



Subsecretaría de Salud Pública  
Departamento de Epidemiología

# INFORME METODOLÓGICO

## ENCUESTA NACIONAL DE SALUD, SEXUALIDAD Y GÉNERO

### 2022-2023



UNIVERSIDAD DE CHILE



UNIVERSIDAD  
ALBERTO HURTADO

**DESUC**  
Dirección de Estudios  
Sociales

# Índice

<b>EQUIPO DE TRABAJO</b> .....	<b>3</b>
<b>RESUMEN EJECUTIVO</b> .....	<b>5</b>
<b>1. OBJETIVOS</b> .....	<b>8</b>
1.1 Objetivo general .....	8
1.2 Objetivos específicos .....	8
<b>2. METODOLOGÍA</b> .....	<b>9</b>
2.1 Universo, muestra y error.....	9
2.1.1 Universo y población objetivo .....	9
2.1.2 Marco muestral .....	10
2.1.3 Tamaño de la muestra y errores de muestreo .....	11
2.1.4 No respuesta y muestra a seleccionar .....	18
a. Diseño muestral .....	19
2.1.5 Estratificación de la muestra.....	19
2.1.6 Etapas de selección.....	21
<b>3. MUESTRA ALCANZADA Y ERRORES EFECTIVOS</b> .....	<b>28</b>
<b>4. REPORTE DEL TRABAJO DE CAMPO</b> .....	<b>31</b>
4.1 Indicadores de evaluación de proceso de encuestas poblacionales .....	31
4.2 Estado final de terreno .....	33
4.3 Caracterización del trabajo de campo .....	37
4.3.1 Resultados de supervisión.....	37
4.3.2 Resultados de control de calidad e inconsistencias.....	44
4.3.3 Duración de las entrevistas .....	47
<b>5 BIBLIOGRAFÍA</b> .....	<b>49</b>

## **EQUIPO DE TRABAJO**

La Encuesta Nacional de Salud, Sexualidad y Género (ENSSEX) 2022-2023 fue desarrollada por el Ministerio de Salud, a través del Departamento de Epidemiología en directa colaboración con la Universidad de Chile, la Universidad Alberto Hurtado y el *Institut National d'Etudes Démographiques* á Paris.

## **JEFATURA DE DEPARTAMENTO DE EPIDEMIOLOGÍA**

- Dr. Christian García Calavaro

## **DIRECTORA DE PROYECTO**

- Soc. Pía Álvarez Iturriaga

## **CO DIRECCIÓN DE PROYECTO**

- Soc. José Ignacio Alarcón Molina

## **OFICINA DE VIGILANCIA DE ENFERMEDADES NO TRASMISIBLES Y ENCUESTAS POBLACIONALES**

- Mat. Jessica Moya Cortez
- Dra. Clelia Vallebuona Stango
- Soc. Rocío Martínez Gutierrez
- Enf Mat. Sylvina Alvarado Fick
- Odont. Jacqueline Cerda Cordova
- Psi. Bert Kaempfe Vasquez
- Hist. Daniela Solís

## **OFICINA DE ESTUDIOS Y ANÁLISIS ESTADÍSTICOS AVANZADOS**

- Est. Natalia Vergara Mallegas
- Est. Valentina Navarro Muñoz

- Soc. Esteban Navarro Leal

- Est. Nora Díaz Sanzana

#### **OFICINA DE VIGILANCIA DE ENFERMEDADES TRASMISIBLES**

- Mat. Karen Cáceres Burtón

#### **INSTITUT NATIONAL D' ETUDES DÉMOGRAPHIQUES Á PARIS.**

- Michel Bozón

#### **UNIVERSIDAD DE CHILE**

- Irma Palma

#### **UNIVERSIDAD ALBERTO HURTADO**

- Jaime Barrientos

#### **DIRECCIÓN DE ESTUDIOS SOCIALES (DESUC)**

#### **JEFE DE PROYECTO**

- Cristian Ayala

#### **JEFA DE PROYECTO Y MUESTRISTRA**

- Cristina Marchant Araya

#### **JEFA DE TERRENO**

- Vicky Rojas Araya

## RESUMEN EJECUTIVO

Este informe tiene como principal objetivo dar cuenta de los objetivos, metodología y trabajo de campo de la **Encuesta Nacional de Salud, Sexualidad y Género (ENSSEX 2022-2023)**.

La encuesta tiene como principal objetivo: “conocer las características de salud, sexualidad y género de la población de 18 años y más, residente en Chile, con representatividad nacional, regional, por tramos de edad y sexo, para la producción de evidencia científica dirigida al diseño, implementación y evaluación de políticas públicas afines”.

Como antecedente, se tiene que, en Chile, sólo se ha realizado una encuesta específica asociada al comportamiento sexual en 1998 bajo la dirección de la entonces Comisión Nacional del SIDA (CONASIDA). Si bien existen en Chile encuestas como Encuesta Nacional de Salud (ENS), la Encuesta Nacional de Calidad de Vida y Salud (ENCAVI) o Encuesta Nacional de la Juventud (INJUV), estas no permiten recoger con especificidad temas asociados a los cambios de comportamientos en salud sexual y reproductiva que se requieren actualmente para el desarrollo de políticas y programas oportunos a la comunidad desde el Gobierno de Chile.

En este sentido, se plantea la necesidad de levantar una encuesta con representatividad nacional que permita conocer las características de salud, sexualidad y género de la población de 18 años y más, residente en Chile.

Para ello, se desarrolló el pilotaje y llevó a cabo el levantamiento de una encuesta de **20.000 casos considerando un muestreo probabilístico a nivel nacional**, que tenga la capacidad de extrapolar los resultados obtenidos de la encuesta a la población mayor de 18 años de todo Chile.

Respecto a los instrumentos de recolección de datos, se revisó y desarrolló un instrumento basado en un cuestionario elaborado por el equipo de MINSAL y académicos de la U. de Chile, U. Alberto Hurtado y l'Institut National d'Etudes Démographiques à Paris, los años previos. Este instrumento considera una modalidad de aplicación presencial mediante dispositivo electrónico, siendo aplicada por un encuestador/a, con módulos autoaplicados asociados a temáticas sensibles.

Luego de una fase de pretesteo, que consideró la aplicación de **entrevistas cognitivas** y una **prueba de campo** se diagnosticaron problemas y dificultades en la aplicación del instrumento, permitiendo incorporar cambios y mejoras que contribuyeran a capturar mejor la información y responder de mejor manera a los objetivos de la investigación.

Teniendo un cuestionario en su versión definitiva, en el mes de agosto se dio inicio oficial al levantamiento de datos de la ENSSEX 2022-2023. Para ello, se desarrolló una metodología que consideró la definición de universos y tipo de muestra, de diseños muestrales, y de una metodología para la elaboración de factores de expansión que permitan que los datos recogidos en la muestra representen a la población objetivo en su conjunto.

El levantamiento de datos dio como resultado una **muestra efectiva de 20.392 casos**. En el apartado 3 se puede ver la distribución de dicha muestra según variables de interés como sexo, tramos de edad y región. Adicionalmente, en el mismo apartado se presentan los errores muestrales efectivos para dos variables sustantivas del estudio (percepción de salud y prevalencia de haber tenido relaciones sexuales). Dichos errores muestrales efectivos se calcularon bajo un modelo de Diseño Muestral Complejo, que es el efectivamente utilizado en el estudio. Dichos errores fueron congruentes con lo solicitado por bases técnicas de licitación, no superando un error relativo efectivo superior al 30% en ninguna subpoblación de interés.

Por otro lado, en paralelo al levantamiento de datos se llevaron a cabo una serie de procesos que permiten el control de la muestra levantada, así como su calidad técnica. Dado ello, en el Apartado 4 se presenta un reporte del trabajo de campo finalizado. En primera instancia, se definen los indicadores que permiten evaluar el proceso de levantamiento de los datos, para luego exponer en detalle el estado de avance de terreno, donde se aborda el logro e indicadores del control de trabajo de campo de AAPOR. Posteriormente, se realiza una caracterización del proceso de recolección de información en torno a tres ejes; los procesos de supervisión, resultados del control de calidad de los datos recolectados y las inconsistencias encontradas en la base de datos con sus respectivas correcciones, y las principales dificultades que se encontraron durante el trabajo de terreno.

**Tabla Metodológica resumen Encuesta Nacional de Salud, Sexualidad y Género  
(ENSSEX)**

<b>Organismo responsable</b>	Departamento de Epidemiología. Subsecretaría de Salud Pública. Ministerio de Salud. Gobierno de Chile
<b>Organismo ejecutor</b>	DESUC Dirección de Estudios Sociales UC
<b>Población objetivo</b>	Personas de 18 años y más, chilenas o extranjeras, que residen habitualmente en viviendas particulares ocupadas localizadas en zonas urbanas de las dieciséis regiones de Chile.
<b>Diseño muestral</b>	Encuesta presencial en hogares, con diseño probabilístico, estratificado geográficamente y multietápico (cuatro etapas: comuna, manzana, vivienda y persona)
<b>Modo de aplicación</b>	Entrevista personal en el hogar, aplicada por encuestador en cuestionario en dispositivo electrónico (CAPI)
<b>Periodo de trabajo de campo</b>	Agosto a diciembre de 2022
<b>Tamaño Muestral</b>	20.392 casos

<b>Error muestral</b>	Error absoluto de $\pm 0,7$ a nivel nacional, bajo supuesto de Muestreo Aleatorio Simple (MAS), para una proporción del 50% a nivel de confianza de 95%.
<b>Tasas AAPOR</b>	Respuesta (RR1): 51%. Rechazo (REF1): 34%.

## 1. Objetivos

### 1.1 Objetivo general

Conocer las características de salud sexual de la población de 18 años y más, con representatividad nacional, regional, tramos de edad y sexo en Chile, con el fin de producir evidencia para el diseño y evaluación de políticas públicas.

### 1.2 Objetivos específicos

- 1) Generar una línea base de conocimiento científico sobre salud, sexualidad y género en Chile, para la elaboración de un marco interpretativo sobre su evolución histórica para contextualización de las políticas públicas.
- 2) Estimar la prevalencia nacional de los temas salud, sexualidad y género definidos para la encuesta, según sexo, edad, regiones y otros descriptores.
- 3) Describir las conductas, prácticas de riesgo y protectoras de enfermedades/infecciones a partir de las temáticas relacionadas a la salud sexual y reproductiva definidos para la encuesta de la población en estudio, según sexo, edad, regiones y otros descriptores.
- 4) Describir las trayectorias y prácticas sexuales de la población de estudio, según sexo, edad, regiones y otros descriptores.
- 5) Explorar las orientaciones normativas en torno a la sexualidad y el género de la población en estudio y según sexo, edad, distribución geográfica y otros descriptores.



## 2. Metodología

A continuación, se presentan las principales características metodológicas de la implementación y desarrollo de la Encuesta Nacional de Salud, Sexualidad y Género 2022-2023.

### 2.1 Universo, muestra y error

En esta sección se detalla lo referente al universo, población objetivo, diseño muestral, tamaño de la muestra y errores de muestreo asociados al diseño metodológico de la encuesta.

Atendiendo a los requerimientos del estudio, se consideró un diseño muestral de áreas probabilístico, estratificado geográficamente y multietápico, para un tamaño **muestral a nivel nacional objetivo de 20.000 casos efectivos**, distribuidos a proporcionalmente en las dieciséis regiones del país.

#### 2.1.1 Universo y población objetivo

El **universo del estudio** se encuentra definido como: hombres y mujeres de 18 años y más, residentes en hogares particulares de áreas urbanas todas las regiones del país.

La población objetivo de la encuesta entonces corresponde a las **personas de 18 años y más**, chilenas o residentes en Chile, que habitan (residen en la vivienda por más de seis meses, no se encuentran en prisión o en casa de reposo, etc.) en viviendas particulares ocupadas localizadas en **zonas urbanas** de las **dieciséis regiones de Chile**.

Quedan excluidas de la población objetivo las personas que residen en áreas clasificadas como “Áreas de Difícil Acceso” (ADA)<sup>1</sup>, por el Instituto Nacional de Estadísticas<sup>2</sup> y DESUC<sup>3</sup>, ubicadas en 22 comunas.

---

<sup>1</sup> Las comunas de difícil acceso definidas por el INE durante el proceso de PreCenso 2011 corresponden a: Ollagüe, Isla de Pascua y Juan Fernández, Chaitén, Futaleufú, Palena, Lago Verde, Guaitecas, Chile Chico, O'Higgins, Tortel, Torres del Paine, Río Verde, Laguna Blanca, San Gregorio, Primavera, Timaukel, Cabo de Hornos y Antártica

<sup>2</sup> Para mayor información respecto a las Áreas de Difícil Acceso ver documento: “Definición de áreas especiales para el levantamiento censal”, Septiembre 2015, Instituto Nacional de Estadísticas.

<sup>3</sup> A la lista de INE, DESUC agrega las comunas de Camarones, y Putre, debido a la verificación en levantamientos de datos previos de dificultades para acceder a ellas por sus condiciones físico-geográficas (por ejemplo, falta de infraestructura vial, o de medios de transporte públicos y privados constantes que permitan conectar con la zona, dificultades climáticas, entre otras).

### 2.1.2 Marco muestral

Para la ejecución de este diseño muestral se utilizó el marco de manzanas y entidades elaboradas para el Censo 2017<sup>4</sup>. INE puso a disposición en julio de 2018 la cartografía y microdatos a nivel de manzanas urbanas y entidades rurales del Censo 2017. Combinando ambas fuentes se obtiene la ubicación geográfica de manzanas –unidad geográfica básica que para fines estadísticos conforman zonas censales en el área urbana-, número de viviendas y personas que cada una de ellas posee (dato relevante para la selección de manzanas).

El marco fue creado a partir del análisis y actualización de las entidades de población realizadas durante el Precenso 2016 y levantamiento censal 2017, de acuerdo con la división político-administrativa y censal. La población objetivo de este estudio está completamente cubierta por la información disponible a partir del Censo.

Al momento de la selección de manzanas o secciones se excluyeron —por razones operativas—, aquellas con 7 o menos viviendas habitadas (Grupo 0 y Sin viviendas, en Tabla 1). Esta exclusión dejó fuera al 1% de las viviendas catastradas en el marco, pero removió del marco al 9% de sus manzanas. Se estima que esa pérdida de cobertura en el marco es compensada por una mayor eficiencia en los recursos de terreno.

**Tabla 1 Número de manzanas y entidades, viviendas y personas en el Censo 2017 según grupo de tamaño.**

Grupo de tamaño	Rango de total de viviendas	Total, de manzanas y entidades	Dist. de manzanas y entidades	Total, de Viviendas	Dist. de viviendas	Total, de Personas	Dist. de personas
	1	2	3	4	5	6	7
Total		180.499	100%	6.499.355	100%	17.574.003	100%
Sin viviendas	0	2	0%	0	0%	439	0%
Grupo 0	1 a 7	16.242	9%	69.492	1%	252.079	1%
Grupo 1 a 4	8 a 23	79.762	44%	1.249.628	19%	3.565.723	20%
Grupo 5 a 9	24 a 44	53.091	29%	1.665.145	26%	4.889.884	28%
Grupo 10 a 19	45 a 81	18.060	10%	1.046.489	16%	2.907.868	17%
Grupo 20 a 28	82 a 154	8.003	4%	882.281	14%	2.320.225	13%
Grupo 29 y 30	155 y más	5.339	3%	1.586.320	24%	3.637.785	21%

Fuente: Elaboración propia a partir de Censo 2017

Una vez seleccionada la muestra de manzanas/secciones, se realizó una **actualización del total de viviendas seleccionadas en la muestra**. Este procedimiento, denominado “empadronamiento” (*listing*, en inglés), busca actualizar el total de viviendas registrado

<sup>4</sup> Ver Base Cartográfica Censal Junio 2018, INE  
Departamento de Epidemiología – Oficina de Vigilancia de Enfermedades No Transmisibles y Encuestas Poblacionales

en el marco muestral a partir de las observaciones realizadas al momento del trabajo de campo. En concreto, el proceso de empadronamiento se realizó de la siguiente manera:

1. Ubicación del empadronador en la esquina nor-poniente de la manzana seleccionada.
2. Recorrido en sentido de las agujas del reloj de toda la manzana seleccionada (misma modalidad de recorrido utilizado en Censos).
3. Listado del total de edificaciones encontradas durante el recorrido de la manzana.
4. Asignación de Código de Disposición de Vivienda (que resume la naturaleza de la edificación en la manzana)
5. Identificación de viviendas elegibles (viviendas particulares ocupadas).

Además, durante este proceso de empadronamiento de manzanas se catastraron algunas características de estas que podrán utilizarse posteriormente para el ajuste explícito de no respuesta en la elaboración de factores de expansión (ver detalles en informe asociado). Ejemplo de estas características son:

- Estado de edificios y casas de la manzana
- Presencia de basura en la manzana
- Presencia de vandalismo, grafitis o daño deliberado a la propiedad en el sector

De esta forma cambios derivados de nuevas construcciones, demoliciones, cambio de uso de las edificaciones y otros quedan incorporados en el estudio de dos formas:

- Las nuevas construcciones (en manzanas/secciones seleccionadas) tienen probabilidad mayor a cero de ser seleccionadas.
- El factor de expansión a través del ajuste de las probabilidades de selección de las viviendas, incorpora el cambio en los pesos de las unidades seleccionadas para incorporar los crecimientos/decrecimientos entre los registros desactualizados del marco y los registros actualizados del empadronamiento.

A su vez, en paralelo al proceso de empadronamiento, este fue supervisado y digitado en un **100%** en las oficinas de DESUC.

### **2.1.3 Tamaño de la muestra y errores de muestreo**

La Tabla 2 presenta la distribución de la población objetivo por región, según las proyecciones poblacionales al año 2021<sup>5</sup> publicadas por INE con base 2017. En las

---

<sup>5</sup> Notar que se presentan totales poblacionales al año 2021 y no 2022, pues la selección de la muestra se realizó considerando los datos poblacionales del año 2021.  
*Departamento de Epidemiología – Oficina de Vigilancia de Enfermedades No Transmisibles y Encuestas Poblacionales*

primeras dos columnas se puede ver el total y la distribución de toda la población residente en áreas urbanas, que alcanzan los 17.430.714 personas. Centrándonos en la población objetivo, en la columna #3 se ve que las personas mayores de 18 años habitando en áreas urbanas son 13.439.667. Respecto a la distribución de dicha población, en la última columna se ve que se distribuye de manera muy similar a la de toda la población urbana. Así, el 46% se concentra en la región Metropolitana, y un 11% en la región de Valparaíso.

**Tabla 2 Distribución de la población objetivo del estudio: personas mayores de 18 años residentes de zonas urbanas de Chile, según región.**

Región	Personas habitando en áreas urbanas		Personas mayores de 18 años y más en áreas urbanas	
	Población 1	% col 2	Población 3	% col 4
1 Tarapacá	371.788	2%	270.308	2%
2 Antofagasta	680.559	4%	508.680	4%
3 Atacama	289.760	2%	213.064	2%
4 Coquimbo	695.749	4%	523.772	4%
5 Valparaíso	1.814.427	10%	1.417.150	11%
6 O'Higgins	754.522	4%	572.789	4%
7 Maule	850.678	5%	646.246	5%
8 Biobío	1.485.892	9%	1.146.717	9%
9 La Araucanía	731.837	4%	556.150	4%
10 Los Lagos	668.895	4%	507.012	4%
11 Aysén	88.130	1%	64.290	0%
12 Magallanes	165.904	1%	129.261	1%
13 Metropolitana	7.937.816	46%	6.199.548	46%
14 Los Ríos	296.130	2%	228.753	2%
15 Arica y Parinacota	234.536	1%	176.290	1%
16 Ñuble	364.091	2%	279.637	2%
<b>Total</b>	<b>17.430.714</b>	<b>100%</b>	<b>13.439.667</b>	<b>100,0%</b>

Nota: Los totales poblacionales incluyen la población en las Áreas de Difícil Acceso.

Fuente: Proyecciones de población al año 2021 publicadas por INE con base 2017.

Para la distribución muestral se realizó una **distribución a proporción** de los casos, en tanto esto permite **optimizar los errores muestrales** a nivel regional.

Ello quiere decir que la distribución de casos objetivos de la muestra no responde exactamente a la proporción de la población en cada región. Así, en algunas regiones se obtendrá una muestra menor a la que correspondería según su peso poblacional, y en otras, por el contrario, se obtendrá una muestra mayor respecto a la que le correspondería si se sigue la distribución real de personas en dichas regiones.

En la Tabla 3 se muestran los tamaños muestrales objetivo de tres escenarios, junto con sus errores absolutos y relativos calculados bajo un supuesto de Muestreo Aleatorio Simple, a 95% nivel de confianza, con una prevalencia de 50% (varianza máxima). En

concreto, para realizar estos cálculos se utilizaron las siguientes fórmulas según lo definido por Lohr (1999)<sup>6</sup>:

$$e_{abs} = Z \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}}$$
$$e_{rel} = \frac{e_{abs}}{p}$$

Donde,

- Z : Nivel de confianza según distribución Z (95%)
- p : Prevalencia estimada
- n : Tamaño muestral
- N : Tamaño poblacional

Respecto a los tamaños muestrales, en todos los escenarios se consideró una muestra objetivo de 20.000 casos, pero la distribución de estos en regiones es diversa.

En el primer escenario, se tiene una distribución bastante proporcional de los casos, aunque con ajustes leves para lograr un mínimo de casos en regiones con baja proporción de población para lograr reducir los errores muestrales. En el segundo escenario, se tiene una distribución mucho menos proporcional, manteniendo tamaños muestrales por región relativamente similares, y con ello, manteniendo errores muestrales bastante análogos en todas las regiones. En el tercer escenario, se tiene una propuesta intermedia, con una distribución más proporcional de los casos, pero aumentando los mínimos regionales, y disminuyendo el máximo regional (en la región Metropolitana).

En cuanto a los errores muestrales, al compartir el mismo tamaño muestral, los tres escenarios alcanzan un error absoluto de 0,7 y relativo de 1,4 puntos. Al analizar los errores absolutos por región, en el primer escenario el máximo error es de 4,9 puntos en las regiones con 400 casos. En el escenario 2 el máximo es de 3,4 puntos en la región de Magallanes. En el tercer escenario el mayor error absoluto se encuentra en la región de Aysén con 4,4 puntos.

---

<sup>6</sup> Cabe señalar que la fórmula utilizada corresponde al cálculo de error muestral absoluto para poblaciones finitas, que corresponde a tamaños poblacionales menores a 100.000. Se utiliza esta fórmula porque la mayoría de las subpoblaciones se cumple con tener una población mayor a 100.000. Sin embargo, y aunque la región de Aysén tiene un tamaño poblacional menor a ese umbral, por parsimonia, se utiliza la misma fórmula para el cálculo de errores en todas las subpoblaciones.  
*Departamento de Epidemiología – Oficina de Vigilancia de Enfermedades No Transmisibles y Encuestas Poblacionales*

En relación a los errores por sexo, todos los escenarios cuentan con un error absoluto de 1 punto y errores relativos cercanos a los 2 puntos. Respecto a los tramos etarios, también se comparten los mismos errores entre los tres escenarios, teniendo un máximo de error absoluto estimado de 4,2 puntos en el rango etario de 75 a 79 años.

Cabe destacar que en ninguno de los escenarios, y en ninguna subpoblación se estimaron errores relativos superiores a los 20 puntos.

Teniendo estos elementos en consideración, es que en conjunto con la contraparte técnica **se acordó utilizar el escenario muestral 3 como abordaje para el levantamiento de los datos.**

Independiente de ello, se debe mencionar que - de todas formas -, la distribución apropiada de los casos del escenario se corrige mediante ajustes de postmuestreo. En concreto, se elaboró un factor de expansión que mediante el método de postestratificación ajusta las distribuciones y tamaños de la muestra a los estándares poblacionales para dar cuenta de la distribución real de la población en las regiones, tramos etarios y sexo.

De manera adicional, en la Tabla 4 y Tabla 5 se muestran los errores de muestreo absolutos y relativos bajo supuesto de Muestreo Aleatorio Simple, pero con prevalencias del 30% y 70% respectivamente.

**Tabla 3 Tamaños muestrales de los Escenarios 1,2 y 3 con errores absolutos y errores relativos al 95% nivel de confianza bajo Muestreo Aleatorio Simple y prevalencia de 50% según subpoblaciones de interés**

Sub poblaciones	Población	% Población	Prevalencia p= 50%	Escenario 1			Escenario 2			Escenario 3		
				Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS
					4	5		6	7		8	9
País	13.439.667	100%	0,5	<b>20.000</b>	0,7	1,4	<b>20.000</b>	0,7	1,4	<b>20.000</b>	0,7	1,4
Tarapacá	270.308	2%	0,5	<b>450</b>	4,6	9,2	<b>930</b>	3,2	6,4	<b>580</b>	4,1	8,1
Antofagasta	508.680	4%	0,5	<b>720</b>	3,6	7,3	<b>1.080</b>	3,0	6,0	<b>880</b>	3,3	6,6
Atacama	213.064	2%	0,5	<b>400</b>	4,9	9,8	<b>1.040</b>	3,0	6,1	<b>520</b>	4,3	8,6
Coquimbo	523.772	4%	0,5	<b>910</b>	3,2	6,5	<b>1.310</b>	2,7	5,4	<b>1.150</b>	2,9	5,8
Valparaíso	1.417.150	11%	0,5	<b>2.130</b>	2,1	4,2	<b>1.330</b>	2,7	5,4	<b>1.640</b>	2,4	4,8
O'Higgins	572.789	4%	0,5	<b>1.070</b>	3,0	6,0	<b>1.020</b>	3,1	6,1	<b>1.410</b>	2,6	5,2
Maule	646.246	5%	0,5	<b>1.210</b>	2,8	5,6	<b>1.210</b>	2,8	5,6	<b>1.410</b>	2,6	5,2
Biobío	1.146.717	9%	0,5	<b>1.840</b>	2,3	4,6	<b>1.280</b>	2,7	5,5	<b>1.750</b>	2,3	4,7
La Araucanía	556.150	4%	0,5	<b>1.060</b>	3,0	6,0	<b>1.190</b>	2,8	5,7	<b>1.580</b>	2,5	4,9
Los Lagos	507.012	4%	0,5	<b>990</b>	3,1	6,2	<b>1.150</b>	2,9	5,8	<b>1.450</b>	2,6	5,1
Aysén	64.290	1%	0,5	<b>400</b>	4,9	9,8	<b>920</b>	3,2	6,4	<b>490</b>	4,4	8,8
Magallanes	129.261	1%	0,5	<b>400</b>	4,9	9,8	<b>840</b>	3,4	6,7	<b>510</b>	4,3	8,7
Metropolitana	6.199.548	45%	0,5	<b>6.790</b>	1,2	2,4	<b>3.870</b>	1,6	3,1	<b>4.380</b>	1,5	3,0
Los Ríos	228.753	2%	0,5	<b>560</b>	4,1	8,3	<b>980</b>	3,1	6,2	<b>650</b>	3,8	7,7
Arica y Parinacota	176.290	1%	0,5	<b>400</b>	4,9	9,8	<b>870</b>	3,3	6,6	<b>550</b>	4,2	8,3
Ñuble	279.637	2%	0,5	<b>670</b>	3,8	7,6	<b>980</b>	3,1	6,2	<b>1.050</b>	3,0	6,0
Hombre	6.496.501	48%	0,5	<b>9.668</b>	1,0	2,0	<b>9.668</b>	1,0	2,0	<b>9.668</b>	1,0	2,0
Mujer	6.943.166	52%	0,5	<b>10.332</b>	1,0	1,9	<b>10.332</b>	1,0	1,9	<b>10.332</b>	1,0	1,9
18-19	460.507	4%	0,5	<b>685</b>	3,7	7,5	<b>685</b>	3,7	7,5	<b>685</b>	3,7	7,5
20-24	1.289.993	11%	0,5	<b>1.920</b>	2,2	4,5	<b>1.920</b>	2,2	4,5	<b>1.920</b>	2,2	4,5
25-29	1.489.525	11%	0,5	<b>2.217</b>	2,1	4,2	<b>2.217</b>	2,1	4,2	<b>2.217</b>	2,1	4,2
30-34	1.505.509	10%	0,5	<b>2.240</b>	2,1	4,1	<b>2.240</b>	2,1	4,1	<b>2.240</b>	2,1	4,1
35-39	1.310.784	9%	0,5	<b>1.951</b>	2,2	4,4	<b>1.951</b>	2,2	4,4	<b>1.951</b>	2,2	4,4
40-44	1.204.573	9%	0,5	<b>1.793</b>	2,3	4,6	<b>1.793</b>	2,3	4,6	<b>1.793</b>	2,3	4,6
45-49	1.130.166	9%	0,5	<b>1.682</b>	2,4	4,8	<b>1.682</b>	2,4	4,8	<b>1.682</b>	2,4	4,8
50-54	1.063.286	9%	0,5	<b>1.582</b>	2,5	4,9	<b>1.582</b>	2,5	4,9	<b>1.582</b>	2,5	4,9
55-59	1.009.793	8%	0,5	<b>1.503</b>	2,5	5,1	<b>1.503</b>	2,5	5,1	<b>1.503</b>	2,5	5,1
60-64	875.019	6%	0,5	<b>1.302</b>	2,7	5,4	<b>1.302</b>	2,7	5,4	<b>1.302</b>	2,7	5,4
65-69	723.764	5%	0,5	<b>1.077</b>	3,0	6,0	<b>1.077</b>	3,0	6,0	<b>1.077</b>	3,0	6,0
70-74	524.494	4%	0,5	<b>781</b>	3,5	7,0	<b>781</b>	3,5	7,0	<b>781</b>	3,5	7,0
75-79	365.956	3%	0,5	<b>545</b>	4,2	8,4	<b>545</b>	4,2	8,4	<b>545</b>	4,2	8,4
80 y más	486.298	3%	0,5	<b>724</b>	3,6	7,3	<b>724</b>	3,6	7,3	<b>724</b>	3,6	7,3

Nota: los tamaños muestrales para sexo y tramos etarios se estiman en función de la distribución de la muestra que se espera tomando como punto base la distribución de dichas variables en la población objetivo. Por lo tanto, dichas estimaciones pueden variar con la muestra alcanzada finalmente.

**Tabla 4 Tamaños muestrales de los Escenarios 1,2 y 3 con errores absolutos y errores relativos al 95% nivel de confianza bajo Muestreo Aleatorio Simple y prevalencia de 30% según subpoblaciones de interés**

Sub poblaciones	Población	% Población	Prevalencia p= 30%	Escenario 1			Escenario 2			Escenario 3		
				Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS
	1	2	3	4	5	6	7	8	9	10	11	12
País	13.439.667	100%	0,3	<b>20.000</b>	0,6	2,1	<b>20.000</b>	0,6	2,1	<b>20.000</b>	0,6	2,1
Tarapacá	270.308	2%	0,3	<b>450</b>	4,2	14,1	<b>930</b>	2,9	9,8	<b>580</b>	3,7	12,4
Antofagasta	508.680	4%	0,3	<b>720</b>	3,3	11,1	<b>1.080</b>	2,7	9,1	<b>880</b>	3,0	10,1
Atacama	213.064	2%	0,3	<b>400</b>	4,5	15,0	<b>1.040</b>	2,8	9,3	<b>520</b>	3,9	13,1
Coquimbo	523.772	4%	0,3	<b>910</b>	3,0	9,9	<b>1.310</b>	2,5	8,3	<b>1.150</b>	2,6	8,8
Valparaíso	1.417.150	11%	0,3	<b>2.130</b>	1,9	6,5	<b>1.330</b>	2,5	8,2	<b>1.640</b>	2,2	7,4
O'Higgins	572.789	4%	0,3	<b>1.070</b>	2,7	9,1	<b>1.020</b>	2,8	9,4	<b>1.410</b>	2,4	8,0
Maule	646.246	5%	0,3	<b>1.210</b>	2,6	8,6	<b>1.210</b>	2,6	8,6	<b>1.410</b>	2,4	8,0
Biobío	1.146.717	9%	0,3	<b>1.840</b>	2,1	7,0	<b>1.280</b>	2,5	8,4	<b>1.750</b>	2,1	7,2
La Araucanía	556.150	4%	0,3	<b>1.060</b>	2,8	9,2	<b>1.190</b>	2,6	8,7	<b>1.580</b>	2,3	7,5
Los Lagos	507.012	4%	0,3	<b>990</b>	2,9	9,5	<b>1.150</b>	2,6	8,8	<b>1.450</b>	2,4	7,9
Aysén	64.290	1%	0,3	<b>400</b>	4,5	14,9	<b>920</b>	2,9	9,8	<b>490</b>	4,0	13,5
Magallanes	129.261	1%	0,3	<b>400</b>	4,5	14,9	<b>840</b>	3,1	10,3	<b>510</b>	4,0	13,2
Metropolitana	6.199.548	45%	0,3	<b>6.790</b>	1,1	3,6	<b>3.870</b>	1,4	4,8	<b>4.380</b>	1,4	4,5
Los Ríos	228.753	2%	0,3	<b>560</b>	3,8	12,6	<b>980</b>	2,9	9,5	<b>650</b>	3,5	11,7
Arica y Parinacota	176.290	1%	0,3	<b>400</b>	4,5	15,0	<b>870</b>	3,0	10,1	<b>550</b>	3,8	12,7
Ñuble	279.637	2%	0,3	<b>670</b>	3,5	11,6	<b>980</b>	2,9	9,5	<b>1.050</b>	2,8	9,2
Hombre	6.496.501	48%	0,3	<b>9.668</b>	0,9	3,0	<b>9.668</b>	0,9	3,0	<b>9.668</b>	0,9	3,0
Mujer	6.943.166	52%	0,3	<b>10.332</b>	0,9	2,9	<b>10.332</b>	0,9	2,9	<b>10.332</b>	0,9	2,9
18-19	460.507	4%	0,3	<b>685</b>	3,4	11,4	<b>685</b>	3,4	11,4	<b>685</b>	3,4	11,4
20-24	1.289.993	11%	0,3	<b>1.920</b>	2,0	6,8	<b>1.920</b>	2,0	6,8	<b>1.920</b>	2,0	6,8
25-29	1.489.525	11%	0,3	<b>2.217</b>	1,9	6,4	<b>2.217</b>	1,9	6,4	<b>2.217</b>	1,9	6,4
30-34	1.505.509	10%	0,3	<b>2.240</b>	1,9	6,3	<b>2.240</b>	1,9	6,3	<b>2.240</b>	1,9	6,3
35-39	1.310.784	9%	0,3	<b>1.951</b>	2,0	6,8	<b>1.951</b>	2,0	6,8	<b>1.951</b>	2,0	6,8
40-44	1.204.573	9%	0,3	<b>1.793</b>	2,1	7,1	<b>1.793</b>	2,1	7,1	<b>1.793</b>	2,1	7,1
45-49	1.130.166	9%	0,3	<b>1.682</b>	2,2	7,3	<b>1.682</b>	2,2	7,3	<b>1.682</b>	2,2	7,3
50-54	1.063.286	9%	0,3	<b>1.582</b>	2,3	7,5	<b>1.582</b>	2,3	7,5	<b>1.582</b>	2,3	7,5
55-59	1.009.793	8%	0,3	<b>1.503</b>	2,3	7,7	<b>1.503</b>	2,3	7,7	<b>1.503</b>	2,3	7,7
60-64	875.019	6%	0,3	<b>1.302</b>	2,5	8,3	<b>1.302</b>	2,5	8,3	<b>1.302</b>	2,5	8,3
65-69	723.764	5%	0,3	<b>1.077</b>	2,7	9,1	<b>1.077</b>	2,7	9,1	<b>1.077</b>	2,7	9,1
70-74	524.494	4%	0,3	<b>781</b>	3,2	10,7	<b>781</b>	3,2	10,7	<b>781</b>	3,2	10,7
75-79	365.956	3%	0,3	<b>545</b>	3,8	12,8	<b>545</b>	3,8	12,8	<b>545</b>	3,8	12,8
80 y más	486.298	3%	0,3	<b>724</b>	3,3	11,1	<b>724</b>	3,3	11,1	<b>724</b>	3,3	11,1

Nota: los tamaños muestrales para sexo y tramos etarios se estiman en función de la distribución de la muestra que se espera tomando como punto base la distribución de dichas variables en la población objetivo. Por lo tanto, dichas estimaciones pueden variar con la muestra alcanzada finalmente.



**Tabla 5 Tamaños muestrales de los Escenarios 1,2 y 3 con errores absolutos y errores relativos al 95% nivel de confianza bajo Muestreo Aleatorio Simple y prevalencia de 70% según subpoblaciones de interés**

Sub poblaciones	Población	% Población	Prevalencia p= 70%	Escenario 1			Escenario 2			Escenario 3		
				Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS	Tamaño muestra objetivo	Error absoluto 95% MAS	Error relativo 95% MAS
	1	2	3	4	5	6	7	8	9	10	11	12
País	13.439.667	100%	0,7	<b>20.000</b>	0,6	0,9	<b>20.000</b>	0,6	0,9	<b>20.000</b>	0,6	0,9
Tarapacá	270.308	2%	0,7	<b>450</b>	4,2	6,0	<b>930</b>	2,9	4,2	<b>580</b>	3,7	5,3
Antofagasta	508.680	4%	0,7	<b>720</b>	3,3	4,8	<b>1.080</b>	2,7	3,9	<b>880</b>	3,0	4,3
Atacama	213.064	2%	0,7	<b>400</b>	4,5	6,4	<b>1.040</b>	2,8	4,0	<b>520</b>	3,9	5,6
Coquimbo	523.772	4%	0,7	<b>910</b>	3,0	4,2	<b>1.310</b>	2,5	3,5	<b>1.150</b>	2,6	3,8
Valparaíso	1.417.150	11%	0,7	<b>2.130</b>	1,9	2,8	<b>1.330</b>	2,5	3,5	<b>1.640</b>	2,2	3,2
O'Higgins	572.789	4%	0,7	<b>1.070</b>	2,7	3,9	<b>1.020</b>	2,8	4,0	<b>1.410</b>	2,4	3,4
Maule	646.246	5%	0,7	<b>1.210</b>	2,6	3,7	<b>1.210</b>	2,6	3,7	<b>1.410</b>	2,4	3,4
Biobío	1.146.717	9%	0,7	<b>1.840</b>	2,1	3,0	<b>1.280</b>	2,5	3,6	<b>1.750</b>	2,1	3,1
La Araucanía	556.150	4%	0,7	<b>1.060</b>	2,8	3,9	<b>1.190</b>	2,6	3,7	<b>1.580</b>	2,3	3,2
Los Lagos	507.012	4%	0,7	<b>990</b>	2,9	4,1	<b>1.150</b>	2,6	3,8	<b>1.450</b>	2,4	3,4
Aysén	64.290	0%	0,7	<b>400</b>	4,5	6,4	<b>920</b>	2,9	4,2	<b>490</b>	4,0	5,8
Magallanes	129.261	1%	0,7	<b>400</b>	4,5	6,4	<b>840</b>	3,1	4,4	<b>510</b>	4,0	5,7
Metropolitana	6.199.548	46%	0,7	<b>6.790</b>	1,1	1,6	<b>3.870</b>	1,4	2,1	<b>4.380</b>	1,4	1,9
Los Ríos	228.753	2%	0,7	<b>560</b>	3,8	5,4	<b>980</b>	2,9	4,1	<b>650</b>	3,5	5,0
Arica y Parinacota	176.290	1%	0,7	<b>400</b>	4,5	6,4	<b>870</b>	3,0	4,3	<b>550</b>	3,8	5,5
Ñuble	279.637	2%	0,7	<b>670</b>	3,5	5,0	<b>980</b>	2,9	4,1	<b>1.050</b>	2,8	4,0
Hombre	6.496.501	48%	0,7	<b>9.668</b>	0,9	1,3	<b>9.668</b>	0,9	1,3	<b>9.668</b>	0,9	1,3
Mujer	6.943.166	52%	0,7	<b>10.332</b>	0,9	1,3	<b>10.332</b>	0,9	1,3	<b>10.332</b>	0,9	1,3
18-19	460.507	3%	0,7	<b>685</b>	3,4	4,9	<b>685</b>	3,4	4,9	<b>685</b>	3,4	4,9
20-24	1.289.993	10%	0,7	<b>1.920</b>	2,0	2,9	<b>1.920</b>	2,0	2,9	<b>1.920</b>	2,0	2,9
25-29	1.489.525	11%	0,7	<b>2.217</b>	1,9	2,7	<b>2.217</b>	1,9	2,7	<b>2.217</b>	1,9	2,7
30-34	1.505.509	11%	0,7	<b>2.240</b>	1,9	2,7	<b>2.240</b>	1,9	2,7	<b>2.240</b>	1,9	2,7
35-39	1.310.784	10%	0,7	<b>1.951</b>	2,0	2,9	<b>1.951</b>	2,0	2,9	<b>1.951</b>	2,0	2,9
40-44	1.204.573	9%	0,7	<b>1.793</b>	2,1	3,0	<b>1.793</b>	2,1	3,0	<b>1.793</b>	2,1	3,0
45-49	1.130.166	8%	0,7	<b>1.682</b>	2,2	3,1	<b>1.682</b>	2,2	3,1	<b>1.682</b>	2,2	3,1
50-54	1.063.286	8%	0,7	<b>1.582</b>	2,3	3,2	<b>1.582</b>	2,3	3,2	<b>1.582</b>	2,3	3,2
55-59	1.009.793	8%	0,7	<b>1.503</b>	2,3	3,3	<b>1.503</b>	2,3	3,3	<b>1.503</b>	2,3	3,3
60-64	875.019	7%	0,7	<b>1.302</b>	2,5	3,6	<b>1.302</b>	2,5	3,6	<b>1.302</b>	2,5	3,6
65-69	723.764	5%	0,7	<b>1.077</b>	2,7	3,9	<b>1.077</b>	2,7	3,9	<b>1.077</b>	2,7	3,9
70-74	524.494	4%	0,7	<b>781</b>	3,2	4,6	<b>781</b>	3,2	4,6	<b>781</b>	3,2	4,6
75-79	365.956	3%	0,7	<b>545</b>	3,8	5,5	<b>545</b>	3,8	5,5	<b>545</b>	3,8	5,5
80 y más	486.298	4%	0,7	<b>724</b>	3,3	4,8	<b>724</b>	3,3	4,8	<b>724</b>	3,3	4,8

Nota: los tamaños muestrales para sexo y tramos etarios se estiman en función de la distribución de la muestra que se espera tomando como punto base la distribución de dichas variables en la población objetivo. Por lo tanto, dichas estimaciones pueden variar con la muestra alcanzada finalmente.

#### 2.1.4 No respuesta y muestra a seleccionar

Para la selección del tamaño de la muestra objetivo se debió considerar el **fenómeno de la no respuesta** y la tasa de no respuesta asociada.

La tasa de no respuesta corresponde al porcentaje de direcciones en donde **no fue posible** concretar una entrevista, ya sea por dificultades para localizar a los residentes, rechazo de los residentes a participar, encontrarse con un local comercial en vez de una vivienda particular, encontrarse con un sitio eriazo en vez de una vivienda particular, etc.

La no respuesta es normal y esperable dentro del trabajo de encuestas. Siguiendo las mejores prácticas (Valliant, Dever, & Kreuter, 2013), el DESUC utiliza la **técnica de sobremuestreo** para administrar la no respuesta esperada de la encuesta. Esto quiere decir que se seleccionó ex ante, y de una sola vez, una muestra ampliada que considere una pérdida esperada.

Para este estudio se estimó que la tasa de no respuesta a nivel país será aproximadamente de **50%**, que corresponde a una tasa superior a la encontrada en estudios de DESUC con similares características en años anteriores. Se realizó esta corrección por la situación sanitaria en la que sería implementado el estudio y que ha afectado negativamente estudios en terreno que se han efectuado durante 2020 y 2021.

La estimación de participación es imprescindible para determinar el número final de casos que se seleccionarán desde el marco. De antemano se extrajeron una mayor cantidad de casos asumiendo que el 50% de estos no se traducirán en encuestas completas.

Esta técnica le pone, de facto, una cota a la posibilidad de reemplazos de la muestra ya que no se permite la selección de muestras adicionales de reemplazo, lo cual **asegura mantener la aleatoriedad y trazabilidad de la muestra**. Si la manzana no es residencial, si no se establece contacto con la vivienda o es rechazado su ingreso, o bien la persona elegida no responde, se registra como manzana inválida, no respuesta de la vivienda seleccionada, generando así mayor transparencia y rigor en el estado final de cada uno de los casos seleccionados.

En la Tabla 6 se puede ver la distribución muestral objetivo en el escenario 3 con 20.000 casos entre 16 estratos considerando las regiones del país. Para lograr la muestra objetivo, entonces, se debieron seleccionar 40.000 viviendas siguiendo la distribución de la población objetivo en el territorio. Esto implica que —dado que se seleccionan 10 viviendas en cada manzana— se seleccionaron de antemano aproximadamente 4.000 manzanas urbanas.

**Tabla 6 Distribución de muestra objetivo y sobredimensionada.**

Región	Escenario 3		
	Muestra objetivo	Casos sobredimensionados	
	5	6	
1	Tarapacá	580	1.160
2	Antofagasta	880	1.760
3	Atacama	520	1.040
4	Coquimbo	1.150	2.300
5	Valparaíso	1.640	3.280
6	O'Higgins	1.410	2.820
7	Maule	1.410	2.820
8	Biobío	1.750	3.500
9	La Araucanía	1.580	3.160
10	Los Lagos	1.450	2.900
11	Aysén	490	980
12	Magallanes	510	1.020
13	Metropolitana	4.380	8.760
14	Los Ríos	650	1.300
15	Arica y Parinacota	550	1.100
16	Ñuble	1.050	2.100
<b>Total</b>	<b>20.000</b>	<b>40.000</b>	

#### **a. Diseño muestral**

El diseño muestral la encuesta se puede caracterizar como un muestreo de áreas probabilístico, estratificado geográficamente y multietápico, donde se seleccionaron **comunas** en una primera etapa, **manzanas censales** en una segunda etapa, y posteriormente, se seleccionan **viviendas ocupadas** en una tercera etapa y **personas mayores de 18 años** en una cuarta etapa, asegurando así **aleatoriedad total en el proceso**.

A continuación, se describen en detalle la metodología implementada para la selección de la muestra.

##### **2.1.5 Estratificación de la muestra**

La estratificación de una muestra corresponde a la agrupación de los elementos de una población previo a la selección de ella. Uno de los motivos para utilizar un diseño estratificado (ver capítulo 4 en Lohr, 1999) es que éste mejora la eficiencia del muestreo para las estimaciones de la población completa, al agrupar unidades que tienen propiedades similares de media y varianza. La estratificación por tamaño, con una distribución eficiente de casos entre estratos, es un ejemplo de lo señalado.

Este método utiliza una variable de tamaño que esté correlacionada con la variable de interés de la encuesta. Para una discusión en más detalle ver Valliant et al. (2013, cap 3) y Valliant et al. (2000, cap. 6). Así, se estratificó la muestra en dos instancias, con la finalidad de atender dos objetivos distintos.

Por una parte, se fijó como objetivo distribuir la muestra de manera apropiada entre ciertos dominios geográficos de regiones en Chile. Y, además, el diseño muestral también consideró la conformación de estratos de muestreo dentro de cada región que considerara comunas grandes (más de 50.000 habitantes), comunas medianas (entre 50.000 y 20.000 habitantes), y comunas pequeñas (menos de 20.000 habitantes).

Así estos estratos, también denominados “dominios de muestreo” (*sampling domains*), se conformaron entonces tomando como base las 16 regiones del país, además de los estratos por tamaño en su interior. Adicional a ello, se definieron ciertas comunas como relevantes o de Inclusión Forzosa (capitales regionales, comunas del Gran Santiago, o de gran tamaño), que se incluyeron con certeza en la muestra. Teniendo todo ello en consideración, se construyeron un total de **108 estratos de muestreo**, dentro de los cuales se incluyen comunas IF (de Inclusión Forzosa).

Por otra parte, se estratificaron implícitamente las manzanas a seleccionar ordenándolas –dentro de los estratos antes planteados– según la población que contenían<sup>7</sup> y seleccionarlas de acuerdo con un muestreo sistemático con probabilidad proporcional al tamaño. Esto con el objetivo de lograr homogeneizar la probabilidad de selección final de las personas participantes en el estudio.

---

<sup>7</sup> Este dato proviene desde el Marco Muestral utilizado para la selección de la muestra.  
*Departamento de Epidemiología – Oficina de Vigilancia de Enfermedades No Transmisibles y Encuestas Poblacionales*

### 2.1.6 Etapas de selección

A continuación, se describirán las distintas etapas de selección de la muestra.

a) Primera etapa de selección: comunas

Las Unidades Primarias de Muestreo (UPM) fueron las comunas. En concreto, dentro de los estratos de muestreo las comunas —que no fueran de inclusión forzosa (IF)— se escogieron con probabilidad de selección proporcional a su tamaño con método aleatorio (no muestreo sistemático), el cual fue medido por la cantidad de personas mayores de 18 años que habitaran en ellas.

La probabilidad de selección de las comunas quedó definida en:

$$P_h(i) = \begin{cases} 1 & \text{si } i \in h; \text{ es IF} \\ \frac{U_{hi}}{U_h} & \text{si } i \in h; \text{ es no IF} \end{cases}$$

Donde,

- $IF$  : Comuna de Inclusión Forzosa (capitales regionales o comunas de Gran Santiago)
- $h$  : Índice de los estratos de muestreo de primera etapa. En este caso, 16 regiones del país y otros estratos conformados según tamaño de comunas.
- $i$  : Índice de la UPM, correspondiente a la parte urbana de la comuna.
- $U_{hi}$  : Total, de personas mayores de 18 años que posee la UPM  $i$ , del estrato  $h$ , según el Marco de muestreo, sin considerar manzanas de 7 viviendas o menos.
- $U_h$  : Total, de personas mayores de 18 años que posee el estrato  $h$ , según el Marco de muestreo, sin considerar manzanas de 7 viviendas o menos.

De esta manera, la muestra de comunas seleccionadas permitió obtener una mayor presencia de comunas más grandes, aunque asegurando la inclusión de comunas más pequeñas en tanto se encontraron dentro de los estratos de muestreo.

En total, se seleccionaron 126 comunas. El listado de ellas se encuentra en la Tabla 7. Adicionalmente, se identifica a aquellas comunas que resultaron de Inclusión Forzosa, ya sea porque correspondían a capitales regionales (16), como por pertenecer al Gran Santiago (54), o poseer más de 150.000 habitantes mayores de 18 años (2).

**Tabla 7 Listado de comunas seleccionadas**

<b>Tarapacá</b>	<b>Maule</b>	<b>Los Lagos</b>	<b>Metropolitana</b>
Alto Hospicio	Cauquenes	Ancud	Cerrillos IF
Iquique IF	Colbún	Frutillar	Cerro Navia IF
Pozo Almonte	Curicó	Llanquihue	Conchalí IF
<b>Antofagasta</b>	Linares	Los Muermos	El Bosque IF
Antofagasta IF	Maule	Osorno	Estación Central IF
Calama	Pencahue	Puerto Montt IF	Huechuraba IF
Taltal	Rauco	Puerto Varas	Independencia IF
Tocopilla	Talca IF	San Juan De La Costa	La Cisterna IF
<b>Atacama</b>	<b>Bíobío</b>	<b>Aysén</b>	La Florida IF
Caldera	Cabrero	Aysén	La Granja IF
Copiapó IF	Chiguayante	Cochrane	La Pintana IF
Diego De Almagro	Concepción IF	Coyhaique IF	La Reina IF
Vallenar	Curanilahue	<b>Magallanes</b>	Las Condes IF
<b>Coquimbo</b>	Florida	Natales	Lo Barnechea IF
Andacollo	Hualpén	Porvenir	Lo Espejo IF
Coquimbo IF	Hualqui	Punta Arenas IF	Lo Prado IF
La Serena	Los Ángeles	<b>Los Ríos</b>	Macul IF
Ovalle	Penco	Los Lagos	Maipú IF
Salamanca	Talcahuano	Paillaco	María Pinto
<b>Valparaíso</b>	Tucapel	Valdivia IF	Melipilla
Limache	Yumbel	<b>Arica y Parinacota</b>	Ñuñoa IF
Quilpué	<b>Araucanía</b>	Arica IF	Pedro Aguirre Cerda IF
San Antonio	Angol	<b>Ñuble</b>	Peñalolén IF
San Esteban	Cholchol	Chillán IF	Pirque
Santo Domingo	Curacautín	Chillán Viejo	Providencia IF
Valparaíso IF	Freire	Coelemu	Pudahuel IF
Viña Del Mar IF	Gorbea	Ñiquén	Puente Alto IF
<b>O'higgins</b>	Lonquimay	Pinto	Quilicura IF
Chépica	Lumaco	San Carlos	Quinta Normal IF
Codegua	Nueva Imperial	Treguaco	Recoleta IF
Graneros	Padre Las Casas		Renca IF
Peralillo	Perquenco		San Bernardo IF
Pichilemu	Pitrufuquén		San Joaquín IF
Rancagua IF	Renaico		San Miguel IF
Rengo	Temuco IF		San Ramón IF
Santa Cruz			Santiago IF
			Vitacura IF

b) Segunda etapa de selección: manzanas

Las unidades secundarias de muestreo (USM) fueron las manzanas censales ubicadas en las áreas urbanas de las comunas seleccionadas del país.

Para este estudio se utilizó un método de selección sistemática con probabilidad proporcional al tamaño de las manzanas, medido éste como la cantidad de personas que habitan en la manzana según los registros del Marco muestral. Con esto se buscó uniformar la probabilidad de selección de las personas en esta muestra.

En concreto, la probabilidad de selección de las manzanas quedó definida en:

$$P_{hi}(j|i) = \frac{n_{hi} M_{hij}}{U_{hi}}$$

Donde,

- $j$  : Índice de la USM, correspondiente a manzana censal.
- $n_{hi}$  : Total, de USM a seleccionar en la UPM  $i$ , del estrato  $h$ .
- $U_{hi}$  : Total, de personas mayores de 18 años que posee la UPM  $i$ , del estrato  $h$ , según el Marco de muestreo, sin considerar manzanas de 7 viviendas o menos.
- $M_{hij}$  : Total, de personas mayores de 18 años que posee la USM  $j$ , en la UPM  $i$ , del estrato  $h$ , según el Marco de muestreo, sin considerar manzanas de 7 viviendas o menos.

En total, se seleccionaron 3.997 manzanas urbanas en distribuidas en las 16 regiones de Chile. En la Tabla 8 se encuentra el detalle del número de manzanas seleccionadas por comuna seleccionada.

**Tabla 8 Número de manzanas seleccionadas por comuna**

<b>Tarapacá</b>		<b>Maule</b>		<b>Los Lagos</b>		<b>Metropolitana</b>	
Alto Hospicio	36	Cauquenes	56	Ancud	30	Cerrillos	10
Iquique	74	Colbún	5	Frutillar	7	Cerro Navia	17
Pozo Almonte	5	Curicó	49	Llanquihue	14	Conchalí	16
<b>Antofagasta</b>		Linares	27	Los Muermos	16	El Bosque	20
Antofagasta	110	Maule	45	Osorno	71	Estación Central	20
Calama	48	Pencahue	19	Puerto Montt	105	Huechuraba	12
Taltal	11	Rauco	2	Puerto Varas	43	Independencia	13
Tocopilla	7	Talca	80	San Juan De La Costa	4	La Cisterna	12
<b>Atacama</b>		<b>Bíobío</b>		<b>Aysén</b>		La Florida	47
Caldera	14	Cabrero	18	Aysén	26	La Granja	15
Copiapó	60	Chiguayante	44	Cochrane	9	La Pintana	21
Diego De Almagro	11	Concepción	60	Coyhaique	63	La Reina	12
Vallenar	18	Curanilahue	34	<b>Magallanes</b>		Las Condes	40
<b>Coquimbo</b>		Florida	3	Natales	13	Lo Barnechea	12
Andacollo	9	Hualpén	67	Porvenir	4	Lo Espejo	13
Coquimbo	80	Hualqui	15	Punta Arenas	85	Lo Prado	13
La Serena	76	Los Ángeles	37	<b>Los Ríos</b>		Macul	16
Ovalle	32	Penco	10	Los Lagos	16	Maipú	66
Salamanca	32	Talcahuano	39	Paillaco	39	María Pinto	4
<b>Valparaíso</b>		Tucapel	5	Valdivia	75	Melipilla	73
Limache	50	Yumbel	19	<b>Arica y Parinacota</b>		Ñuñoa	29
Quilpué	66	<b>Araucanía</b>		Arica	110	Pedro Aguirre Cerda	13
San Antonio	53	Angol	37	<b>Ñuble</b>		Peñalolén	30
San Esteban	23	Cholchol	7	Chillán	107	Pirque	12
Santo Domingo	6	Curacautín	14	Chillán Viejo	32	Providencia	20
Valparaíso	60	Freire	7	Coelemu	26	Pudahuel	28
Viña Del Mar	69	Gorbea	12	Ñiquén	2	Puente Alto	70
<b>O'higgins</b>		Lonquimay	3	Pinto	22	Quilicura	25
Chépica	21	Lumaco	14	San Carlos	20	Quinta Normal	14
Codegua	7	Nueva Imperial	44	Treguaco	2	Recoleta	21
Graneros	33	Padre Las Casas	21			Renca	18
Peralillo	12	Perquenco	5			San Bernardo	36
Pichilemu	16	Pitrufquén	20			San Joaquín	12
Rancagua	98	Renaico	6			San Miguel	14
Rengo	65	Temuco	126			San Ramón	11
Santa Cruz	30					Santiago	58
						Vitacura	11



c) Tercera etapa de selección: viviendas

Las unidades terciarias de muestreo (UTM) corresponden a las viviendas en las manzanas ubicadas en las zonas urbanas del país. Se definió un número fijo de **10 viviendas seleccionadas cada manzana**.

A diferencia de las USM, las viviendas no son directamente identificables en el marco muestral que se dispone. Esto quiere decir que, previo a implementar la selección de las viviendas se debió “construir” un marco muestral, en terreno, para las USM seleccionadas en la muestra. El proceso de identificación y listado de las viviendas en las manzanas seleccionadas en la muestra se denomina “empadronamiento” (*listing*). Para realizar este procedimiento, el empadronador o encuestador contó con una Hoja de Ruta 1 que guía el proceso, y un mapa de la manzana seleccionada (el cual es elaborado por DESUC).

Durante este proceso de empadronamiento (el cual fue supervisado en su 100%) el encuestador hace un recorrido completo por la manzana seleccionada, lista las direcciones de todos los sitios y asigna un código único de acuerdo con el tipo de edificación que se encuentra en cada sitio. Las manzanas con edificaciones de uso habitacional particular son consideradas “elegibles” para el estudio. Las manzanas que no son de uso residencial (ej. sitios eriazos, plazas, establecimientos comerciales, educacionales, etc.) se consideran “no elegibles”.

En las **manzanas empadronadas** (y con viviendas elegibles) el mecanismo de selección consistió en los siguientes pasos:

- identificar en el listado las viviendas “elegibles” en la manzana empadronada;
- calcular el total de viviendas elegibles y dividir por 10 para obtener el periodo del salto sistemático;
- seleccionar al azar una vivienda en cada manzana; y
- partiendo de la vivienda seleccionada, identificar las siguientes 9 viviendas a entrevistar aplicando el salto sistemático.

Este método, denominado **muestreo sistemático**, asignó igual probabilidad de selección a cada vivienda en la manzana seleccionada, la que queda definida según:

$$P_{hij}(k|j) = \frac{m_{hij}}{M_{hij}^*}$$

Donde,

$k$  : Índice de la UTM, corresponde a la vivienda.

$m_{hij}$  : Total, de viviendas a seleccionar en la USM  $j$ , en la UPM  $i$ , del estrato  $h$ . Esto se encuentra definido por diseño en 10.

$M_{hij}^*$  : Total, de viviendas particulares elegibles (residenciales ocupadas) que posee la USM  $j$ , en la UPM  $i$ , del estrato  $h$ , según lo registrado en el proceso de empadronamiento

Este proceso de selección se realizó en paralelo a la ejecución de entrevistas dentro de las viviendas seleccionadas. Finalmente, al terminar del proceso de empadronamiento, y habiéndose realizado las supervisiones al 100% de estos, se procedió a la digitación de una base de empadronamiento, la cual tendrá tantas filas como manzanas seleccionadas en la muestra.

d) Cuarta etapa de selección: personas

Las unidades últimas de muestreo (UUM) en este estudio corresponden a las personas mayores de 18 años o más que son residentes habituales de las viviendas seleccionadas. La persona seleccionada a este nivel será el único informante por entrevistar.

Tampoco se dispone de un “listado” de las unidades (las personas) que cumplen con estos requerimientos, por lo tanto, se debió implementar un procedimiento que permitiera identificar a las personas elegibles en cada vivienda y luego seleccionar al informante.

Para la encuesta se utilizó un método totalmente probabilístico, donde se seleccionó a una persona elegible dentro de la vivienda del siguiente modo:

1. El encuestador enlista en el software de aplicación de encuestas en su dispositivo electrónico a todas las personas de la vivienda, registrando datos como nombre de pila, sexo, edad.
2. La función de la edad de las personas, el software definirá la característica de elegible o no elegible para cada una de las personas. Cabe señalar que, en consonancia con lo estipulado en las Bases de Licitación, se excluyeron del listado y, por tanto, de la categoría de elegibles a aquellas personas que mostraran actitudes violentas con los encuestadores. Esto queda debidamente registrado en los códigos de disposición final de casos.
3. El software automáticamente escoge entre las personas elegibles, a una persona que es la seleccionada para participar del estudio.

Adicionalmente, y atendiendo a lo solicitado por la contraparte técnica, se decidió sobremuestrear a la población de menor de 30 años, en tanto constituye una subpoblación de interés en elementos relativos a diversidad sexual e Infecciones de Transmisión Sexual (ITS). Dado ello, dicha población tuvo el doble de probabilidad de ser escogidos.

Independiente de ello, este mecanismo de selección se trata de un método probabilístico, ya que asigna una probabilidad de selección conocida a cada persona en la vivienda, preservando el carácter probabilístico del diseño muestral aplicado en las etapas de selección previas. En específico, la probabilidad de selección de la persona queda definido por:

$$P_{hijk}(l|k) = \frac{q_l}{\sum q_l}$$

Donde,

$q_l$  : Es 2 cuando persona tiene <30 años, y 1 cuando persona tiene  $\geq$  30 años.

Dado que este proceso se encuentra incorporado en el software de aplicación, esto asegura que no existan errores en el proceso de selección de la persona.

### 3. Muestra alcanzada y errores efectivos

Los tamaños muestrales y las estimaciones de errores muestrales recién descritos en la Tabla 3, Tabla 4 y Tabla 5 —que suponen un muestreo aleatorio simple— sirven como referencia ya que: 1) la muestra efectivamente alcanzada difirió de lo definido previamente, y 2) el diseño muestral implementado para la ENSSEX 2022-2023 es probabilístico, estratificado y multietápico, lo que se conoce como un diseño muestral complejo.

Dado eso, para poder estimar los errores absolutos y relativos asociados a diversas variables de interés del estudio es necesario considerar los tamaños muestrales efectivos, y de un factor corrector llamado *efecto diseño*.

El *efecto diseño* es una medida de la eficiencia relativa de un diseño muestral cualquiera con respecto al Muestreo Aleatorio Simple (MAS) del mismo tamaño muestral (Kish, 1949). En términos estadísticos, esta eficiencia se puede expresar como la razón entre la varianza asociada a un muestreo complejo y la varianza asociada a un muestreo aleatorio simple.

En términos formales, el efecto diseño de una media muestral  $\bar{Y}$  se puede expresar como (para mayores detalles revisar (Valliant, Dever, & Kreuter, 2013)):

$$Def_{\bar{Y}} = \frac{V(\bar{Y})}{V_{MAS}(\bar{Y})}$$

Donde,

$V(\bar{Y})$  : Varianza de la media bajo Diseño Muestral Complejo

$V_{MAS}(\bar{Y})$  : Varianza de la media bajo Muestreo Aleatorio Simple.

Ahora bien, existe un efecto diseño para cada variable en una encuesta. A modo de ilustración para este informe, se utilizó el efecto diseño asociado a dos variables: percepción de tener “buena o muy buena salud”, y la declaración de haber tenido “relaciones sexuales voluntarias”. Considerando ello, se calculó del efecto diseño para cada región, sexo y tramo etario de manera independiente.

Con ese dato, las columnas #3, #4, #6 y #7 de la Tabla 9 presentan una estimación de los errores de muestreo absolutos y relativos observados para ENSSEX 2022-2023 al considerar un diseño muestral complejo.

Para calcular dichos errores bajo un diseño muestral complejo con un efecto diseño similar, se utiliza la misma fórmula de errores muestrales absolutos, pero corrigiendo el tamaño muestral según el efecto diseño. De esta manera, podemos expresar el impacto del efecto diseño en el tamaño muestral de la siguiente manera (Groves, y otros, 2011):

$$n_{eff} = \frac{n}{Def(\bar{y})}$$

Donde,

$n$  : Tamaño muestral

$Def(\bar{y})$  : Efecto diseño de la media bajo Diseño Muestral Complejo.

Así, los errores quedarían reflejados en:

$$eDMC_{abs} = Z \sqrt{\frac{p(1-p)}{n_{eff}}} \sqrt{\frac{N - n_{eff}}{N - 1}}$$

En concreto, se ve que para la variable asociada a percepción de salud se tiene un error absoluto nacional de 1,3 y relativo de 2,0. El error absoluto mayor se encuentra en la región de Magallanes con 7,1 y el relativo mayor en el grupo etario de 80 años y más, con un valor de 16,5. Para el caso de la prevalencia de relaciones sexuales voluntarias, se tiene un error absoluto y relativo nacional de 0,7. El mayor error absoluto se encuentra en la región de Aysén con 7,9, y el mayor error relativo en la misma región con 8,6.

Finalmente, también es importante notar que tanto en ninguna de las celdas de estimación (región, sexo y tramos etarios) se alcanza un error relativo superior al 30% bajo el diseño muestral complejo empleado en ENSSEX 2022-2023.

**Tabla 9 Errores muestrales absolutos y relativos asociados a muestra efectiva bajo Diseño Muestral Complejo**

Sub poblaciones	Percepción buena o muy buena salud				Ha tenido relaciones sexuales voluntarias		
	N muestra l	P	Error absoluto 95% DMC	Error relativo 95% DMC	P	Error absoluto 95% DMC	Error relativo 95% DMC
	1	2	3	4	5	6	7
<b>País</b>	<b>20.392</b>	<b>0,66</b>	<b>1,3</b>	<b>2,0</b>	<b>0,93</b>	<b>0,7</b>	<b>0,7</b>
Tarapacá	616	0,65	2,7	4,1	0,84	6,0	7,2
Antofagasta	892	0,72	4,0	5,5	0,95	1,6	1,7
Atacama	530	0,68	4,6	6,7	0,97	0,9	0,9
Coquimbo	1.176	0,72	4,1	5,7	0,95	2,1	2,2
Valparaíso	1.713	0,70	3,3	4,7	0,95	1,2	1,3
O'Higgins	1.422	0,64	6,5	10,2	0,95	1,7	1,8
Maule	1.451	0,57	5,1	8,9	0,90	3,3	3,7
Biobío	1.869	0,76	2,5	3,3	0,89	1,7	1,9
La Araucanía	1.495	0,62	4,3	6,9	0,96	0,9	0,9
Los Lagos	1.526	0,68	4,4	6,5	0,88	4,4	5,1
Aysén	502	0,74	7,0	9,4	0,91	7,9	8,6
Magallanes	451	0,75	7,1	9,4	0,98	1,4	1,5
Metropolitana	4.392	0,63	2,4	3,9	0,94	1,1	1,2
Los Ríos	663	0,69	5,7	8,2	0,94	2,2	2,3
Arica y Parinacota	524	0,65	5,9	9,1	0,90	6,0	6,6
Ñuble	1.170	0,68	5,5	8,0	0,91	3,2	3,5
Hombre	6.838	0,72	1,7	2,3	0,94	0,9	1,0
Mujer	13.554	0,60	1,7	2,9	0,93	0,8	0,8
18-24	3.028	0,80	2,1	2,6	0,71	2,9	4,1
25-29	2.189	0,77	2,6	3,4	0,95	1,1	1,2
30-34	2.129	0,76	3,1	4,0	0,96	1,3	1,3
35-39	1.695	0,76	2,8	3,6	0,97	1,3	1,3
40-44	1.502	0,69	3,7	5,4	0,97	1,1	1,1
45-49	1.599	0,70	3,5	5,1	0,98	1,1	1,2
50-54	1.625	0,64	3,8	6,0	0,97	1,0	1,0
55-59	1.550	0,57	4,7	8,2	0,97	1,2	1,2
60-64	1.453	0,50	4,1	8,1	0,95	2,0	2,1
65-69	1.317	0,43	5,0	11,5	0,98	0,8	0,8
70-74	1.000	0,45	5,1	11,4	0,97	1,2	1,2
75-79	735	0,43	4,9	11,3	0,94	2,9	3,1
80 y más	570	0,41	6,8	16,5	0,94	2,5	2,6

## 4. Reporte del trabajo de campo

A continuación, se entrega un reporte del trabajo de campo con el levantamiento de datos finalizado. Para ello, en primer lugar se abordan los indicadores que permiten evaluar el proceso de terreno, para luego entregar el estado final del levantamiento de datos, y por último, caracterizar dicho proceso.

### 4.1 Indicadores de evaluación de proceso de encuestas poblacionales

La información del levantamiento de datos en terreno es realizada de acuerdo con los códigos recomendados por AAPOR (*American Association of Public Opinion Research*, ver en [www.aapor.org](http://www.aapor.org)). Esta asociación, estudió una serie de clasificaciones para dar cuenta de los códigos de disposición final de casos, generando una clasificación en cuatro grandes grupos para las encuestas presenciales en hogar (AAPOR, 2016)<sup>8</sup>:

- Entrevistas (logradas): Pueden ser completas o parciales.
- Casos elegibles que no han sido entrevistados: Se subdividen en tres categorías generales, rechazos, no contactos y otros no elegibles. Los rechazos refieren a aquellos que han declinado su participación en la encuesta, el no contacto puede darse por la dificultad de entrar a la vivienda o que los entrevistados no se encontraban disponibles durante las visitas. Otros casos no elegibles, refieren a un entrevistado elegible que no ha rechazado la entrevista, pero ésta no se puede obtener dada su inhabilidad mental o física, problemas de lenguaje, muerte u otras razones.
- Casos de elegibilidad desconocida: Refiere a situaciones donde no se sabe si una vivienda elegible existe y aquellos donde la vivienda elegible existe, pero no se tiene información respecto a si existe un respondiente elegible, entre otros.
- Casos no elegibles: Estos casos incluyen a modo general viviendas fuera de muestra, viviendas no residenciales, viviendas vacías, y viviendas con respondientes no elegibles.

Estas categorías se clasifican generalmente de la siguiente manera:

- E= Entrevista completa
- P= Entrevista parcial

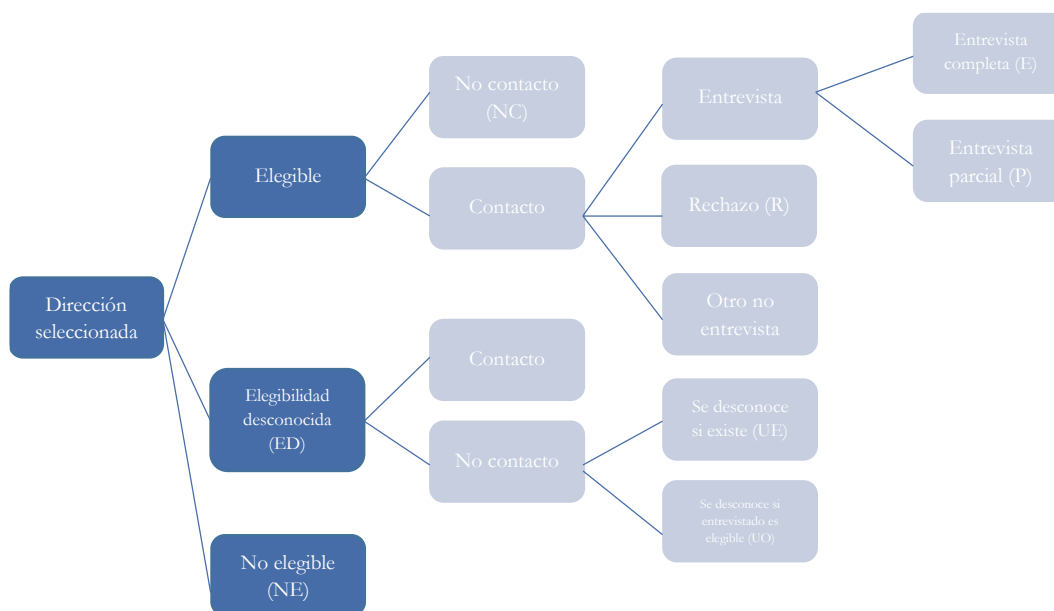
---

<sup>8</sup> Las clasificaciones presentadas refieren de manera general a los códigos AAPOR y su numeración da cuenta de la referencia utilizada.

- R= Rechazo o quiebre de entrevista por parte de entrevistado.
- NC= No contacto con la unidad.
- O= Otro motivo de no entrevista.
- UE= Se desconoce si el hogar es elegible
- UO= Se desconoce si el entrevistado es elegible u existe otra razón de elegibilidad desconocida.

Estos códigos suelen organizarse de acuerdo con lo estipulado en el siguiente diagrama.

**Figura 1. Clasificación de Códigos de Disposición Final de Casos**



Bajo estos códigos se generó el seguimiento de terreno respecto a la muestra objetivo, lo cual se abordará en la siguiente subsección.

Los códigos de disposición final de casos permiten calcular cuatro tipos de tasas según los estándares de AAPOR, que serán entregados en la siguiente sección de manera preliminar, y de manera definitiva al finalizar el levantamiento de terreno:

- **Tasa de respuesta (RR1):** Corresponde al número de entrevistas completas y parciales, partido por el total de casos elegibles seleccionados. En la versión 2016 de AAPOR, se considera en RR2 la elegibilidad desconocida como denominador de esta tasa, para una versión conservadora.
- **Tasa de rechazo (REF1):** Corresponde a la proporción de rechazos (se considera rechazo hogar, persona y encuestas incompletas) del total de casos elegibles (se



excluyen viviendas deshabitadas, casas de veraneo, locales comerciales, es decir, los casos no elegibles).

- **Tasa de cooperación (COOP1):** Corresponde a la proporción de casos en donde se realizó una entrevista completa, del total de casos contactados de la muestra.
- **Tasa de contacto (CON1):** Corresponde a la proporción de casos en donde se llegó a algún tipo de contacto (ya sea encuesta completa o rechazo), del total de casos elegibles de la muestra.

## 4.2 Estado final de terreno

Considerando lo abordado previamente, en la Tabla 10 se muestran los resultados de terreno finales según región.

Las primeras dos columnas corresponden a la muestra objetivo y sobredimensionada, respectivamente.

Luego, en la tercera columna se presentan los casos de encuestas aplicadas, que corresponden a un total de **20.392 casos**. Dentro de estos casos se encuentran exclusivamente encuestas que superaron todos los procesos de supervisión (ver sección 4.3.1). Respecto a la distribución, la región Metropolitana es la que presenta una mayor cantidad de casos aplicados, con 4.392, lo que se corresponde con el tamaño de su muestra objetivo. En Magallanes, por el contrario, se tiene la menor cantidad, con 451 casos finalmente aplicados, lo que también se vincula con su baja muestra objetivo en términos comparativos.

En la cuarta columna se presentan los casos donde existió algún tipo de rechazo, ya sea a nivel de hogar, como a nivel de persona seleccionada. Como se ve, se registraron un total de 13.555 rechazos. El máximo número de rechazos registrados se encuentra en la región Metropolitana con 2.965 casos, y el mínimo en la región de Tarapacá, con 263.

Continuando, en la quinta columna se muestra los no contactos<sup>9</sup>, que corresponden a casos donde no se ha logrado establecer el contacto con la vivienda seleccionada, o con la persona seleccionada. En esta situación se encontraron 2.653 casos de la muestra sobredimensionada. El máximo de este número se encuentra en la región de Metropolitana con 573 casos, lo que se encuentra en línea con el tamaño de la muestra objetivo, y también con las dinámicas de respuesta propias de esta región. El mínimo, por el contrario, está en la región de Coquimbo, con 71 casos donde no se logró establecer contacto con la unidad (vivienda o persona) seleccionada.

---

<sup>9</sup> Se incluyen en los no contacto también aquellos casos que fueron rechazados por procesos de supervisión.  
*Departamento de Epidemiología – Oficina de Vigilancia de Enfermedades No Transmisibles y Encuestas Poblacionales*

Por otro lado, en la sexta columna se muestran los casos no elegibles, que corresponden a viviendas que no cumplen con el criterio de elegibilidad del estudio, el cual se define en contar con residentes habituales mayores de 18 años. En esta situación se encontraron 18 casos, teniendo su máximo en la región de Atacama con 9 casos.

Finalmente, en la séptima columna se presentan los casos con elegibilidad desconocida. Esta situación agrupa a casos donde se intentó realizar contacto con la vivienda seleccionada, pero se desconoce la elegibilidad de la misma (por ejemplo, no se tiene información sobre el uso de edificaciones de la dirección seleccionada), o también, casos que no fueron enviados a terreno, y por tanto, se desconoce la elegibilidad de dichas unidades seleccionadas. En esta situación se encuentra un total de 3.352 casos. La mayor cantidad de estos se encuentra en – como es de esperar – la región Metropolitana con 810 casos en este estado. El mínimo, por su parte, está en la región de Aysén con 60 casos donde se desconoce información de la elegibilidad de la vivienda.

**Tabla 10 Resultados de terreno según región**

Región	Muestra objetivo	Muestra sobre-dimensión	Aplicadas	Rechazos	No contacto	No elegibles	Elegibilidad desconocida
	1	2	3	4	5	6	7
1 Tarapacá	580	1.150	616	263	61	0	210
2 Antofagasta	880	1.760	892	685	73	0	110
3 Atacama	520	1.030	530	269	72	9	150
4 Coquimbo	1.150	2.290	1.176	903	71	0	140
5 Valparaíso	1.640	3.270	1.713	1.224	213	0	120
6 O'Higgins	1.410	2.820	1.422	838	347	2	211
7 Maule	1.410	2.830	1.451	1.070	109	0	200
8 Biobío	1.750	3.510	1.869	1.230	221	0	190
9 La Araucanía	1.580	3.160	1.495	980	284	0	401
10 Los Lagos	1.450	2.900	1.526	1.018	114	2	240
11 Aysén	490	980	502	329	89	0	60
12 Magallanes	510	1.020	451	334	105	0	130
13 Metropolitana	4.380	8.740	4.392	2.965	573	0	810
14 Los Ríos	650	1.300	663	380	93	4	160
15 Arica y Parinacota	550	1.100	524	363	82	1	130
16 Ñuble	1.050	2.110	1.170	704	146	0	90
<b>Total</b>	<b>20.000</b>	<b>39.970</b>	<b>20.392</b>	<b>13.555</b>	<b>2.653</b>	<b>18</b>	<b>3.352</b>

Teniendo estos resultados de terreno expuestos, es que en la Tabla 11 se muestra el cálculo de las Tasas AAPOR según región, además de la tasa de logro.

Respecto a esta última, que se presenta en la columna 3, corresponde a la división entre las encuestas completas, y la muestra objetivo. Al finalizar el terreno, se tuvo una **tasa de logro de 101%**. En cuanto a la distribución por regiones, se tiene un máximo de 111% en la región del Ñuble. El mínimo se encuentra en la región de Aysén, con 88 %.

Pasando ahora a tasas AAPOR, en la columna 4 se muestra la **tasa de respuesta**, que corresponde a la cantidad de encuestas completas divididas por el total de casos elegibles (donde se asume que los casos de elegibilidad desconocida corresponden a casos elegibles). Esta tasa fue del **51% en el estudio general**, teniendo el máximo en Ñuble, con 55%, y un mínimo de 44% en Aysén.

En la quinta columna, por su parte, se expone la **tasa de contacto**, que corresponde a la división entre casos contactados (ya sea con encuestas completas o rechazos), divididos por el total de casos elegibles. Esta tasa es de un **85% a nivel nacional**, teniendo el máximo en Coquimbo con 91%, y el mínimo en Tarapacá con 76%.

Continuando, en la sexta columna se presenta la **tasa de cooperación**, construida a partir de la división de casos aplicados, dividido por la cantidad de casos contactados. **A nivel general, esta tasa es de 60%**, teniendo su máximo de 70% en la región de Tarapacá, y el mínimo se encuentra en las regiones de Antofagasta, Coquimbo y Aysén con una tasa de cooperación de 57%.

Finalmente, en la última columna se muestra la **tasa de rechazo**, correspondiente a la división entre rechazos del hogar o a nivel de persona, por el total de casos elegibles. Esta **tasa fue de 34%**, teniendo su máximo de 39% en la región de Coquimbo, y su mínimo en Tarapacá con 23%.

**Tabla 11 Tasas AAPOR de resultados de terreno**

Región		Muestra objetivo	Muestra sobre- dimensión	Tasa de logro	Tasa de respuesta RR1	Tasa de contacto CON1	Tasa de cooperación COOP1	Tasa de rechazo REF1
		1	2	3	4	5	6	7
1	Tarapacá	580	1.150	106%	54%	76%	70%	23%
2	Antofagasta	880	1.760	101%	51%	90%	57%	39%
3	Atacama	520	1.030	102%	52%	78%	66%	26%
4	Coquimbo	1.150	2.290	102%	51%	91%	57%	39%
5	Valparaíso	1.640	3.270	104%	52%	90%	58%	37%
6	O'Higgins	1.410	2.820	101%	50%	80%	63%	30%
7	Maule	1.410	2.830	103%	51%	89%	58%	38%
8	Biobío	1.750	3.510	107%	53%	88%	60%	35%
9	La Araucanía	1.580	3.160	95%	47%	78%	60%	31%
10	Los Lagos	1.450	2.900	105%	53%	88%	60%	35%
11	Aysén	490	980	102%	51%	85%	60%	34%
12	Magallanes	510	1.020	88%	44%	77%	57%	33%
13	Metropolitana	4.380	8.740	100%	50%	84%	60%	34%
14	Los Ríos	650	1.300	102%	51%	80%	64%	29%
15	Arica y Parinacota	550	1.100	95%	48%	81%	59%	33%
16	Ñuble	1.050	2.110	111%	55%	89%	62%	33%
<b>Total</b>		<b>20.000</b>	<b>39.970</b>	<b>102%</b>	<b>51%</b>	<b>85%</b>	<b>60%</b>	<b>34%</b>

### **4.3 Caracterización del trabajo de campo**

A continuación, se presentará una caracterización del trabajo de campo de acuerdo a tres elementos; resultados del proceso de supervisión, resultados asociados al control de calidad de la base de datos, y por último, se abordarán las principales dificultades que se han dado durante la recolección de los datos.

#### **4.3.1 Resultados de supervisión**

Para asegurar la calidad de los datos, se establecieron cuatro tipos de supervisiones, que fueron llevadas a cabo siguiendo un proceso lineal. En concreto, los procesos de supervisión a los que se sometieron las encuestas aplicadas fueron:

- Supervisión de audios/videos.
- Supervisión de GPS.
- Supervisión telefónica.
- Supervisión en terreno.

La linealidad del proceso implica que, por ejemplo, ninguna encuesta puede participar de alguna etapa posterior a la supervisión de audios y videos sin haber pasado exitosamente por esta.

A continuación, se presentarán los resultados de cada una de estas etapas de supervisión.

##### a) Supervisión de audio/video

La supervisión de audio y video implica la escucha de todos los audios, y la visualización todos los videos que el programa de captura de datos toma de las encuestas aplicadas, con el objetivo de determinar los siguientes aspectos:

- Correcta aplicación de las preguntas: verificación de la existencia de un interlocutor válido.
- Coherencia con las preguntas aplicadas: verificación de que la respuesta que entrega el interlocutor coincida con lo que registrado en el dispositivo, y por tanto, en la base de datos.
- Correcta formulación de las preguntas: verificación de que las preguntas hayan sido realizadas de la forma apropiada, sin alteraciones que alteren el sentido de la pregunta.

Para mejorar y elevar los criterios de exigencia en el proceso de supervisión, las preguntas que seleccionadas a pasar por este proceso van cambiando aleatoriamente a lo largo del proceso de levantamiento de datos. Adicionalmente, debe tenerse en consideración que, dada la naturaleza de los módulos autoaplicados, las preguntas incluidas en estos se excluyen del proceso de supervisión de audio y video.

Una vez terminado el proceso de supervisión de audio y video de una encuesta, y en el caso que se identifique alguna falta, se da a los encuestadores la posibilidad de que justificar las falencias del proceso. Si las razones son atendibles, las encuestas que forman parte de la apelación pasan por un proceso de supervisión telefónico y/o presencial más extenso (que se describirá más adelante), en el cual se busca verificar la correcta aplicación de la encuesta.

Los resultados de la supervisión de audio y video se muestran en la Tabla 12. Como se puede ver, se supervisaron en total 21.441 encuestas, de las cuales 20.443 fueron aceptadas por este proceso, correspondiendo a una tasa del 95%. Al analizar este dato por región, se tiene que el máximo porcentaje de rechazos de este tipo se han dado en la región de Coquimbo.

**Tabla 12 Resultados de proceso de supervisión audio y video**

<b>Región</b>	<b>Aceptado</b>	<b>Rechazado</b>	<b>Total encuestas</b>
Tarapacá	97,0%	3,0%	635
Antofagasta	98,8%	1,2%	1548
Atacama	90,1%	9,9%	558
Coquimbo	85,0%	15,0%	533
Valparaíso	95,6%	4,4%	4599
O'Higgins	99,4%	0,6%	673
Maule	92,6%	7,4%	568
Biobío	92,3%	7,7%	1269
La Araucanía	98,7%	1,3%	907
Los Lagos	95,4%	4,6%	561
Aysén	99,1%	0,9%	1189
Magallanes	99,4%	0,6%	1725
Metropolitana	87,6%	12,4%	1629
Los Ríos	98,2%	1,8%	1482
Arica y Parinacota	95,8%	4,2%	1951
Ñuble	93,2%	6,8%	1614
<b>Total encuestas</b>	<b>20.443</b>	<b>998</b>	<b>21.441</b>

#### b) Supervisión de GPS

Una vez que la encuesta pasa por el proceso de supervisión de audios/videos, se procede a la supervisión de GPS, la cual se realiza para verificar que las encuestas se estén aplicando en la manzana seleccionada. Este proceso se realiza comparando los puntos de ubicaciones donde se realizaron las encuestas (que son capturadas por el programa) y

compararlos con los polígonos de las manzanas seleccionadas en la muestra. Se aplica mediante el siguiente procedimiento:

1. Se realiza la lectura de los puntos en el mapa que el programa captura.
2. Se comparan con los polígonos de las manzanas seleccionadas.
3. Se coteja con la dirección registrada en la encuesta, y por ende, en la base de datos.

Tal como ocurre en el proceso de supervisión telefónica, una vez culminada la etapa de supervisión de GPS se les entrega a los encuestadores la posibilidad de apelar a sus rechazos. Esta posibilidad no existe para encuestadores cuyas encuestas hayan sido rechazadas porque más del 50% de ellas fue hecha con el GPS apagado. En este proceso de apelación, se indaga al encuestado por la dirección de la vivienda, además de la ubicación geográfica respecto de la manzana seleccionada.

Los resultados del proceso de supervisión de GPS se exponen en la Tabla 13. Como se ve, se supervisaron de esta manera un total de 20.453 encuestas, siendo rechazadas en este proceso solo 10 de ellas.

**Tabla 13 Resultados de proceso de supervisión GPS**

<b>Región</b>	<b>Aceptado</b>	<b>Rechazado</b>	<b>Total encuestas</b>
Tarapacá	98,7%	1,3%	624
Antofagasta	100,0%	0,0%	1530
Atacama	100,0%	0,0%	503
Coquimbo	100,0%	0,0%	453
Valparaíso	100,0%	0,0%	4397
O'Higgins	100,0%	0,0%	669
Maule	100,0%	0,0%	526
Biobío	100,0%	0,0%	1171
La Araucanía	100,0%	0,0%	895
Los Lagos	100,0%	0,0%	535
Aysén	99,9%	0,1%	1179
Magallanes	100,0%	0,0%	1714
Metropolitana	99,9%	0,1%	1428
Los Ríos	100,0%	0,0%	1455
Arica y Parinacota	100,0%	0,0%	1870
Ñuble	100,0%	0,0%	1504
<b>Total, encuestas</b>	<b>20.443</b>	<b>10</b>	<b>20.453</b>

Adicionalmente, cabe destacar que las encuestas que son realizadas con el GPS apagado son derivadas automáticamente al proceso de supervisión telefónico y/o presencial extenso (al que se aludió en el apartado anterior). Sin embargo, si un mismo encuestador hace más de un 50% de sus encuestas con el GPS apagado, estas son rechazadas de manera automática.



### c) Supervisión telefónica

La supervisión telefónica consiste en extraer una muestra de las encuestas aplicada por cada encuestador, y corroborar que éstas hayan sido aplicadas efectivamente, y de una manera apropiada. Para esto, se establece un protocolo que contempla re indagar sobre las siguientes preguntas del cuestionario:

- Número de residentes en el hogar.
- Sexo asignado al nacer.
- Edad.
- Nivel educacional.
- Nacionalidad.
- País de residencia de la madre cuando el encuestado nació.

Además de esto, se aplica un protocolo orientado a establecer el comportamiento de cada encuestador y la evaluación general del proceso por parte del encuestado.

La selección de las preguntas que se aplican en este proceso se justifica en tanto deben ser preguntas cuya respuesta se espera que no se modifique desde el momento en que se aplicó la encuesta y el momento en que se realiza la supervisión telefónica.

Para determinar si una encuesta es rechazada en este proceso, se construye un índice que le da un peso de 2 al número de residentes del hogar, y de 1 al resto de las variables que se preguntan, con excepción de la edad, para la cual se tolera un rango de 2 años de error, entendiéndose que el encuestado puede haberse tenido confusiones en la declaración de edad, o haber estado de cumpleaños entre la aplicación de la encuesta y de la re-entrevista. Si la encuesta obtiene un puntaje de 5 o más, es aprobada en la supervisión telefónica. De lo contrario, es rechazada.

Adicionalmente, si la persona, luego de 3 llamados, no responde al proceso de supervisión de encuesta, se deja como una encuesta aceptada, por cuanto se asume que puede ocurrir que el encuestado entregó un número incorrecto a propósito, o no desea responder a los llamados. Sin embargo, los números que no responden pasan forzosamente a formar parte de las encuestas a supervisar en terreno.

Los resultados de la supervisión telefónica se muestran en la Tabla 14. Como se puede observar, se supervisaron un total de 20.453 encuestas, de las cuales 11.588 fueron supervisadas telefónicamente. De estas, 11.485 fueron aprobadas por este proceso de supervisión, correspondiendo al 99%, sin grandes variaciones por región.

**Tabla 14 Resultados de proceso de supervisión telefónica**

Región	Encuestas aplicadas	% encuestas supervisadas	% encuestas aceptadas (sobre supervisadas)
Tarapacá	624	59,8%	99,8%
Antofagasta	1530	74,2%	99,9%
Atacama	503	73,8%	99,8%
Coquimbo	453	68,9%	99,6%
Valparaíso	4397	66,8%	100,0%
O'Higgins	669	38,4%	100,0%
Maule	526	39,0%	99,8%
Biobío	1171	37,1%	99,9%
La Araucanía	895	57,1%	100,0%
Los Lagos	535	64,9%	96,3%
Aysén	1179	23,7%	99,9%
Magallanes	1714	38,0%	100,0%
Metropolitana	1428	46,0%	95,1%
Los Ríos	1455	69,3%	99,9%
Arica y Parinacota	1870	59,1%	99,9%
Ñuble	1504	66,7%	100,0%
<b>Total, encuestas</b>	<b>20453</b>	<b>11588</b>	<b>11485</b>

#### d) Supervisión en terreno

La supervisión en terreno es análoga a la supervisión telefónica, aunque se realiza revisitando a los encuestados a los domicilios registrados durante el proceso de levantamiento de datos. Para escoger los casos que son visitados, se toman en cuenta los siguientes criterios:

- Casos que no respondieron a la supervisión telefónica.
- Casos de encuestadores que tienen rechazos por audio/video.
- Criterios sustantivos de cercanía entre manzanas.

Este proceso es llevado a cabo por un equipo de supervisores independiente de los encuestadores, siendo el único que no se realiza en paralelo con los demás. Esto implica que, para que una encuesta sea supervisada en terreno, debe haber pasado necesariamente por la supervisión de audio/video y GPS.

Durante esta supervisión se vuelven a aplicar las preguntas ya descritas en el proceso de supervisión. Los resultados de este proceso se muestran en la Tabla 15.

Como se puede ver, del total de 20.453 encuestas aplicadas, fueron supervisadas en terreno 2.422, correspondiendo a una tasa del 12%. De estos casos, fueron aprobadas en esta etapa 2.384 encuestas, correspondiendo al 98%, siendo relativamente estable entre regiones.

**Tabla 15 Resultados de proceso de supervisión en terreno**

<b>Región</b>	<b>Encuestas aplicadas</b>	<b>% encuestas supervisadas</b>	<b>% encuestas aceptadas (sobre supervisadas)</b>
Tarapacá	568	10,1%	100,0%
Antofagasta	697	14,6%	99,9%
Atacama	410	16,9%	100,0%
Coquimbo	1188	16,8%	100,0%
Valparaíso	1723	13,6%	99,9%
O'Higgins	1623	9,6%	99,1%
Maule	1475	7,2%	99,8%
Biobío	1950	7,9%	100,0%
La Araucanía	1470	11,6%	99,7%
Los Lagos	1498	15,3%	99,1%
Aysén	369	4,7%	99,9%
Magallanes	346	8,8%	99,9%
Metropolitana	2630	10,5%	99,7%
Los Ríos	502	13,5%	99,9%
Arica y Parinacota	567	12,4%	100,0%
Ñuble	1267	14,0%	99,4%
<b>Total, encuestas</b>	<b>20.453</b>	<b>2.422</b>	<b>2.384</b>

#### 4.3.2 Resultados de control de calidad e inconsistencias

Una vez consolidada la base de datos proveniente del software de levantamiento en tablet se procedió a validar la base de datos construida en un Software ad hoc como SPSS, Stata o R. Este procedimiento se realizó a partir del análisis del flujo lógico, no respuesta, e identificación de inconsistencias o valores fuera de rango dentro de lo establecido en el instrumento.

Para lo anterior se construyen **sintaxis y/o mallas de validación** que permitan identificar al menos los siguientes aspectos (el proceso de validación será reformulado una vez obtenido el cuestionario definitivo):

- No respuesta a nivel total de la encuesta, y no respuesta de preguntas claves: La no respuesta o *missing* es frecuente en las encuestas a hogares debido a la fatiga del informante, al desconocimiento de la información solicitada o al rechazo de las personas a informar acerca de temas sensibles. Uno de los elementos claves a observar se relaciona al porcentaje máximo de omisiones que pueden ser aceptables; como no existen criterios unificados para guiar esta pregunta cada equipo investigador debe establecer sus parámetros (Medina & Galván, 2007). Lynch (2003) otorga algunos criterios derivados de los procedimientos de imputación; menciona que, para el análisis de datos, las variables que tienen un 5% de *missing* no requieren mayor tratamiento, entre un 5% y un 10% es importante evaluar si los casos se distribuyen aleatoriamente o presentan algún sesgo en particular, pero cuando éstos sobrepasan el 10% ya es importante tomar medidas. De esta manera, se podrá establecer un parámetro derivado de los datos establecidos por la literatura que permita evaluar y dar cuenta de las medidas a tomar según los porcentajes de no respuesta encontrados en los procesos de validación.
- Por otro lado, la validación deberá aclarar los casos de no respuesta del encuestado, de aquella que no corresponde debido a filtros.
- Casos fuera de rango (*outliers*): Se deriva del chequeo de los valores mínimos y máximos posibles en una pregunta. Por ejemplo, chequeos de días a la semana, horas de trabajo máximo a la semana.
- Consistencia en datos socio demográficos: Se chequeará la consistencia en datos demográficos como sexo y edad. Para el caso de edad, si el cuestionario solicita fecha de nacimiento, se revisará la consistencia entre la edad calculada y la declarada.
- Chequeos de consistencia en eventos con tiempos cronológicos.
- Verificación lógica del cálculo de variables al interior del instrumento.

Dado esto, en la Tabla 16 se reportan los errores de validación encontrados en la base de datos y el porcentaje de incidencia de este error en los casos aplicados.

Como se puede ver, se cotejaron 23 posibles errores, los cuales tienen una incidencia que varía desde el 0,01% de los casos, al 7,1% de los respondentes.

**Tabla 16 Errores de validación, solución propuesta y porcentaje de incidencia en base de datos**

<b>Variable</b>	<b>Error</b>	<b>% de incidencia</b>
P22	Peso declarado 0	3,0%
P23	Estatura declarada 0	2,3%
P27	Cantidad de hijos deseada superior a 15	0,1%
P43	Edad de primera eyaculación nocturna con valor negativo	7,2%
P45	Personas que no declaran haber dado su primer beso, declaran haber tenido relaciones sexuales voluntarias	0%
P52	Edad de pareja en primera relación sexual voluntaria con valor negativo	4,7%
P69	Personas que declaran 0 parejas sexuales en el transcurso de su vida, pero señalan haber tenido relaciones sexuales voluntarias	0%
P78	Personas que declaran tener más de 50 parejas afectivas relevantes	0,01%
P99	Edad de primera masturbación con valor negativo	6,6%
P103_1	Persona declara haber consumido alcohol en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido alcohol en su vida	1,3%
P103_2	Persona declara haber consumido cocaína en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido cocaína en su vida	0,1%
P103_3	Persona declara haber consumido marihuana en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido marihuana en su vida	0,2%
P103_4	Persona declara haber consumido drogas inyectables en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido drogas inyectables en su vida	0,04%
P103_5	Persona declara haber consumido tranquilizantes en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido tranquilizantes en su vida	0,1%
P103_6	Persona declara haber consumido alucinógenos en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido alucinógenos en su vida	0,04%
P103_7	Persona declara haber consumido poppers en al menos una de las relaciones sexuales que ha tenido en el último mes, y a la vez, declara nunca haber consumido poppers en su vida	0,05%

P113a	Edad de primer consumo de prostitución con valor negativo	0,2%
P113b	Edad de último consumo de prostitución con valor negativo	0,3%
P113b	Edad de último consumo de prostitución menor a edad de primer consumo de prostitución	0,1%
P138	Edad en que empezó a sentir que su género era distinto al que se asignaron es mayor que edad del encuestado	0,03%
P160	Se declara estar embarazada actualmente, pero se menciona 0 embarazos en la vida	0,01%
P224a	Semanas de gestación al momento del aborto superior a 40	0,4%
P224b	Meses de gestación al momento del aborto superior a 12	0,2%

Por su parte, en la Tabla 17 se ve la distribución de la cantidad de errores detectados. En concreto, el 79,1% de los respondentes no posee ningún error de validación. Un 16,1 presenta solo 1, y un 4,7% dos errores.

Estos resultados se encuentran alineados con la modalidad CAPI (*Computer Assisted Personal Interview*), donde la gran mayoría de los errores de validación son evitados mediante una correcta programación de filtros lógicos y control de valores fuera de rango.

**Tabla 17 Distribución de errores detectados**

Cantidad de errores detectados	Casos	%
0	16.138	79,1
1	3.282	16,1
2	750	4,7
3	186	0,9
4	23	0,1
5	13	0,1

Adicionalmente, luego de una revisión de la contraparte técnica se detectaron sugerencias de cambios y correcciones a realizarle a la base de datos general. Estas correcciones fueron ya realizadas en la base de datos que se adjunta al informe, y los cambios realizados se detallan en la Tabla 18.

**Tabla 18 Modificaciones dado inconsistencias en base de datos**

Variable	Error	Casos modificados
P03	Persona responde Nr, en vez de "Prefiere no responder".	1
P49	Inconsistencia dada respuesta de "No" en P47. Se imputa valor perdido.	1

P66_det	Se reportan meses superiores a 12. Se editó para mantener respuestas en años.	6
P79a	Suma de parejas sexuales según sexo no congruente. Se corrigió mediante supervisión telefónica.	1
P80a	Número de parejas con las que ha vivido no congruente. Se corrigió el dato.	1
P80c	Valores inválidos según respuestas. Se corrigió dato según respuesta original.	2.438
P96	Incongruencia en respuesta de si ha mantenido relaciones sexuales con su pareja el último mes. Se corrige dato.	1
P162	Personas declaran haber tenido 0 cantidad de embarazos no planificados, a pesar de haber declarado previamente que los tuvieron. Se corrige dato.	294
P164	Respuestas inválidas en edad de primer embarazo no deseado. Se corrige dato.	11
P175	Respuesta 0 en la semana del primer control natal. Se corrige dato.	573
P199	Respuestas inválidas en edad de menopausia. Se corrige dato según supervisión telefónica	12
i_2_p203a	Respuestas inválidas en edad de diagnóstico de Hepatitis B. Se corrige dato cuando aplica.	2
i_4_p203a	Respuestas inválidas en edad de diagnóstico de Hepatitis C. Se corrige dato cuando aplica.	6
P216	Respuestas extremas en cantidad de veces de uso de anticoncepción de emergencia. Se corrige dato.	2
P284	Casos con No respuesta se estipulan con valor 9 de ingresos totales, que es un monto válido. Se modifica por 9999	4.512

### 4.3.3 Duración de las entrevistas

Por último, se presenta la mediana de duración de las entrevistas. Tal como se observa en la tabla 19, la mediana nacional de tiempo de aplicación de la entrevista fue de 43 minutos, observando diferencias para mujeres (44 minutos) en comparación a hombres (40). Cabe destacar que las mujeres completaban un número mayor de consultas en los módulos de salud sexual y reproductiva.

**Tabla 19 Mediana tiempo de aplicación de la entrevista**

<b>Variable</b>	<b>categoría</b>	<b>Mediana</b>
	Total, País	43 minutos
<b>Sexo</b>	Mujer	44 minutos
	Hombre	40 minutos
<b>Tramos de edad (años)</b>	30 a 34	42 minutos
	18 a 24	37 minutos
	55 a 59	45 minutos
	25 a 29	41 minutos
	50 a 54	45 minutos
	75 a 79	46 minutos
	65 a 69	45 minutos
	70 a 74	44 minutos
	60 a 64	45 minutos
	40 a 44	44 minutos
	35 a 49	43 minutos
	45 a 49	44 minutos
80 y más	44 minutos	



## 5 Bibliografía

- Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples*. New York: Springer.
- Lohr, S. L. (1999). *Sampling: design and analysis*. Brooks.
- Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American statistical Association*, 44(247), 380-387.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. (Vol. 561): John Wiley & Sons.
- AAPOR. (2016). *Standard definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*.
- Medina, F., & Galván, M. (2007). Estudios estadísticos y prospectivos. Imputación de datos: Teoría y práctica. *División Estadística y proyecciones Económicas Naciones Unidas. CEPAL*.
- Lynch, S. M. (2003). Using statistics in social research. *Using Statistics in Social Research: A Concise Approach*, 1-15.
- Kalton, G., & Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics*, 81-97.
- Rosenbaum, P. R., & Rubin, D. B. (1985). The bias due to incomplete matching. *Biometrics*, 103-116.
- Deville, J. C., & Särndal, C. E. (1994). Variance estimation for the regression imputed Horvitz-Thompson estimator. *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM*, 381.

Kreuter, F., & Valliant, R. (2007). A survey on survey statistics: What is done and can be done in Stata. *The Stata Journal*, 7(1), 1-21.