionario individual*, la fecha, la hora y los miembros del equipo de producción y control de la calidad de los dispositivos portátiles. En dicho documento de aprobación se debe indicar que se han ultimado los dispositivos portátiles para la capacitación para la encuesta completa o el trabajo sobre el terreno y que no se realizarán más modificaciones en la programación ni se utilizarán los dispositivos con ningún otro fin hasta que concluya la encuesta GATS en el país.

## Apéndice C: Fase posterior a la recogida de datos

### C.1 Códigos de disposición finales y cálculos de la tasa de respuesta

Tras un examen detenido de las publicaciones sobre encuestas nacionales de hogares relacionadas y no relacionadas con el tabaco llevadas a cabo mediante la modalidad de recogida de datos por teléfono o presencial en el hogar, se ha confeccionado una lista de códigos de disposición finales y cálculos de la tasa de respuesta aplicables a la encuesta GATS. A tal efecto, se analizaron varias encuestas, entre ellas las del Instituto Brasileño de Geografía y Estadística (IBGE), la Encuesta Nacional de Demografía y Salud Sexual y Reproductiva de 2004 (ENDSSR-2004), la Encuesta de Salud Reproductiva de Georgia, las encuestas demográficas y de salud (EDS), la Encuesta Nacional de Salud Familiar, el Sistema de Vigilancia de Factores de Riesgo del Comportamiento, el Sistema de Vigilancia de las Conductas de Riesgo en los Jóvenes y la Asociación Estadounidense de Investigación de la Opinión Pública.

Los códigos de disposición finales y la presentación de información sobre la tasa de respuesta se describen en las tablas que figuran más adelante en esta sección. En las *tablas C-1 y C-2* se proporcionan los códigos de disposición finales del *Cuestionario para hogares* y el *Cuestionario individual* que deben asignarse en relación con diversos códigos de resultados finales de ambos cuestionarios. (En el *Manual del encuestador sobre el terreno de la encuesta GATS* se puede consultar información más detallada sobre los códigos de resultados). En la *tabla C-3* se facilita una plantilla ilustrativa del modo de calcular e indicar las tasas de respuesta y los códigos de disposición en el *Informe del país de la encuesta GATS*.

**Tabla C-1: Códigos de disposición y de resultados finales del Cuestionario para hogares**

Códigos de disposición finales	Códigos de resultados finales	Nombre y descripción
1	200	<i>Cuestionario para hogares</i> cumplimentado, con una persona seleccionada —El Cuestionario para hogares se considera cumplimentado si se ha finalizado la lista de preselección.
2	201	<i>Cuestionario para hogares</i> cumplimentado, sin ninguna persona seleccionada —No hay nadie mayor de 15 años. —No hay nadie que considere que el hogar sea su lugar de residencia habitual. —El hogar se asignó a hombres y no reside ningún hombre en el hogar o el hogar se asignó a mujeres y no reside ninguna mujer en el hogar (si se ha utilizado la aleatorización por género).
3	202	<i>Cuestionario para hogares</i> parcialmente cumplimentado, sin que haya sido posible finalizar la lista de preselección —Entrevista incompleta.
4	203	<i>Cuestionario para hogares</i> no cumplimentado, sin que haya sido posible identificar a un encuestado adecuado para la evaluación —No hay nadie mayor de 18 años en casa. —El miembro del hogar mayor de 18 años disponible no es competente. —En estos hogares puede haber o no residentes que cumplan los requisitos para participar en la encuesta.
5	209	Nadie en casa
6	204	Negativa del hogar
7	205	Casa desocupada
8	206	La dirección seleccionada no corresponde a un hogar

---

9	208	Otros casos de falta de respuesta del hogar
---	-----	---

---

**Tabla C-2: Códigos de disposición y de resultados finales del *Cuestionario individual***

Códigos de disposición finales	Códigos de resultados finales	Nombre y descripción
11	400, 402*	<i>Cuestionario individual</i> cumplimentado —Cuestionario cumplimentado hasta la pregunta E01. —Ninguna de las preguntas clave (por ejemplo, B01/B02/B03 y C01/C02/C03 según corresponda) = No sabe o No contesta.
12	402*	Incompleto —Cuestionario interrumpido antes de la pregunta E01.
13	403	Determinación <i>a posteriori</i> de que la persona seleccionada no cumplía los requisitos para participar en la encuesta GATS —Edad < 15. —La persona no considera que el hogar sea su lugar de residencia habitual. —Género incorrecto (si se ha utilizado la aleatorización por género).
14	409	El encuestado seleccionado no se encontraba en casa
15	404	El encuestado seleccionado se niega a responder a la encuesta
16	407	El encuestado seleccionado está incapacitado
17	408	Otros casos de falta de respuesta individual

\* A los casos con un código de resultado final 402 (*Cuestionario individual* parcialmente cumplimentado) se les asignará un código de disposición de 11 (*Cuestionario individual* cumplimentado) o 12 (incompleto). Véanse las correspondientes normas resumidas en la *sección 5.1.7*.

**Tabla C-3: Tabla 3.1 del informe del país de la encuesta GATS: Plantilla de presentación de información sobre las tasas de respuesta**

	Residencia				Total	
	Entorno urbano		Entorno rural		Número	Porcentaje
	Número	Porcentaje	Número	Porcentaje		
<i>Hogar seleccionado</i>						
Cumplimentado (HC)						
Cumplimentado – Nadie reúne las condiciones (HCNRC)						
Incompleto (HINC)						
Ninguna persona para responder a la encuesta (HNPRE)						
Nadie en casa (HNC)						
Negativa a responder (HNR)						
Desocupado (HD)						
La dirección no es una vivienda (HDNV)						
Otros <sup>1</sup> (HO)						
Total de hogares seleccionados		100		100		100
<i>Tasa de respuesta de los hogares (TRH) (%)<sup>2</sup></i>						
<i>Persona seleccionada</i>						
Cumplimentado (PC)						
Incompleto (PINC)						
No reúne las condiciones (PNRC)						
Nadie en casa (PNC)						
Negativa a responder (PNR)						
Incapacitada (PI)						
Otros <sup>1</sup> (PO)						
Número total de personas incluidas en la muestra		100		100		100
<i>Tasa de respuesta de las personas (TRP) (%)<sup>3</sup></i>						
<i>Tasa de respuesta de los hogares (TRH) (%)<sup>4</sup></i>						
<sup>1</sup> Otros, incluye cualquier otro resultado no enumerado.						
<sup>2</sup> La tasa de respuesta de los hogares (TRH) se calcula de la siguiente manera:						
$\frac{HC * 100}{HC + HINC + HNPRE + HNC + HNR + HO}$						
<sup>3</sup> La tasa de respuesta de las personas (TRP) se calcula como:						
$\frac{PC * 100}{PC + PINC + PNC + PNR + PI + PO}$						
<sup>4</sup> La tasa de respuesta total (TRT) se calcula como: $(TRH \times TRP) / 100$						
<b>Notas:</b>						
— Se consideró que una entrevista de hogar incompleta (es decir, la lista de preselección no pudo terminarse) constituía una falta de respuesta a la encuesta GATS. Por consiguiente, estos casos (HINC) no se incluyeron en el numerador de la tasa de respuesta de los hogares.						
— El número total de personas incluidas en la muestra debe ser igual al número de entrevistas de hogares con el cuestionario cumplimentado [HC].						
— El concepto de entrevista personal con el cuestionario cumplimentado [PC] abarca a los encuestados que hayan cumplimentado al menos la pregunta E01 y que hayan aportado respuestas válidas a las preguntas B01/B02/B03 (y C01/C02/C03 cuando proceda). En el caso de los encuestados que no se ajustan a estos criterios, las entrevistas se consideran incompletas (PINC) y, por lo tanto, no se incluyen en el numerador de la tasa de respuesta de las personas.						

## **C.2 Patrón de ajustes de calibración de las ponderaciones con posterioridad a la estratificación entre celdas de ajuste**

### **Contexto**

El último paso en la generación de las ponderaciones de la muestra conlleva la calibración de tales ponderaciones con los recuentos de la población mediante correlaciones conocidas de mediciones clave de los resultados del estudio, denominadas variables de calibración (por ejemplo, género, educación, edad, entorno urbano o rural y región, como se sugiere en el *Manual de ponderación de la muestra de la encuesta GATS*). Desde el punto de vista procedimental, la calibración con posterioridad a la estratificación implica la creación de «celdas de ajuste» mediante la clasificación cruzada de las mediciones de correlación. El «ajuste con posterioridad a la estratificación» en cada una de estas celdas de ajuste será  $<1,00$  si la submuestra de esa categoría está sobrerrepresentada en la muestra una vez tomada en cuenta la falta de respuesta y la selección de la muestra, y el «ajuste con posterioridad a la estratificación» será  $>1,00$  en aquellas categorías en las que la muestra esté infrarrepresentada. El análisis del patrón de ajustes con posterioridad a la estratificación entre celdas de ajuste ayuda a entender en qué parte de la muestra se ha mantenido el desequilibrio una vez corregido el desequilibrio de la muestra debido a la falta de respuesta.

### **Fuente(s) de datos**

Para estos cálculos se debe utilizar un archivo de datos que contenga el conjunto final de ponderaciones de la muestra ajustadas en función de la falta de respuesta, así como todos los componentes computacionales de la ponderación final correspondiente a cada encuestado.

### **Software de cálculo**

Ninguno en particular, ya que los cálculos requeridos se pueden programar fácilmente en cualquier entorno de software.

### **Cálculo**

Los pasos del procedimiento para calcular los ajustes con posterioridad a la estratificación se describen en la *sección 3.3 del Manual de ponderación de la muestra de la encuesta GATS*. El conjunto de ajustes con posterioridad a la estratificación ha de organizarse a partir de los resultados finales de tales cálculos en forma de tabla multidireccional en función de las variables de correlación definitorias de las celdas de ajuste (es decir, las variables de calibración). Cabe esperar que todos los ajustes con posterioridad a la estratificación estén cerca de 1,00, con valores ligeramente superiores a 1,00 en algunos casos y situándose el resto levemente por debajo de 1,00.

### **Interpretación**

La constatación de grandes desviaciones de 1,00 en estos ajustes con posterioridad a la estratificación sugiere la existencia de graves problemas con el desequilibrio de la muestra debido a la cobertura diferencial de la muestra de hogares, a problemas no detectados con los cálculos de las ponderaciones de la muestra o a cambios demográficos importantes en la población a la que estaba dirigida la encuesta entre la fecha de los recuentos de población con los que se calibró la muestra y la fecha de muestreo y recogida de datos en la encuesta de la encuesta GATS.

**Tabla C-4: Tabla de ejemplo hipotético de ajustes de calibración de las ponderaciones con posterioridad a la estratificación, entre regiones, entornos urbano o rural, géneros e intervalos de edad.**

Región	Entorno urbano o rural	Género	Edad	Ajuste	
1	Urbano	Hombre	15 a 24 25 a 44 45 a 59 60 o más		
		Mujer	15 a 24 25 a 44 45 a 59 60 o más		
	Rural	Hombre	15 a 24 25 a 44 45 a 59 60 o más		
		Mujer	15 a 24 25 a 44 45 a 59 60 o más		
	2	Urbano	Hombre	15 a 24 25 a 44 45 a 59 60 o más	
			Mujer	15 a 24 25 a 44 45 a 59 60 o más	
Rural		Hombre	15 a 24 25 a 44 45 a 59 60 o más		
		Mujer	15 a 24 25 a 44 45 a 59 60 o más		
3		Urbano	Hombre	15 a 24 25 a 44 45 a 59 60 o más	
			Mujer	15 a 24 25 a 44 45 a 59 60 o más	
	Rural	Hombre	15 a 24 25 a 44 45 a 59 60 o más		
		Mujer	15 a 24 25 a 44 45 a 59 60 o más		

### C.3 Efecto de las ponderaciones variables de la muestra en la precisión de las estimaciones de la encuesta

#### Contexto

La variación de las ponderaciones de la muestra suele generar un aumento de la varianza (es decir, una reducción de la precisión) de las estimaciones de la encuesta. Este tipo de variación supondrá un problema importante en los diseños de muestra de los países en los que el muestreo por entorno urbano o rural o por región haya resultado excesivo. Kish (1965, sección 11.7) propuso un modelo sencillo pero aplicado desde hace mucho tiempo para medir el aumento multiplicativo de la varianza cuando el tamaño de las ponderaciones no esté correlacionado con la medición de la encuesta utilizada para estimar una característica simple de la población ( $\theta$ ). Usaremos  $Meff_w$  para indicar este efecto multiplicativo.

$Meff_w$  no es un parámetro específico de la estimación, pero sí del ámbito de presentación de información de las estimaciones. Por consiguiente, se aplicará a todas las estimaciones del ámbito en cuestión en el marco de la encuesta GATS. Por ejemplo, el valor de  $Meff_w$  calculado para las estimaciones generales de la población ha de calcularse a partir de toda la muestra, pero el valor de  $Meff_w$  calculado para las estimaciones indicadas en relación con las personas que viven en zonas rurales debe calcularse únicamente a partir de los encuestados de dichas zonas.

#### Fuente(s) de datos

Para estos cálculos, debe emplearse el archivo final de datos ponderados utilizado para el análisis.

#### Software de cálculo

No se necesita ningún software especial. El cálculo de  $Meff_w$  se puede programar fácilmente partiendo de los resultados del software de análisis de encuestas utilizado.

#### Cálculo

Centrándose en la ponderación final ( $w_i$ ) calculada para cada uno de los  $n$  encuestados a los que se aplica  $Meff_w$ , calcúlese el promedio simple de las ponderaciones para obtener  $\bar{w}$  y, a continuación, úsese la fórmula,  $s_w^2 = \sum_{i=1}^n (w_i - \bar{w})^2 / (n - 1)$  para calcular la varianza de las ponderaciones de la muestra. Se puede calcular  $Meff_w$  a partir del conjunto final de ponderaciones de la muestra para los encuestados de la encuesta GATS de la siguiente manera:

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

Obsérvese que  $Meff_w \geq 1$ , dado que  $\bar{w}^2$  y  $s_w^2$  siempre serán, en ambos casos, valores positivos.

#### Interpretación

Si bien es claramente preferible que  $Meff_w$  esté lo más cerca posible de uno, un valor de  $Meff_w > 2,00$  podría considerarse significativo y, por lo tanto, requerir algún tipo de medida correctiva. Las estrategias

de recorte de ponderaciones o de censura constituyen el recurso más habitual en el caso de ponderaciones excesivamente variables (Potter, 1988). La decisión final sobre la necesidad de recortar las ponderaciones depende de la posibilidad de encontrar un equilibrio entre el beneficio en términos de precisión derivado del recorte y el efecto en las estimaciones de la encuesta derivado de la variación de las ponderaciones. Si el recorte de las ponderaciones reduce el valor de  $Meff_w$  pero no supone una variación apreciable de las estimaciones ponderadas para las mediciones clave de los resultados del estudio, el paso del recorte puede estar justificado habida cuenta del tiempo y el esfuerzo necesarios para llevarlo a cabo.

### **Ejemplo**

Dada una muestra de la encuesta GATS, el promedio de las ponderaciones finales de dicha muestra entre la totalidad de los XXXX encuestados es  $\bar{w} = YYY, YYYY$  y su varianza simple es  $s_w^2 = ZZZZ, ZZ$ . El efecto multiplicativo de las ponderaciones variables sobre las estimaciones nacionales generales de un país es  $Meff_w = 1 + \frac{YYY,YYYY}{XXXX,ZZ} = U, UU$

Una muestra con  $Meff_w = U, UU < 2,00$  puede considerarse una muestra en la que la variación de las ponderaciones es lo suficientemente discreta como para no requerir ninguna medida de reducción del efecto de las ponderaciones variables en las estimaciones, por ejemplo, mediante el recorte de ponderaciones (véase Potter, 1988).

## **C.4 Efecto general del diseño en la precisión de las estimaciones de la encuesta y la homogeneidad entre agrupamientos dentro de la unidad primaria de muestreo de las correspondientes estimaciones principales de la encuesta**

### **Contexto**

Kish (1965) definió inicialmente el concepto de «efecto del diseño», indicado mediante «Deff», con el fin de medir la variación multiplicativa de la varianza de una estimación de encuesta debida al uso del muestreo por conglomerados para un determinado diseño de muestra, en comparación con la varianza de una estimación similar obtenida a partir de una muestra aleatoria simple del mismo tamaño. Cada una de las mediciones clave de los resultados de la encuesta en el marco de la encuesta GATS constituye una proporción ( $p$ ), o una tasa expresable como porcentaje o promedio (media). El modelo original de Kish relativo al efecto del diseño (debido al muestreo por conglomerados) de una estimación ( $\hat{p}$ ) de  $p$  es simplemente,

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

donde  $\rho_{\theta}$  es una medición referente a encuestados dentro de la unidad primaria de muestreo de homogeneidad relativa dentro del agrupamiento con respecto a la medición de la encuesta utilizada para la estimación de  $p$  (medición del grado en que los miembros del mismo agrupamiento de muestra se asemejan más entre sí que los miembros de la población en su conjunto),  $\bar{m} = n / a$  es el número promedio de encuestados por unidad primaria de muestreo de la muestra,  $n$  es el tamaño global de la muestra de encuestados y  $a$  es el número de unidades primarias de muestreo de la muestra.

En los casos en que una muestra de encuesta conlleva tanto agrupamientos como ponderaciones desiguales de muestras, Gabler, et al. (1999) han demostrado que el efecto general del diseño en las

estimaciones de la encuesta, indicado mediante  $Deff(\hat{\theta})$ , es el producto tanto del muestreo por conglomerados (véase lo explicado anteriormente) como del efecto multiplicativo de las ponderaciones variables de la muestra (véase el análisis del  $Meff_w$ ); es decir:

$$Deff(\hat{\theta}) \equiv \frac{V(\hat{p}; \text{diseño}, n)}{V(\hat{p}_{srs}; srs, n)}$$

$$= [Meff_w][Deff_{CS}(\hat{\theta})] = \left[1 + \frac{S_w^2}{\bar{w}^2}\right][1 + \rho_\theta(\bar{m} - 1)]. \quad (1)$$

### Fuente(s) de datos

Para estos cálculos, debe emplearse el archivo final de datos ponderados utilizado para el análisis.

### Software de cálculo

Los mismos paquetes de software de análisis de encuestas aplicables a la tarea relativa al margen de error (véase el *apéndice C.5*).

### Cálculo

La varianza estimada o el error cuadrático medio (es decir,  $v(\hat{p}; \text{diseño}, n)$  o  $\sqrt{v(\hat{p}; \text{diseño}, n)}$  en la ecuación 1, respectivamente) se obtienen mediante el software de análisis de encuestas. Para estimar  $\rho_\theta$  podemos simplemente resolver  $\rho_\theta$  en la ecuación 1 y obtenemos lo siguiente:

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

A fin de calcular  $\rho_\theta$  para cada estimación de la encuesta, podemos calcular  $Meff_w$  a partir de la muestra; conocemos  $\bar{m}$  al saber el número de unidades primarias de muestreo de la muestra ( $a$ ) y el tamaño global de la muestra de encuestados ( $n$ ) y, en numerosos tipos de software de análisis de encuestas (por ejemplo, SUDAAN), podemos solicitar una estimación de  $Deff(\hat{\theta})$  junto con la estimación ( $\hat{\theta}$ ). Si no podemos obtener una estimación de  $Deff(\hat{\theta})$  directamente a partir de los resultados del software para su uso en la determinación de  $Deff(\hat{\theta})$  como se define en la ecuación 1 o de  $\rho_\theta$  según lo visto en la ecuación 2, entonces podemos estimar el efecto general del diseño para las proporciones estimadas de la siguiente manera:

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

dado que  $v(\hat{p}_{srs}; srs, n) = \hat{p}(1-\hat{p})/n$ . La estimación de  $\rho_\theta$  se puede obtener directamente a partir de la ecuación 2.

## Interpretación

Lo esperable suele ser que el efecto general del diseño relativo a una estimación (es decir  $Deff(\hat{\theta})$ ) sea inferior a 2,50. Las estimaciones de  $\rho_{\theta}$  (para la homogeneidad dentro de la unidad primaria de muestreo) suelen ser números positivos pequeños entre 0 y 1 (más habitualmente entre 0 y 0,05). En algunos casos, las estimaciones de  $\rho_{\theta}$  serán números pequeños negativos.

Para resumir la magnitud de  $deff(\hat{\theta})$  entre todas las estimaciones generadas en el marco de la encuesta GATS, se recomienda calcular  $Deff(\hat{\theta})$  para varias mediciones clave de los resultados de la encuesta (por ejemplo, la prevalencia actual del tabaquismo, la prevalencia del tabaquismo ocasional, las tasas de abandono del hábito, etcétera) en función de diversos ámbitos de presentación de información (por ejemplo, por género, edad, entorno urbano o rural, región y diversas clasificaciones cruzadas de estas variables). A continuación habría que resumir, para cada una de las mediciones de resultados clave, los valores estimados de  $Deff(\hat{\theta})$  entre todos los ámbitos de presentación de información indicando las siguientes estadísticas resumidas para los valores estimados de  $Deff(\hat{\theta})$ : mediana y valores mínimo y máximo.

Resúmanse los valores de  $\hat{\rho}_{\theta}$  (mediana y valores mínimo y máximo) entre todas las estimaciones de encuesta obtenidas en el marco de la encuesta GATS siguiendo el mismo procedimiento aplicado para  $deff(\hat{\theta})$ .

En el caso de algunos sistemas u organizaciones de encuestas, se prefiere presentar la información sobre el efecto del diseño indicándolo mediante el parámetro «Deft» en lugar de «Deff». El parámetro «Deft» se calcula para cada estimación como la proporción entre el error cuadrático medio obtenido al utilizar el diseño de muestra en cuestión y el error cuadrático medio que se habría obtenido de haber usado una muestra aleatoria simple.

## C.5 Margen de error de las estimaciones principales de la encuesta

### Contexto

Es necesario establecer algunas definiciones básicas de las medidas de precisión estadística obtenidas de las muestras de encuestas. En primer lugar, hay que tener en cuenta que es probable que todas las estimaciones obtenidas de las muestras difieran en algún grado de lo estimado en la población. Esta diferencia suele denominarse **error de muestreo**. Las muestras probabilísticas como las utilizadas en la encuesta GATS nos permiten generar mediciones sinópticas del error de muestreo que indican la precisión de las estimaciones. Existen diferentes mediciones sinópticas de la precisión de las estimaciones de una encuesta. A la hora de definir las, utilizaremos el símbolo  $\theta$  para indicar la característica de la población que estamos estimando (por ejemplo, la tasa de prevalencia de las personas que actualmente fuman cigarrillos). Asimismo, emplearemos  $\hat{\theta}$  para representar la estimación de  $\theta$  a partir de la muestra.

La primera medición sinóptica habitual de la precisión es la **varianza de la estimación**, expresada como  $V(\hat{\theta})$ . Como veremos, el resto de las mediciones están relacionadas de alguna manera con  $V(\hat{\theta})$ . La varianza de las estimaciones de la encuesta, y todas las demás mediciones conexas, constituyen

indicadores cuantitativos del grado previsto de variación de las estimaciones de la muestra entre todas las muestras posibles que podría generar el diseño de muestra utilizado en la encuesta. Una segunda medición de la precisión es el **error cuadrático medio de la estimación**, definido como  $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}$ . Por tanto, el error cuadrático medio de una estimación es simplemente la raíz cuadrada de la varianza de esa estimación y se calcula en las mismas unidades que la estimación  $\hat{\theta}$ . Una tercera medición de la precisión estadística es el **error cuadrático medio relativo** de la estimación, definido como  $RSE(\hat{\theta}) = SE(\hat{\theta}) / \theta = \sqrt{V(\hat{\theta})} / \theta$ . Dado que  $RSE(\hat{\theta})$  mide la precisión en relación con el tamaño de lo que se está estimando, no se expresa en unidades y, por ende, constituye un indicador de precisión más comparable entre las estimaciones que  $SE(\hat{\theta})$ .

La medición de la precisión recomendada para que los países indiquen las estimaciones principales en el marco de la encuesta GATS (por ejemplo, la tasa de prevalencia del consumo de tabaco, etcétera) es el **margen de error**, definido como  $MOE(\hat{\theta}) = [Z][SE(\hat{\theta})]$ , donde Z es una medida del nivel de confianza para la medición y  $SE(\hat{\theta})$  es el error cuadrático medio de  $\hat{\theta}$ . En el caso de la mayoría de las estimaciones nacionales en el marco de la encuesta GATS, podríamos utilizar  $Z = 1,96$  con un nivel de confianza del 95%. El valor estimado ( $moe(\hat{\theta})$ ) indicado para  $\hat{\theta}$  se interpreta de la siguiente manera:

*Estamos seguros al 95% de que el valor indicado ( $\hat{\theta}$ ) está dentro de la cantidad  $moe(\hat{\theta})$  de  $\theta$ .*

Se insta a los analistas de la encuesta GATS a que indiquen el valor de  $moe(\hat{\theta})$  correspondiente a todas las estimaciones clave, incluida la totalidad de las estimaciones nacionales de los principales indicadores del consumo de tabaco. También deberá señalarse el valor de  $moe(\hat{\theta})$  relativo a las correspondientes estimaciones regionales si el diseño de la muestra requiere un aumento de los tamaños de las muestras regionales a efectos de incorporar la generación de estimaciones regionales ajustadas a las normas de precisión de la encuesta GATS.

Dado que las medidas de precisión pueden obtenerse fácilmente con programas de software estadístico, se insta firmemente a los países de la encuesta GATS a solicitar de manera rutinaria estimaciones de precisión con respecto a todas las estimaciones que se generen, aun cuando solo sea necesario indicar las principales estimaciones nacionales (y posiblemente regionales). Las estimaciones cuyo valor de  $MOE(\hat{\theta})$  se aproxime a la magnitud de la propia estimación o la supere, deberán, si acaso, indicarse con precaución. Como medida adicional de precaución, es importante señalar que, al igual que ocurre con  $\hat{\theta}$ , las estimaciones de  $V(\hat{\theta})$ ,  $SE(\hat{\theta})$ ,  $MOE(\hat{\theta})$ , o  $RSE(\hat{\theta})$  están sujetas a errores de muestreo. Este aspecto es importante, ya que todas estas estimaciones pueden ser imprecisas si se basan en muestras pequeñas, en particular aquellas que abarquen menos de entre 10 y 20 unidades primarias de muestreo de la muestra.

### **Fuente(s) de datos**

Para estos cálculos, debe emplearse el archivo final de datos ponderados utilizado para el análisis.

## Software de cálculo

Las estimaciones de las características de la población ( $\hat{\theta}$ ) y sus correspondientes valores de  $moe(\hat{\theta})$  deben tener en cuenta las características del diseño de la muestra importantes desde el punto de vista estadístico. Las características estimadas deben ponderarse y las mediciones de la precisión conexas deben adecuarse al uso de la estratificación, el muestreo por conglomerados, el muestreo sin reemplazamiento y las ponderaciones de la muestra. Es bien sabido que, el hecho de no hacerlo así, generalmente contribuye a la obtención de estimaciones sesgadas (concretamente en términos de precisión) y, por lo tanto, a la obtención de estimaciones de intervalos inapropiadas y pruebas de significación de carácter inadecuado. Por lo tanto, se insta firmemente a los analistas de los países de la encuesta GATS a que utilicen un software de análisis que permita tener plenamente en cuenta el diseño de la muestra utilizado para producir los datos de la encuesta. Ello implica el uso de las ponderaciones de la muestra para generar todas las estimaciones (de  $\hat{\theta}$  y  $moe(\hat{\theta})$ ) a efectos del análisis descriptivo. Asimismo, este proceso conlleva la utilización de un software que aplique una metodología ampliamente aceptada para la estimación de las varianzas y los errores cuadráticos medios de las estimaciones de la encuesta. Los tres principales métodos para la estimación de la varianza son la linealización mediante series de Taylor, la replicación repetida equilibrada y la replicación repetida de Jackknife (Wolter, 1985). La mayoría de los paquetes de análisis de encuestas utilizan la linealización mediante series de Taylor para la estimación de la varianza, por lo que, por motivos de comparabilidad entre las conclusiones de los países de la encuesta GATS, se recomienda un software que aplique dicha metodología.

Varias organizaciones estadísticas de todo el mundo han diseñado programas de software para analizar datos procedentes de muestras complejas como ocurre en el caso de la encuesta GATS<sup>1</sup>. Tales programas no solo generarán estimaciones del consumo de tabaco (es decir,  $\hat{\theta}$ ) sino que también podrán generar estimaciones de precisión (es decir, normalmente  $V(\hat{\theta})$  o  $SE(\hat{\theta})$ ) que tengan debidamente en cuenta las características fundamentales del diseño en el marco de la encuesta GATS, a saber, el muestreo por conglomerados, el uso de la estratificación y las distintas probabilidades de selección (es decir, las ponderaciones de la muestra):

## Información adicional

Para obtener información más detallada sobre las definiciones expuestas anteriormente, consúltese la obra *Encyclopedia of Survey Research Methods* (Enciclopedia de Métodos de Investigación por Encuesta), (2007). Para obtener información en línea de todos los paquetes de software de análisis de encuestas, accédase al sitio web patrocinado por la sección de métodos de investigación por encuesta de la asociación American Statistical Association (<http://www.hcp.med.harvard.edu/statistics/survey-soft/>). Además de los enlaces al software de análisis, en dicho sitio web figuran asimismo otros muchos enlaces y recursos útiles en relación con el diseño y la realización de encuestas (véase <http://www.amstat.org/sections/SRMS/index.html>).

---

<sup>1</sup> El centro coordinador de datos de la encuesta GATS en los CDC brinda asistencia técnica en el uso de los siguientes paquetes de software estadístico: SAS, SPSS y STATA.

## Cálculo

Los productos del software enumerados anteriormente indicarán un valor de  $\hat{\theta}$ , así como su varianza estimada, señalada mediante  $v(\hat{\theta})$ , o su error cuadrático medio, expresado como  $se(\hat{\theta})$ . A partir de estos valores, se puede calcular el margen de error estimado para  $\hat{\theta}$ , de la siguiente manera:

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

## Interpretación

El valor de  $moe(\hat{\theta})$  indicado para  $\hat{\theta}$  se interpreta del siguiente modo:

Estamos seguros al 95% de que el valor estimado ( $\hat{\theta}$ ) está dentro de la cantidad de  $moe(\hat{\theta})$  de  $\theta$ .

## C.6 Estimaciones de errores de muestreo

Las estimaciones de una encuesta por muestreo se ven afectadas por dos tipos de error: 1) errores ajenos al muestreo y 2) errores de muestreo. Los *errores ajenos al muestreo* son el resultado de errores o equivocaciones no atribuibles al muestreo cometidos al ejecutar la recogida y el procesamiento de datos, como errores en la cobertura, errores de respuesta, errores de falta de respuesta, cuestionarios defectuosos, errores de registro por parte del encuestador, errores de procesamiento de datos, etcétera. Aunque durante la ejecución de la encuesta GATS se hayan realizado numerosos esfuerzos en [el país] encaminados a reducir al mínimo esos errores, los errores ajenos al muestreo son imposibles de evitar y difíciles de evaluar estadísticamente.

La muestra de encuestados seleccionados en el marco de la encuesta GATS de [país] es solo una de las muestras que podrían haberse seleccionado de la misma población, utilizando el mismo diseño y tamaño de muestra. Cada una de estas muestras habría arrojado resultados que diferirían algo de los resultados de la muestra real seleccionada. Los *errores de muestreo* constituyen una medición de la variabilidad entre todas las muestras posibles. El alcance de la variabilidad no se conoce exactamente, pero se puede estimar estadísticamente a partir de los resultados de la encuesta.

Para cada uno de los indicadores seleccionados, se aportan las siguientes mediciones del error de muestreo:

- **Valor (R):** Estimación de la prevalencia ponderada del indicador.
- **Error cuadrático medio (SE):** Los errores de muestreo generalmente se miden en términos de errores cuadráticos medios de muestreo para una estimación o indicador en concreto (R). Por tanto, el error cuadrático medio de una estimación es simplemente la raíz cuadrada de la varianza de esa estimación y se calcula en las mismas unidades que la estimación en cuestión.
- **Tamaño de la muestra (n):** Número total de observaciones utilizadas para calcular la estimación de la prevalencia (R).
- **Efecto del diseño:** El efecto del diseño, indicado mediante el parámetro «*Deff*», es la proporción entre la varianza real de un indicador, calculada con arreglo al método de muestreo utilizado en la encuesta, y la varianza calculada bajo el supuesto de un muestreo aleatorio simple. La raíz cuadrada del efecto del diseño, indicada mediante el parámetro «*Deft*», sirve para mostrar la

eficiencia del diseño de la muestra y se calcula, para cada estimación, como la proporción entre el error cuadrático medio obtenido con el diseño concreto de la muestra y el error cuadrático medio que se produciría de utilizar una muestra aleatoria simple. Un valor de DEFT de 1,0 indica que el diseño de la muestra es tan eficiente como una muestra aleatoria simple, mientras que un valor de DEFT por encima de 1,0 indica un aumento del error cuadrático medio debido al uso de un diseño de muestra más complejo. En general, en una encuesta bien diseñada, DEFT suele situarse entre 1 y 3. No obstante, es habitual que DEFT sea mucho mayor y pueda llegar a alcanzar un valor de 7 u 8.

- **Error cuadrático medio relativo (RSE):** El error cuadrático medio relativo, también denominado «coeficiente de variación» (CV), es la proporción entre el error cuadrático medio y el valor del indicador.
- **Margen de error (MOE):** El margen de error se calcula como el producto de la medición de confianza deseada y el error cuadrático medio de la estimación. El nivel de confianza generalmente se basa en un valor (Z) de la distribución normal típica. Por ejemplo, para un nivel de confianza del 95%, podemos usar  $Z=1,96$ .
- **Límites de confianza ( $R \pm 1,96SE$ ):** Estos límites se calculan para mostrar el intervalo en el que cabe suponer razonablemente que se situará el valor real registrado en la población. Dada cualquier estadística calculada a partir de la encuesta, el valor de esta se situará en un intervalo de más o menos dos veces el error cuadrático medio de la estadística en el 95% de todas las muestras posibles de tamaño y diseño idénticos.

### **Cálculo del error cuadrático medio**

Si la muestra de encuestados se hubiera seleccionado como una muestra aleatoria simple, habría sido posible utilizar fórmulas sencillas para calcular los errores de muestreo. Sin embargo, la muestra de la encuesta GATS [año] es fruto de un diseño estratificado en múltiples etapas, con la consiguiente necesidad de utilizar fórmulas más complejas. Para el cálculo de los errores de muestreo a partir de los datos de la encuesta GATS de [país], se recurrió a [versión del software estadístico]. En el caso de las estimaciones de la encuesta que constituyen promedios o proporciones, el método utilizado para estimar la varianza fue la linealización mediante series de Taylor.

El método de linealización mediante series de Taylor trata cualquier porcentaje o promedio como una estimación de la proporción,  $r = y/x$ , donde  $y$  representa el valor total de la muestra para la variable  $y$ , y  $x$  representa el número total de casos en el grupo o subgrupo analizado. La varianza de  $r$  se calcula mediante la fórmula indicada a continuación:

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

en la cual,  $Z_{hi} = y_{hi} - rx_{hi}$ , y  $Z_h = y_h - rx_h$

donde  $h$  ( $=1$  o  $2$ ) representa el estrato correspondiente a entorno urbano o rural,  $m_h$  es el número total de unidades primarias de muestreo seleccionadas en el  $h^o$  estrato,  $y_{hi}$  es la suma de los valores ponderados de la variable  $y$  en la  $i^a$  unidad primaria de muestreo en el  $h^o$  estrato,

$x_{hi}$  es la suma del número ponderado de casos en la  $i^{\text{a}}$  unidad primaria de muestreo en el estrato  $h^{\text{o}}$ , y  $f$  es la fracción de muestreo global, cuyo valor es tan pequeño que se ignora.

En el presente apéndice se exponen los resultados correspondientes a todo el país, a las zonas urbanas y rurales y a cada género. En la *tabla C-5* se indican el tipo de estadística (promedio, proporción o tasa) y la población de base para cada variable o indicador. En la *tabla C-6* figuran, además del error cuadrático medio (SE) descrito anteriormente, el valor de la estimación (R), el tamaño de la muestra, el efecto del diseño (DEFF o DEFT), el error cuadrático medio relativo (SE/R), el margen de error (MOE) y los límites de confianza del 95% ( $R \pm 1.96SE$ ) referentes a cada variable o indicador.

**Tabla C-5: Lista de los indicadores para la estimación de errores de muestreo, GATS de [país] [año]**

Indicador	Estimación	Población de base
Consumidores actuales de tabaco	Proporción	Adultos con edad $\geq 15$ años
Fumadores actuales de tabaco	Proporción	Adultos con edad $\geq 15$ años
Fumadores actuales de cigarrillos	Proporción	Adultos con edad $\geq 15$ años
Consumidores actuales de tabaco sin humo	Proporción	Adultos con edad $\geq 15$ años
Consumidores de tabaco a diario	Proporción	Adultos con edad $\geq 15$ años
Fumadores de tabaco a diario	Proporción	Adultos con edad $\geq 15$ años
Fumadores de cigarrillos a diario	Proporción	Adultos con edad $\geq 15$ años
Consumidores de tabaco sin humo a diario	Proporción	Adultos con edad $\geq 15$ años
Exconsumidores de tabaco a diario entre el total de adultos	Proporción	Adultos con edad $\geq 15$ años
Exconsumidores de tabaco entre quienes alguna vez han consumido tabaco a diario	Proporción	Consumidores de tabaco que alguna vez han consumido tabaco a diario con edad $\geq 15$ años
Exfumadores de tabaco a diario entre el total de adultos	Proporción	Adultos con edad $\geq 15$ años
Exfumadores de tabaco entre quienes alguna vez han fumado tabaco a diario	Proporción	Fumadores de tabaco que alguna vez han fumado a diario con edad $\geq 15$ años
Exconsumidores de tabaco sin humo a diario entre el total de adultos	Proporción	Adultos con edad $\geq 15$ años
Exconsumidores de tabaco sin humo a diario entre quienes alguna vez han consumido tabaco sin humo a diario	Proporción	Consumidores de tabaco sin humo que alguna vez han consumido tabaco sin humo a diario con edad $\geq 15$ años
Tiempo transcurrido hasta el primer consumo de tabaco dentro de los 5 minutos posteriores al momento de despertarse	Proporción	Consumidores de tabaco a diario con edad $\geq 15$ años
Tiempo transcurrido hasta el primer consumo de tabaco entre los 6 y los 30 minutos posteriores al momento de despertarse	Proporción	Consumidores de tabaco a diario con edad $\geq 15$ años
Intento de dejar de fumar en los últimos 12 meses	Proporción	Fumadores actuales y exfumadores que han estado sin fumar durante menos de 12 meses
Intento de abandono del consumo de tabaco sin humo en los últimos 12 meses	Proporción	Consumidores actuales de tabaco sin humo y exconsumidores de tabaco sin humo que han estado sin consumirlo durante menos de 12 meses
Profesional sanitario consultado acerca del tabaquismo	Proporción	Fumadores actuales y exfumadores que han estado sin fumar durante menos de 12 meses y que han acudido a un profesional sanitario en los últimos 12 meses

<b>Indicador</b>	<b>Estimación</b>	<b>Población de base</b>
Recomendación por parte de un profesional sanitario de dejar de fumar	Proporción	Fumadores actuales y exfumadores que han estado sin fumar durante menos de 12 meses y que han acudido a un profesional sanitario en los últimos 12 meses
Uso de la farmacoterapia para el abandono del tabaco	Proporción	Fumadores actuales y exfumadores que han estado sin fumar durante menos de 12 meses
Recurso a orientación o asesoramiento o a líneas telefónicas de ayuda al abandono del tabaco	Proporción	Fumadores actuales y exfumadores que han estado sin fumar durante menos de 12 meses
Piensa dejar de fumar, se está planteando dejar de fumar o va a dejar de fumar	Proporción	Fumadores actuales con edad $\geq$ 15 años
Exposición al humo ajeno en el hogar	Proporción	Adultos con edad $\geq$ 15 años
Exposición al humo ajeno en el lugar de trabajo	Proporción	Adultos que trabajan en interior
Exposición al humo ajeno en edificios públicos u oficinas gubernamentales	Proporción	Adultos con edad $\geq$ 15 años
Exposición al humo ajeno en centros de salud	Proporción	Adultos con edad $\geq$ 15 años
Exposición al humo ajeno en restaurantes	Proporción	Adultos con edad $\geq$ 15 años
Exposición al humo ajeno en el transporte público	Proporción	Adultos con edad $\geq$ 15 años
Última compra de cigarrillos en establecimiento	Proporción	Fumadores actuales de cigarrillos fabricados con edad $\geq$ 15 años
Última compra de cigarrillos a vendedores callejeros	Proporción	Fumadores actuales de cigarrillos fabricados con edad $\geq$ 15 años
Última compra de cigarrillos en quiosco	Proporción	Fumadores actuales de cigarrillos fabricados con edad $\geq$ 15 años
Conciencia de haber visto u oído información en contra del consumo de tabaco en radio o televisión	Proporción	Adultos con edad $\geq$ 15 años
Conciencia de haber visto etiquetas de advertencia sanitaria en paquetes de cigarrillos	Proporción	Fumadores actuales con edad $\geq$ 15 años
Planteándose dejar de fumar debido a las etiquetas de advertencia sanitaria en los paquetes de cigarrillos	Proporción	Fumadores actuales con edad $\geq$ 15 años
Conciencia de haber visto anuncios o campañas promocionales de cigarrillos	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que el hábito de fumar tabaco provoca enfermedades graves	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que el hábito de fumar tabaco provoca accidentes cerebrovasculares	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que el hábito de fumar tabaco provoca ataques cardíacos	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que el hábito de fumar tabaco provoca cáncer de pulmón	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que el consumo de tabaco sin humo provoca enfermedades graves	Proporción	Adultos con edad $\geq$ 15 años
Creencia de que la exposición al humo ajeno provoca enfermedades graves a los no fumadores	Proporción	Adultos con edad $\geq$ 15 años
Número de cigarrillos fumados al día (por quienes fuman a diario)	Promedio	Fumadores actuales que fuman cigarrillos a diario con edad $\geq$ 15 años
Tiempo transcurrido desde el abandono del hábito de fumar (en años)	Promedio	Exfumadores con edad $\geq$ 15 años

<b>Indicador</b>	<b>Estimación</b>	<b>Población de base</b>
Gasto mensual en cigarrillos fabricados	Promedio	Fumadores actuales de cigarrillos fabricados con edad $\geq$ 15 años
Edad de iniciación en el hábito de fumar a diario	Promedio	Fumadores que alguna vez han fumado a diario con edad $\geq$ 15 años



Indicador	Estimación (R)	Error cuadrático medio (SE)	Tamaño de la muestra (n)	Efecto del diseño (DEFF o DEFT)	Error relativo (SE/R)	Margen de error (MOE)	Límites de confianza	
							Límite inferior (R-1,96SE)	Límite superior (R+1,96SE)
Última compra de cigarrillos a vendedores callejeros								
Última compra de cigarrillos en quiosco								
Conciencia de haber visto u oído información en contra del consumo de tabaco en radio o televisión								
Conciencia de haber visto etiquetas de advertencia sanitaria en paquetes de cigarrillos								
Planteándose dejar de fumar debido a las etiquetas de advertencia sanitaria en los paquetes de cigarrillos								
Conciencia de haber visto anuncios o campañas promocionales de cigarrillos								
Creencia de que el hábito de fumar tabaco provoca enfermedades graves								
Creencia de que el hábito de fumar tabaco provoca accidentes cerebrovasculares								
Creencia de que el hábito de fumar tabaco provoca ataques cardiacos								
Creencia de que el hábito de fumar tabaco provoca cáncer de pulmón								
Creencia de que el consumo de tabaco sin humo provoca enfermedades graves								
Creencia de que la exposición al humo ajeno provoca enfermedades graves a los no fumadores								
Número de cigarrillos fumados al día (por quienes fuman a diario)								
Tiempo transcurrido desde el abandono del hábito de fumar (en años)								
Gasto mensual en cigarrillos fabricados								
Edad de iniciación en el hábito de fumar a diario								

Notas:

1. Es preciso elaborar tablas relativas a los subgrupos (por ejemplo, género, entorno urbano o rural y región). Véanse las *Definiciones de los indicadores de la encuesta GATS* para consultar las definiciones y los cálculos específicos de los indicadores que figuran a continuación).
2. Los indicadores que figuran en esta tabla se generan a partir del cuestionario básico y pueden adaptarse a cada país en función de la necesidad. Además, los países tal vez deseen elaborar estimaciones para un subconjunto de los indicadores, en lugar de hacerlo para todos, según proceda.
3. Puede que cada país desee calcular el valor de «Deff» o «Deft» basándose en la práctica generalizada en el país o la organización en cuestión.
4. Las tablas anteriores pueden incluirse en el apéndice del informe del país a modo de estimaciones de los errores de muestreo para describir la calidad de los datos.

## C.7 Tasa de cobertura del marco de hogares

### Contexto

Una parte fundamental del proceso de muestreo en una encuesta de hogares es el conteo de los hogares en la fase de muestreo en la que se eligen las direcciones de estos. Se pueden utilizar diferentes metodologías para el conteo de hogares, siendo lo habitual que una parte neta de los hogares quede excluida del conteo resultante y que, debido a ello, la cobertura de la muestra final de hogares sea incompleta. Dado que el sesgo provocado por la infracobertura depende tanto de i) la diferencia agregada entre las principales mediciones del estudio relativas a los hogares abarcados por el marco y los que quedan fuera de este como de ii) el alcance de la falta de cobertura (es decir, el 100% menos la tasa de cobertura de los hogares) en la muestra, resulta útil estimar la tasa de cobertura del marco aplicable a los hogares. La tasa de cobertura del marco (FCR) se define de la siguiente manera:

$$FCR = \frac{N_F}{N} = \frac{\text{número de hogares contabilizados en el marco de los hogares}}{\text{número de hogares en la población}}. \quad (1)$$

### Fuente(s) de datos

Para estos cálculos se necesitarán dos fuentes de datos. Una de ellas es el archivo final de datos ponderados utilizado para el análisis, que sirve para estimar  $N_F$  en la ecuación 1. La otra es la mejor fuente disponible de recuento de hogares de todo el país, que servirá como medida de  $N$  en la ecuación 1. Esta última fuente puede ser el censo más reciente, las actualizaciones más recientes del último censo elaborado con métodos demográficos o una encuesta nacional reciente en la que el tamaño y la calidad de la muestra sean al menos de la misma magnitud que los de la encuesta realizada en el marco de la encuesta GATS.

### Software de cálculo

Ninguno en particular, ya que los cálculos requeridos se pueden programar fácilmente en cualquier entorno de software.

### Cálculo

El número de hogares contabilizados en el marco de hogares ( $N_F$ ) puede estimarse de dos maneras. Ambas requieren conocer algunas de las probabilidades de selección específicas de cada etapa utilizadas para calcular la ponderación básica de cada encuestado. Una de ellas es la estimación ponderada de  $N_F$  a partir de las ponderaciones básicas para el muestreo de «segmentos» en el diseño de la muestra (el *método de suma ponderada*); la otra es la suma de las ponderaciones básicas correspondientes a todos los hogares seleccionados en la muestra de la encuesta GATS, que también constituye una estimación de  $N_F$  (el *método de suma de las ponderaciones de los hogares*).

*Método de suma ponderada:* En la mayoría de los diseños de la encuesta GATS, los segmentos serán las unidades secundarias de muestreo, pero generalmente suelen ser las unidades de muestreo de zonas en las que se realiza el conteo de hogares con el fin de elaborar los marcos de muestreo para la elección de aquellos. Suponiendo que el segmento sea la unidad secundaria de muestreo, se definirá  $W_j$  como la ponderación básica para el  $j^{\circ}$  segmento de la muestra, calculado como el dividido entre el producto de las probabilidades de selección para la unidad primaria de muestreo en la que se encuentra

el segmento y la probabilidad de selección de segmentos dentro de la unidad primaria de muestreo. Si  $M_j$  hace referencia al número de hogares incluidos en el conteo del  $j^o$  segmento de la muestra, entonces la estimación de  $N_F$  con arreglo a este método se calcula entonces de la siguiente manera:

$$\hat{N}_F^{(i)} = \sum_j \overset{\text{Todos los segmentos de la muestra}}{W_j} M_j . \quad (2)$$

**Método de suma de las ponderaciones de los hogares:** Este método utiliza la ponderación básica correspondiente a cada uno de los hogares seleccionados. En el conjunto de hogares seleccionados figuran aquellos que se asignan sobre el terreno para fomentar su participación durante la recogida de datos de la encuesta GATS. Supongamos que  $W_{jk}$  indica la ponderación básica correspondiente al  $k^o$  hogar seleccionado en el  $j^o$  segmento de la muestra. Calculamos  $W_{jk}$  como  $W_j$  dividido entre la probabilidad de selección del  $k^o$  hogar dentro del segmento. La estimación de  $N_F$  con arreglo a este método se calcula entonces de la siguiente manera:

$$\hat{N}_F^{(ii)} = \sum_{j,k} \overset{\text{Todos los segmentos de la muestra}}{W_{jk}} . \quad (3)$$

Las mejores fuentes para obtener un valor adecuado del número real de hogares en la población (N) son las siguientes: 1) el censo de población más reciente (siempre que se considere en general que su grado de calidad es alto), o 2) un recuento estimado de hogares para  $N_F$  (obtenido mediante las ecuaciones 2 o 3) realizado a partir de una encuesta por muestreo reciente y superior desde el punto de vista estadístico (por ejemplo, la muestra maestra de la que se haya seleccionado a su vez aleatoriamente la muestra de la encuesta GATS).

### **Interpretación**

Una estimación de la tasa de cobertura del marco del 95% o superior sería indicativa de un grado de calidad aceptable del conteo de hogares en la muestra de la encuesta GATS.





## Global Adult Tobacco Survey (GATS)