FOTO DE HAFM JANSEN

La elaboración de una estrategia de muestreo

Los temas que se abordan en este capítulo son:

Consideraciones sobre el muestreo en estudios cualitativos
Consideraciones sobre el muestreo en encuestas para investigaciones cuantitativas

Nunca es suficiente el énfasis que se ponga en la importancia de establecer una muestra apropiada para el tipo de diseño de investigación seleccionado. Si bien en la investigación cualitativa y cuantitativa se utilizan diferentes enfoques para seleccionar a las personas o grupos que se van a estudiar, en todos los estudios es crucial que se planifique con cuidado la estrategia de muestreo. Especialmente en el caso de encuestas basadas en la población, una muestra mal seleccionada puede afectar la credibilidad de un estudio aun si el resto del estudio se ejecuta bien.

Por lo general, los estudios cualitativos enfocan a profundidad un número relativamente pequeño de casos seleccionados **a propósito**. Por otra parte, los métodos cuantitativos típicamente dependen de muestras más grandes seleccionadas al azar. Estas tendencias surgen del propósito subyacente del muestreo en las dos tradiciones de investigación. En la investigación cuantitativa, el objetivo del muestreo es maximizar la **representatividad** de la muestra para así poder generalizar las conclusiones de la muestra en una población más grande. En la indagación cualitativa, la meta es seleccionar para enriquecer la información y así arrojar luz sobre las preguntas bajo estudio.[1]

En este capítulo se abordan los principales temas que deben tomarse en cuenta al diseñar una muestra para la investigación cualitativa o cuantitativa, y se brindan ejemplos sobre cómo se han utilizado diferentes estrategias para satisfacer las necesidades y circunstancias específicas de los proyectos de investigación.

CONSIDERACIONES SOBRE EL MUESTREO EN ESTUDIOS CUALITATIVOS

No existen normas rígidas en lo que respecta a los tamaños de las muestras en la investigación cualitativa. Como lo señala Hudelson, "el tamaño de la muestra dependerá del propósito de la investigación, de las preguntas de investigación específicas que se abordarán, de lo que será útil, de lo que tendrá credibilidad, y de lo que se puede hacer con el tiempo y los recursos a disposición".[2]

En el muestreo cualitativo, la selección de personas encuestadas usualmente continúa hasta el punto de **redundancia** o

TABLA 7.1 TIPOS DE ESTRATEGIAS DE MUESTREO PARA ESTUDIOS CUALITATIVOS

Tipos de muestreo	*Propósito*	*Ejemplo*
Muestreo de intensidad	Proporcionar abundante información a partir de unos cuantos casos seleccionados, que manifiesten el fenómeno intensamente (pero que no sean casos extremos).	Entrevista a sobrevivientes de violación sexual cometida durante el cortejo, para saber más sobre cómo el sexo bajo coerción afecta la sexualidad de las mujeres.
Muestreo de casos divergentes	Aprender de las manifestaciones altamente inusuales del fenómeno en cuestión.	Entrevista a hombres que no maltratan a sus parejas en una cultura en la que el maltrato contra la pareja es culturalmente aceptable.
Muestreo estratificado intencional	Ilustrar características de subgrupos de interés particular a fin de facilitar las comparaciones.	Entrevista a distintos tipos de proveedores de servicios (policía, trabajadores sociales, médicos, miembros del clero) a fin de comparar sus actitudes y trato a las víctimas de abuso.
Muestreo por bola de nieve o en cadena (ubicándolo a una o dos personas claves y pidiendo que mencionen a otras personas que puedan brindar información)	Facilitar la identificación de casos difíciles de encontrar.	Encontrar trabajadoras sexuales comerciales, que sean referidas a través de redes de amistades para entrevistarlas sobre experiencias de abuso sexual.
Muestreo de máxima variación (seleccione intencionalmente una amplia gama de variación en las dimensiones de interés)	Documentar diversas variaciones. Puede ayudar a identificar patrones comunes en las variaciones.	Investigación de variaciones en normas de aceptación del maltrato a la pareja por medio de grupos focales con distintos subgrupos: mujeres jóvenes urbanas, mujeres mayores urbanas, hombres jóvenes rurales, hombres mayores rurales, mujeres que han sido maltratadas, mujeres que no han sido maltratadas).
Muestreo de conveniencia (seleccione a quien sea más fácil, más cercano, etc.)	Ahorrar tiempo, dinero y esfuerzo. La información recopilada suele tener poca credibilidad.	Formación de grupos focales con las personas disponibles ese día en el centro comunitario local, en vez de basarse en criterios claros.
Muestreo por criterios	Investigar a profundidad un "tipo" de caso particular; identificar todas las fuentes de variación.	Entrevistar específicamente sólo a mujeres maltratadas que se han separado de sus parejas en el último año, a fin de entender mejor la variedad de factores que las animan a abandonarlos.

(Tomado de Patton, 1990).[3]

saturación. Esto significa que cuando nuevas entrevistas ya no aportan nueva información y todas las potenciales fuentes de variación han sido debidamente exploradas, el muestreo puede detenerse. Para la mayoría de los estudios cualitativos 10 a 30 entrevistas o 4 a 8 grupos focales serán suficientes. En la tabla 7.1 se da un resumen de una serie de diferentes enfoques hacia el muestreo cualitativo.

Así pues, en la investigación cualitativa, la estrategia de muestreo se debe selec-

cionar para ayudar a dilucidar la pregunta que se está planteando. Por ejemplo, el equipo de investigación de la organización nicaragüense Puntos de Encuentro, inició un proyecto para recopilar información que fuera útil para el diseño de una campaña mediática nacional que invitaba a los hombres a renunciar a la violencia en sus relaciones íntimas. Los investigadores querían entender las creencias y actitudes prevalecientes en la cultura nicaragüense que respaldan el comportamiento violento hacia las mujeres. Más importante aún, querían averiguar si había algunos "beneficios" en la no violencia que pudieran promoverse para alentar a los hombres a reconsiderar su comportamiento (véase el cuadro 5.6).

En lugar de concentrarse en recopilar información sobre normas y actitudes de hombres nicaragüenses "típicos", el grupo de investigación decidió utilizar un muestreo de "caso divergente" y concentrarse en encuestar a hombres que ya tenían la reputación de no ser violentos y de haber renunciado al machismo.[4] Estaban interesados en preguntarle a esos hombres qué beneficios —si los hubiere— les aportaba su decisión, y qué eventos en el transcurso de sus vidas, influencias o individuos les habían empujado en esa dirección. La meta era investigar qué aspiraciones y experiencias de la vida contribuyen a crear relaciones íntimas "saludables". Los resultados se utilizaron para diseñar una campaña informativa dirigida a atraer a más hombres hacia un estilo de vida sin violencia.

CONSIDERACIONES SOBRE EL MUESTREO EN ENCUESTAS PARA INVESTIGACIONES CUANTITATIVAS

En contraste con la investigación cualitativa, que generalmente utiliza un muestreo no probabilístico o "intencional", la investigación cuantitativa se basa en el muestreo aleatorio de las personas encuestadas. Una muestra probabilística o "representativa" es un grupo de personas encuestadas seleccionadas de la población, de manera que los resultados puedan generalizarse para toda la población.

FOTO DE HAFM JANSEN

Cuando a una muestra se le llama **aleatoria**, esto significa que se han utilizado técnicas específicas para asegurar que toda persona que satisfaga ciertos criterios de elegibilidad tiene la misma probabilidad de ser incluida en el estudio. El no cumplimiento de esas técnicas puede introducir errores o **sesgos** en la muestra, lo que puede mermar la validez del estudio. Por ejemplo, si en una encuesta a hogares en torno a la violencia sólo se realizaran entrevistas durante el día, las personas encuestadas con mayores probabilidades de ser incluidas en el estudio serían las mujeres que trabajan en el hogar, y aquellas mujeres que trabajan fuera del hogar

tendrían menos probabilidades de ser entrevistadas. Dado que las mujeres que trabajan fuera del hogar pueden tener diferentes experiencias en torno a la violencia, los resultados del estudio estarían sesgados hacia las mujeres que trabajan en el hogar. Una manera de reducir este particular sesgo sería regresar a los hogares por la noche o durante los fines de semana para elevar la posibilidad de abarcar a todas las mujeres.

*Una muestra probabilística o **representativa** es un grupo de informantes seleccionadas de la población de tal manera que los resultados del estudio puedan generalizarse a toda la población.*

La manera como se selecciona la muestra afecta su posibilidad de **generalización** o el grado al cual la situación encontrada entre una muestra particular en un momento dado puede aplicarse de una manera más general. Existen muchas técnicas para el muestreo, cada una de ellas con sus propias ventajas relativas en términos de costo, esfuerzo y potencial para generar resultados estadísticamente significativos. Algunas estrategias, tales como el muestreo aleatorio sencillo, puede no ser factible donde hay poca información disponible sobre la población del estudio. A continuación se brinda una breve descripción de las técnicas de muestreo más comunes que se utilizan.

*Cuando a una muestra se le llama **aleatoria o al azar**, esto significa que se han utilizado técnicas estadísticas específicas para asegurar que cada persona que satisfaga ciertos criterios de elegibilidad tiene la misma probabilidad de que se le incluya en el estudio.*

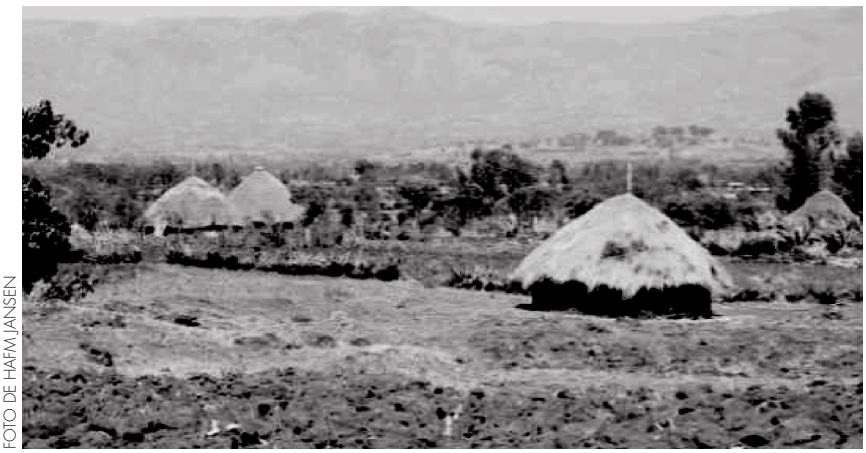


FOTO DE HAFM JANSEN

Muchas personas subestiman el desafío de obtener muestras bien diseñadas. A menudo se cometen errores debido a la confusión en torno al significado del término **selección aleatoria**. Una selección aleatoria no significa que los participantes simplemente se seleccionan sin ningún orden en particular. De hecho, las técnicas para obtener una muestra verdaderamente aleatoria son bastante complejas, e investigadores sin experiencia deben consultar a un experto en muestreo antes de proceder. Utilizar un cuestionario cuidadosamente pensado y probado en una muestra mal diseñada arrojará resultados sin sentido.

A menudo las muestras aleatorias se confunden con muestras de conveniencia o de cuotas. Una **muestra de conveniencia** es cuando las personas que se van a entrevistar se seleccionan según quién esté disponible, sin ningún orden en particular. En una **muestra de cuota**, se selecciona un número fijo de un cierto tipo de personas para entrevistarlas. Ninguna de las dos estrategias resultará en una muestra aleatoria apropiada para la encuesta de la investigación.

Muestreo aleatorio sencillo

Esta técnica de muestreo implica la selección al azar de una lista de la población, conocida como el **marco de muestreo**. Si se conduce debidamente, asegura que cada persona tiene una **misma** e **independiente** oportunidad de ser incluida en la muestra. En este caso, independencia significa que la selección de la persona no influye de ninguna manera en la selección de cualquier otra. La palabra "sencillo" no significa que este método sea más fácil, sino más bien que se toman las medidas para asegurar que sólo el azar influye en la selección de las personas que se van a entrevistar. La selección aleatoria se puede lograr mediante el uso de un método de lotería, de tablas de números aleatorios (que se encuentran en muchos libros sobre

estadísticas), o de un programa computarizado como por ejemplo Epi Info. Para evitar el sesgo, es importante que en el marco del muestreo sólo se incluya a personas que son elegibles para ser entrevistadas por criterios tales como edad, sexo o residencia. De igual manera, si a ciertas personas se les deja fuera de la lista original debido a un censo que ya no está vigente y que no incluye a aquellas personas recién llegadas al área, entonces esas omisiones podrían sesgar los resultados, especialmente si la migración es el resultado de crisis, tales como guerra, desastres naturales o colapso económico. En esos casos será preciso actualizar la lista del muestreo.

Muestreo sistemático

En el muestreo aleatorio, cada persona o cada hogar se seleccionan al azar. En contraste, el muestreo sistemático inicia en un punto aleatorio en el marco de muestreo, y se selecciona cada enésima persona. Por ejemplo, si se quiere una muestra de cien mujeres de un marco de muestreo de 5,000 mujeres, entonces se selecciona al azar un número entre uno y cincuenta para empezar la secuencia, y luego a partir de ahí se selecciona cada quincuagésima mujer. Tanto el muestreo aleatorio como el sistemático requieren una lista completa de la población para poder hacer una selección. También es importante saber cómo se elaboró la lista, y si las personas están colocadas al azar o en algún tipo de orden. Si los habitantes de la misma vivienda o con ciertas características están agrupados, esto podría resultar en una muestra sesgada en la que las personas con esas características estén o bien sobrerepresentados o subrepresentados.

Muestreo estratificado

El **muestreo estratificado** se pude utilizar junto con el muestreo aleatorio sencillo o con el muestreo sistemático. Esto garantiza que la muestra es lo más cercana posible a la población de estudio en lo que concierne a ciertas características tales como edad, sexo, etnia, o condición socioeconómica. En ese caso, la población de estudio se clasifica en **estratos** o subgrupos y luego se seleccionan al azar a personas de cada estrato. Dado que la estratificación implica un esfuerzo adicional, sólo tiene sentido si la característica que se está estratificando está relacionada con el resultado bajo estudio. Para fines de análisis es más fácil si el número de personas seleccionados de cada estrato es proporcional a su distribución en la población (véase el cuadro 7.1 sobre muestras autoponderadas). Por ejemplo, en una muestra estratificada conforme a la residencia urbana/rural, la proporción de mujeres rurales en la muestra sería la misma que la proporción de mujeres rurales en la población de estudio.

Una muestra estratificada ponderada podría ser preferible cuando en la población haya algunos grupos que son proporcionalmente pequeños, pero que son relevantes para el propósito del estudio, tales como personas de una cierta región geográfica o grupo étnico. Asegurar que estos grupos estén debidamente representados podría requerir un tamaño de muestra excesivamente grande si se usan

FOTO DE HAFM JANSEN

Un mapa de las calles utilizado para ubicar los hogares en el estudio de la OMS en el Japón.

CUADRO 7.1 MUESTREO POR CONGLOMERADOS AUTOPONDERADOS

La manera de escoger la muestra influye enormemente en la utilidad de los estimados resultantes. Supongamos que hay un distrito que sólo tiene dos poblados:

- En el poblado A hay 4,000 mujeres, de las cuales 800 (20%) han sido maltratadas.
- En el poblado B hay 800 mujeres, de las cuales 40 (5%) han sido maltratadas.

La prevalencia real de maltrato en este distrito se calcularía como sigue:

$$\frac{\text{Total de casos de violencia}}{\text{Número total de mujeres}} = \frac{(800 + 40) \times 100}{(4000 + 800)} = 17.5\%$$

Sin embargo, si decidiéramos determinar la prevalencia del abuso en este distrito sobre la base de una muestra aleatoria de 100 mujeres en cada poblado, encontraríamos lo siguiente:

- 20 entre 100 mujeres del poblado A denuncia maltrato.
- 5 entre 100 mujeres del poblado B denuncia maltrato.

Al combinar estas dos cifras, encontraríamos que 25 entre 200 mujeres entrevistadas fueron maltratadas, lo que nos daría una prevalencia de 12.5%.

¿Qué ha sucedido aquí?

El procedimiento de muestreo llevó a subestimar la prevalencia porque el número de informantes seleccionados en cada poblado no guardaba proporción con el tamaño relativo de cada poblado. Se puede suponer que si se hubiera sabido cuál era el tamaño relativo de los poblados, se habría podido realizar un análisis ponderado, según el cual los resultados del poblado A serían cinco veces más altos que los del poblado B. Sin embargo, es preferible, por lo general, obtener una muestra ponderada. Una manera de hacerlo sería seleccionar cinco veces más encuestadas en el poblado A que en el B. Otro método sería seleccionar los poblados con una probabilidad proporcional al tamaño, lo cual significa que si se tiene una lista de poblados, uno grande como el A tendría cinco veces más probabilidades de que se le seleccionara para la muestra que un poblado del tamaño del B. Después de la selección de los poblados, se podría seleccionar una cantidad igual de encuestadas de cada uno (véase el cuadro 7.3 para un ejemplo de cómo se obtuvo una muestra autoponderada en Perú).

(Tomado de Morison, 2000).[5]

técnicas de muestreo aleatorio sencillo. En ese caso, podría ser apropiado tener una muestra más grande de lo necesaria o seleccionar de ese estrato un número desproporcionadamente grande de personas a entrevistar. Esto resulta en una muestra ponderada que tendrá que tomarse en cuenta en el proceso de análisis.

Muestreo en múltiples etapas y en conglomerados

El muestreo en múltiples etapas se utiliza a menudo para extraer muestras de poblaciones muy grandes que cubren una vasta área geográfica, e implica seleccionar la muestra por etapas o tomar muestras de las muestras. Primero se divide la población en conglomerados que ocurren de forma natural (como son los poblados o vecindarios). Luego se procede a extraer una muestra aleatoria de esos conglomerados para la encuesta. Esta es la primera etapa del muestreo. La segunda etapa podría implicar o bien seleccionar todas las unidades de muestreo (personas u hogares) en los conglomerados seleccionados, o seleccionar un grupo de unidades de muestreo desde dentro de los conglomerados. A veces se requieren más de dos etapas, para lo cual, como en la segunda etapa, se podrían escoger al azar poblados de los distritos seleccionados. En una tercera etapa se seleccionará de los conglomerados a las personas que se van a entrevistar. En cada etapa se podrían utilizar técnicas aleatorias sencillas, sistemáticas o estratificadas. Es recomendable consultar a una persona experta en estadísticas si considera válido utilizar un esquema de muestreo en múltiples etapas.

La ventaja del muestreo en múltiples etapas es que sólo se necesita un marco de muestreo, por ejemplo la lista de los hogares, para los conglomerados seleccionados, poblados y no para toda la población de estudio. Asimismo, la logística será más fácil debido a que el muestreo está restringido a los conglomerados seleccionados y no necesitan cubrir toda el área de estudio. En el cuadro 7.3 se incluye un ejemplo de una estrategia de muestreo en múltiples etapas en Perú.

Las desventajas de un muestreo en múltiples etapas es que el tamaño de la muestra tiene que ser considerablemente más grande que si la muestra se seleccionara mediante el muestreo aleatorio sencillo. Asimismo, puede ser más complicado obtener una muestra autoponderada. Otra dificultad con el muestreo en múltiples etapas puede ser la definición de los

conglomerados si el área de estudio es un área urbana grande, por ejemplo. Podría ser que estos ya hayan sido definidos para censos o encuestas anteriores, pero de lo contrario se deben crear a partir de un mapa o basarse en otros criterios tales como áreas de captación para la escuela o el centro de salud.

¿Qué tan grande debe ser la muestra?

El tamaño ideal de la muestra para una encuesta depende de varios factores:

- **¿Qué tan seguro se quiere estar de las conclusiones?** Por lo general, un tamaño de muestra más grande aumenta la precisión de los resultados o la confianza con la cual se puede decir que representan una medida confiable de los fenómenos bajo estudio.

- **¿Cuáles son las características de la población de estudio?** Entre más variada sea, más grande necesitará ser el tamaño de la muestra.

- **¿Qué tan común es el fenómeno bajo estudio?** Si cualquiera de las condiciones que se desean medir en el estudio son muy raras, como por ejemplo, la mortalidad infantil o la mortalidad materna, entonces se necesitará un tamaño de muestra muy grande.

- **¿Cuál es el propósito de la investigación?** El cálculo del tamaño de la muestra también dependerá de si simplemente se desea medir la prevalencia de una condición en una población o si se desea medir una diferencia esperada entre dos grupos. Los programa tales como el Epi Info contienen dos fórmulas diferentes para estos dos enfoques distintos.

- **¿Qué tipo de análisis estadístico se utilizará?** Esto subraya la necesidad de considerar desde un inicio cómo se van a analizar los datos. El tamaño de la muestra debe ser lo suficientemente grande para proporcionar los niveles de exactitud deseados en las estimaciones de prevalencia y para hacer pruebas entre diferentes variables sobre la importancia de las diferencias.

CUADRO 7.2 ENCUESTA DE POBLACIÓN CON MUESTREO ALEATORIO (TAMAÑO Y PODER ESTADÍSTICO DE LA MUESTRA UTILIZANDO EL PROGRAMA STATCALC)

Tamaño de la población	100,000	10,000	10,000	10,000
Frecuencia prevista	30%	30%	20%	30%
Peor frecuencia aceptable	25%	25%	15%	28%
Nivel de confianza	***Tamaño de la muestra***	***Tamaño de la muestra***	***Tamaño de la muestra***	***Tamaño de la muestra***
80%	138	136	104	794
90%	227	222	170	1,244
95%	322	313	270	1,678
99%	554	528	407	2,583
99.9%	901	834	648	3,624
99.99%	1,256	1,128	883	4,428

- **¿Qué tipo de estrategia de muestreo se utilizará?** Las fórmulas y los paquetes informáticos de uso común para el tamaño de la muestra, suponen que se está utilizando un muestreo aleatorio sencillo. Si se contempla utilizar muestreo en múltiples etapas o en conglomerados, podría ser preciso aumentar el tamaño de la muestra para lograr la precisión que se necesita. Considere consultar con una persona experta en estadística para que le ayude a decidir a cuanto tiene que aumentar el tamaño de la muestra.

Es mejor recopilar datos excelentes de pocos encuestados que recopilar datos de dudosa validez y confiabilidad de muchos entrevistados. Los paquetes estadísticos computarizados o las fórmulas matemáticas

CUADRO 7.3 MUESTREO EN PERÚ

Se realizó el estudio de la OMS sobre VCM en dos sitios en Perú; en Lima, la capital, y en Cuzco, una región montañosa con una población mayoritaria de habla quechua. El equipo de investigación utilizó dos estrategias de muestreo muy distintas para el estudio.

Lima

1. El equipo obtuvo una lista de conglomerados seleccionados del Instituto Nacional de Estadísticas e Información (INEI), según la cual Lima estaba dividida en 12,000 conglomerados, cada uno de los cuales tenía entre 50 y 150 casas. El equipo seleccionó de forma aleatoria 166 conglomerados y utilizó una probabilidad proporcional al tamaño (asignando un valor ponderado a los conglomerados más grandes para que tuvieran mayor probabilidad de que se les seleccionara). El INEI proveyó al equipo de mapas que mostraban dónde estaba cada uno de los conglomerados respecto de los demás y un mapa de cada conglomerado seleccionado que señalaba edificaciones como casas, tiendas y bodegas.
2. El equipo de Perú tuvo que actualizar los mapas, que habían sido preparados unos años antes, para tomar en cuenta la migración. Se envió a un equipo de enumeradores a cada conglomerado unas cuantas semanas antes de empezar el trabajo de campo, que fue a cada edificio para verificar cuántas casas había en cada uno. Asimismo, obtuvo información sobre la cantidad de mujeres elegibles en cada casa. Se enumeró un total de 21,322 casas en este proceso.
3. Basándose en esta información, el equipo calculó que necesitarían un tamaño de muestra de 2,000 para completar 1,500 entrevistas, tomando en cuenta las negativas y las casas sin mujeres elegibles. Por consiguiente, se hizo una selección de 12 hogares en cada uno de los 166 conglomerados. El intervalo de muestreo variaba según el tamaño del conglomerado. En un conglomerado de 120 casas, se seleccionó una de cada diez; en uno de 72 casas, una de cada seis, y así sucesivamente. Se utilizó una estrategia distinta en los conglomerados con menos de 48 casas. Para evitar entrevistar a mujeres que vivían muy cerca unas de otras, se podía perder la confidencialidad, se estableció un intervalo mínimo de cuatro casas entre los hogares seleccionados, es decir que en un conglomerado con 36 casas, sólo se entrevistaría a nueve mujeres y no a 12.
4. Para la selección de las casas, el supervisor del equipo escogió al azar un hogar en cada conglomerado utilizando la misma técnica para empezar la numeración. Por ejemplo, la casa más al norte en una cuadra. Luego, los entrevistadores siguieron el intervalo de muestreo para seleccionar las casas donde se entrevistaría a mujeres. No se permitió reemplazar ningún hogar donde no hubiera mujeres elegibles.

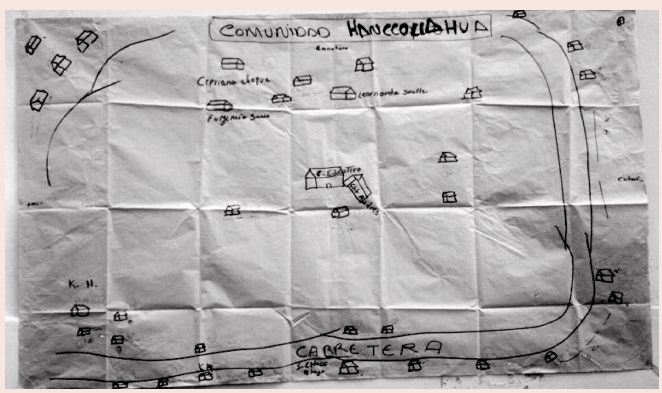
Líderes comunitarios trazando un mapa de las provincias en Cuzco, Perú

Cuzco

En Cuzco, donde hay unos cuantos pueblos rurales y el resto de la población se encuentra dispersa por las montañas, se utilizó una estrategia de muestreo en múltiples etapas. La región tiene 12 provincias y comprende la capital regional, Cuzco, donde vive gran parte de la población. Las otras provincias son rurales en su mayoría.

1. Se realizó una selección aleatoria de cuatro provincias, que comprendían la ciudad de Cuzco y tres provincias rurales.
2. El INEI seleccionó 46 conglomerados para la ciudad de Cuzco y 66 para las tres provincias. Se realizó la selección de cada conglomerado con una probabilidad proporcional al tamaño.
3. Para los conglomerados rurales sólo se disponía de mapas generales que señalaban la ubicación de los poblados más grandes, en los cuales el equipo realizó un rápido inventario de casas y elaboró un croquis. La cantidad total de hogares se dividió entre 23 para obtener el intervalo de muestreo. El primero fue seleccionado al azar y el resto de manera sistemática.

4. No había información disponible en la zona rural. El equipo sostuvo reuniones con los líderes de las comunidades rurales y clubes locales de mujeres para elaborar mapas detallados. Se informó a los líderes acerca de las metas generales del estudio y las fechas aproximadas en las que se llevaría a cabo el trabajo de campo en sus comunidades. Los líderes elaboraron mapas de cada poblado, con todas las casas y puntos de referencia indicados, y luego prepararon conjuntamente un mapa de las rutas que recorrería el personal de campo para llegar a los poblados.
5. Con la ayuda de los mapas se seleccionó al azar una casa y desde este punto de partida,se realizaron visitas a todos los hogares siguiendo una dirección en espiral hasta visitar 23 casas donde habían mujeres elegibles. Puesto que los hogares estaban muy dispersos, no fue necesario tener un intervalo de muestreo mínimo. Si un poblado no tenía 23 casas, entonces se visitaba el poblado más cercano. Se siguió el mismo procedimiento hasta completar 23 entrevistas.



Mapa de la comunidad con escuelas y rutas señaladas

(Tomado de Guezmes et al., 2002).[7]

se pueden utilizar para determinar el tamaño de la muestra para un estudio. El cuadro 7.2 presenta una tabla producida por el programa STATCALC de Epi Info para los cálculos del tamaño ideal de la muestra. Este programa está disponible en línea en http://www.cdc.gov/epiinfo

Si el análisis que se quiere realizar implica estudiar subgrupos particulares de la muestra, será preciso ampliar el tamaño de la muestra según sea necesario. Por ejemplo, para determinar la prevalencia de la violencia, usted podría requerir una muestra de sólo 300 mujeres, pero si desea saber si la prevalencia de la violencia varía según la edad, la educación o el grupo socioeconómico, entonces se necesitará un tamaño de muestra lo suficientemente grande para poder hacer comparaciones entre estos grupos.

El cálculo inicial se hizo en base a una muestra aleatoria sencilla de una población de estudio de 100,000 mujeres, en la que se supuso que un 30% de las mujeres habían experimentado violencia y que un margen de error del 10% sería aceptable (5% hacia arriba y 5% hacia abajo). Si estas suposiciones son correctas, la tabla indica que con un tamaño de muestra de 322 mujeres, se obtendría un intervalo de confianza del 95% para una prevalencia real del 25 y 35%. Obsérvese, sin embargo, que si las estimaciones utilizadas para calcular el tamaño de la muestra son inexactas, podría ser que no se obtuviera la precisión requerida.

La tabla también muestra que las diferencias en el tamaño de la población de estudio no influyen mucho en el tamaño de la muestra, mientras que los cambios en la frecuencia esperada y especialmente en el nivel de precisión que se necesita, pueden tener un enorme efecto en los cálculos del tamaño de la muestra. Cabe señalar que el tamaño de la muestra deberá aumentar si se utiliza una muestra en múltiples etapas. En vista de que estos cálculos pueden ser bastante complejos, se aconseja a investigadores inexpertos que consulten con alguien que tenga conocimientos de las técnicas de muestreo para encuestas.

Para explorar con mayor precisión las consecuencias de la violencia para la salud y comparar la incidencia de la violencia en diferentes sitios dentro de cada país, el estudio de la OMS sobre VCM utiliza una estrategia de muestreo en múltiples etapas cuyo objetivo es realizar 3,000 entrevistas en dos sitios; 1,500 en la capital y 1,500 en una provincia. Sin embargo, para contar al final con 1,500 entrevistas completas, usualmente es preciso aumentar el tamaño estimado de la muestra entre un 10 y un 20% para tomar en cuenta los hogares que no se cubrieron, o donde no hay una mujer elegible o donde la persona seleccionada para la entrevista no está disponible o se niega a participar en el estudio.

¿Qué tan grande debe ser la muestra?

Esta es una de las preguntas más comunes que se hacen a especialistas en estadísticas. Una respuesta frecuente pero errónea es "tan grande como sea posible", cuando en realidad debería ser "tan pequeña como sea posible".

También es importante subrayar que la cantidad de información que se puede obtener de una muestra depende de su tamaño absoluto y no de la fracción del muestreo, o de su tamaño como proporción del tamaño de la población. De hecho es cierto que 99 en un millón dice tanto sobre el millón como 99 en mil dice sobre ese mil.

(Tomado de Persson y Wall, 2003).[6]

La obtención de un marco muestral

Como se señaló anteriormente, para seleccionar una muestra aleatoria es necesario tener una lista o un mapa de todos los hogares o personas, dependiendo de la unidad de muestreo, en la población de estudio. Cuando sea posible se deben utilizar **marcos muestrales** elaborados con anterioridad, como por ejemplo:

- Datos oficiales de los censos
- Registros de los votantes

CUADRO 7.4 MUESTREO CREATIVO EN LIBERIA

En un estudio realizado en Liberia sobre violencia contra la mujer durante la guerra civil, se realizaron encuestas en diferentes sitios, tales como plazas de mercado, escuelas secundarias, comunidades urbanas y campamentos de desplazados, con el ánimo de ampliar al máximo la representatividad del estudio. Cada sitio requirió una estrategia de muestreo distinta. Aunque cada estrategia arrojó una buena aproximación a la situación de las mujeres que pertenecían al grupo meta específico de la muestra, cabe señalar que no se pudo combinar a los tres grupos siguientes para obtener el cálculo de prevalencia entre las mujeres liberianas en su conjunto.

- **Muestreo de mujeres mayores:** La unidad de muestreo en las plazas de mercado era un puesto de venta. Las encuestadoras elaboraron un mapa detallado del número y ubicación de los puntos entre los cuales se seleccionó al azar el intervalo de muestreo calculado y las mesas específicas. Ya que la lotería es común en Liberia, el equipo preparó unos coloridos "boletos de la suerte" para las mesas que habían seleccionado de manera aleatoria. Después de presentarse a una comerciante del mercado, la entrevistadora colocaba el boleto sobre la mesa al mismo tiempo que decía "boleto de la suerte" e invitaba a la mujer a participar en el estudio.
- **Muestreo de mujeres jóvenes:** El equipo elaboró un plan de muestreo para las escuelas de secundaria a partir de información suministrada por docentes sobre el número de muchachas de noveno a doceavo grado que asistirían a clase ese día. Después de calcular el intervalo de muestreo, las muchachas en cada clase contaban desde "uno" hasta el intervalo de muestreo; por ejemplo, si el intervalo de muestreo era cuatro, a la muchacha que le tocaba el número cuatro en el conteo se le invitaba a la encuesta. El muestreo se realizó públicamente de manera que todos supieran que no se había seleccionado a las muchachas porque tuvieran alguna experiencia en particular, previniendo así cualquier estigma que pudiera atribuirse al hecho de ser seleccionada para participar en un estudio sobre violencia sexual.
- **Muestreo de refugiados:** La disponibilidad de datos de censo para la distribución de alimentos facilitó el muestreo en las comunidades y en los campamentos de personas desplazadas. Después de calcular el intervalo de muestreo, se realizó una selección sistemática que empezó por escoger de forma aleatoria un hogar utilizando los últimos dos dígitos del número de serie de un billete de $5 dólares liberianos seleccionado al azar.

(Tomado de Swiss et al., 1998).[8]

- Censos realizados por el Ministerio de Salud para campañas de inmunización o contra la malaria
- Marcos de muestreo elaborados por otros estudios de gran escala realizados en el país, tales como encuestas demográficas y de salud o encuestas de medición del nivel de vida

Cada una de esas opciones deberá revisarse con cuidado para evaluar la calidad de la información. En ciertos tipos de datos oficiales, ciertos individuos podrían haberse omitido, por ejemplo, extranjeros residentes o empleados domésticos. Asimismo, a menos que los datos se hayan recopilado recientemente, es probable que sea necesario actualizar los registros. De todas maneras esto requerirá menos tiempo y será menos costoso que elaborar un marco muestral propio. Si no hay nada disponible de antemano entonces será preciso elaborar un marco muestral propio. A menudo esto se hace enumerando (levantado una lista y un mapa) los hogares o las personas elegibles para la entrevista en los conglomerados seleccionados. En el cuadro 7.4 se da un ejemplo de algunos otros enfoques utilizados en Liberia en una situación de conflicto.

¿Quién es elegible para la entrevista?

En otras secciones se aborda la importancia de que por razones de seguridad sólo se entreviste a una mujer por hogar. Sin embargo, cada equipo investigador deberá considerar con cuidado qué criterios utilizará para la elegibilidad. En otras palabras, ¿cuáles mujeres dentro del hogar pueden considerarse como posibles fuentes de información? Por ejemplo, ¿se incluirá a las empleadas domésticas?, ¿qué hacer con las amistades o las parientes que estén de visita, o inquilinas —estudiantes— que estén alquilando un cuarto en la casa?

Las y los investigadores han resuelto estos asuntos de diferentes maneras y cada solución tiene sus propias ventajas y desventajas. Por ejemplo, si se utiliza como criterio básico el hecho de que una mujer haya dormido en el hogar la noche anterior, entonces se podría incluir a las empleadas domésticas en su lugar de trabajo. Sin embargo, en ese caso, gran parte de la información sobre la situación económica recopilada sobre el hogar no reflejará la verdadera condición de esa mujer. Por otra parte, si las empleadas domésticas no se consideran elegibles en su lugar de trabajo, entonces es importante asegurarse que puedan tener la oportunidad de que se les incluya en sus hogares

permanentes. Esto coloca una carga adicional para los y las encuestadoras debido a que para evitar una alta tasa de falta de respuestas, en algunos casos será necesario regresar en días específicos o por las noches cuando la mujer esté en casa. La clave para no introducir sesgos en la muestra es recordar que los criterios de elegibilidad deben determinarse de tal manera que todas las mujeres elegibles dentro del conglomerado tienen una misma oportunidad de ser seleccionadas como fuentes de información, ya sea en sus hogares, su lugar de trabajo o en la escuela.

El estudio de la OMS sobre VCM decidió considerar como fuentes de información elegibles a las inquilinas permanentes y a las empleadas domésticas que pasen menos de dos noches por semana fuera del hogar. No se incluiría a aquellas mujeres que trabajaban en el hogar como empleadas domésticas pero que pasaban por lo menos dos noches a la semana fuera del mismo. Estas mujeres podrían potencialmente resultar seleccionadas para entrevistarlas en sus propios hogares en sus días libres.

Para evitar los sesgos en la muestra, es importante que los y las encuestadoras seleccionen a la mujer que van a entrevistar utilizando métodos aleatorios y no simplemente la primera mujer que abra la puerta o la de mayor edad o la que parece estar más disponible. En el estudio de la OMS sobre VCM esto se logró sacando los nombres de las mujeres de una bolsa. En Nicaragua se utilizó un cuadro de números aleatorios para seleccionar a las mujeres que serían entrevistadas. En el cuadro 7.5 se describe cómo utilizar esos métodos. Si se utiliza una lista de nombres de mujeres generada a partir de un marco de muestreo de personas, y es posible establecer si dos mujeres están viviendo en el mismo hogar, entonces se puede seleccionar a una mujer de cada par alternando entre la mujer de más edad y la más joven.

CUADRO 7.5 DOS MÉTODOS PARA SELECCIONAR A QUIEN ENTREVISTAR

FOTO DE HAFM JANSEN

Lotería

Elabore una lista de todas las mujeres de la casa que sean elegibles para participar en el estudio, según los criterios establecidos. Por ejemplo, ser mayor de 15 años y haber tenido pareja alguna vez. Asegúrese de incluir a todas las mujeres aunque no se encuentren en la casa en el momento de la visita. Escriba el nombre de cada una de las mujeres en un pedazo de papel diferente y coloque los papeles en una bolsa de papel. Pida a alguna persona del hogar, quizá la de mayor edad, que saque un papel de la bolsa. Las ventajas de este método son la simplicidad de su administración y que para todos los de la casa quede claro que la selección de la persona se realizó al azar y no por alguna característica especial. La desventaja de este método es que es difícil controlar si los entrevistadores utilizan el sistema correctamente.

Gráfica aleatoria

Una gráfica aleatoria puede ser útil para seleccionar a la mujer que se entrevistará una vez que haya completado la lista de todos los miembros de la casa. Este método permite controlar con bastante facilidad si el personal de campo utiliza correctamente el procedimiento, pero es menos transparente para los miembros del hogar.

Revise el último dígito del número que aparece en el cuestionario, el cual indica qué fila utilizar. Luego, revise en la lista de los miembros de la casa la cantidad total de mujeres entre 15 y 49 años que alguna vez han estado casadas o han tenido pareja, para determinar qué columna utilizar. El número que aparece en la intersección de la fila y la columna representa la cantidad de mujeres que se debe entrevistar.

Último dígito en el cuestionario	*Cantidad de mujeres elegibles en la casa*							
	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***
0	1	2	2	4	3	6	5	4
1	1	1	3	1	4	1	6	5
2	1	2	1	2	5	2	7	6
3	1	1	2	3	1	3	1	7
4	1	2	3	4	2	4	2	8
5	1	1	1	1	3	5	3	1
6	1	2	2	2	4	6	4	2
7	1	1	3	3	5	1	5	3
8	1	2	1	4	1	2	6	4
9	1	1	2	1	2	3	7	5

1. Crabtree B, Miller W. *Doing Qualitative Research*. Newbury Park: Sage; 1992.
2. Hudelson P. *Qualitative Research for Health Programs*. Geneva: Division of Mental Health, World Health Organization; 1994.
3. Patton M. *Qualitative Evaluation and Research Methods*. 2nd ed. Newbury Park: Sage Publications; 1990.
4. Montoya O. *Nadando contra Corriente: Buscando Pistas para Prevenir la Violencia Masculina en las Relaciones de Pareja*. Managua, Nicaragua: Fundación Puntos de Encuentro; 1998.
5. Morison L. Self-weighting in Cluster Samples. Personal Communication. London; 2002.
6. Persson LÅ, Wall S. *Epidemiology for Public Health*. Umeå, Sweden: Umeå International School of Public Health; 2003.
7. Guezmes A, Palomino N, Ramos M. *Violencia Sexual y Física contra las Mujeres en el Perú*. Lima: Flora Tristan, Organización Munidial de la Salud, Universidad Peruana Cayetano Heredia; 2002.
8. Swiss S, Jennings PJ, Aryee GV, et al. Violence against women during the Liberian civil conflict. *Journal of the American Medical Association*. 1998;279:625 - 629.