

Encuesta de infección por coronavirus tipo 2 del síndrome respiratorio agudo grave (SARS-CoV-2), a nivel comunitario en habitantes de un barrio vulnerable urbano de la ciudad de Buenos Aires: un ejemplo de muestreo probabilístico

Vanina Pagotto, Mercedes Soriano, Diego H. Giunta y Silvana Figar

RESUMEN

Entre el 1 y el 26 de junio se llevó a cabo el estudio de investigación “Encuesta de infección por coronavirus tipo 2 del síndrome respiratorio agudo grave (SARS-CoV-2), nivel comunitario en habitantes de un barrio vulnerable urbano de la ciudad de Buenos Aires”, que determinó que un 54,3% de los habitantes del barrio presentaban anticuerpos inmunoglobulina tipo G para SARS-CoV-2. El objetivo de este artículo es proporcionar un ejemplo de un muestreo probabilístico que fue utilizado para estimar la prevalencia de seropositividad en este estudio.

Palabras clave: muestreo, probabilidad, COVID-19.

SURVEY OF INFECTION BY SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2 (SARS-COV-2), COMMUNITY LEVEL IN INHABITANTS OF A MARGINAL URBAN NEIGHBORHOOD OF THE CITY OF BUENOS AIRES: AN EXAMPLE OF A PROBABILITY SAMPLING

ABSTRACT

Between 1st and 26th of June, a research named “Survey of infection by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), community level in inhabitants of a marginal urban neighborhood of the city of Buenos Aires” was carried on. The study showed that 54.3% of the that 54.3% of the people of the neighborhood had antibodies immunoglobulin type G for SARS-CoV-2. The objective of this article is to provide an example of a probability sampling carried out in the study, to measuring the prevalence of seropositivity.

Key words: sample, probability, COVID-19.

Rev. Hosp. Ital. B.Aires 2020; 40(4): 233-236.

Obtener una muestra o realizar un muestreo implica seleccionar una serie de unidades de una población con el propósito de que la información recolectada de esas unidades permita realizar inferencias sobre toda la población¹. En la práctica se habla de muestreo probabilístico y de muestreo no probabilístico. En un muestreo de tipo probabilístico se realiza una selección aleatoria en la que cada elemento de la población tiene una probabilidad conocida no nula de ser seleccionado, con lo cual cada elemento de la muestra representa a un sector de la población y su totalidad a toda la población. Por lo tanto, a partir de la muestra se pueden hacer inferencias sobre el total de la población. En un muestreo no probabilístico, las muestras no son aleatorias, por lo que podrían no ser representativas de la población de donde se extrajeron².

Se presenta a continuación el muestreo probabilístico empleado para la realización de la encuesta de seroprevalencia de SARS-CoV-2 a nivel comunitario en habitantes del Barrio Mujica, un barrio vulnerable urbano de la ciudad de Buenos Aires, repartido entre la Comuna 1 (donde se ubica el barrio Retiro) y la Comuna 2 (donde se ubica el barrio Recoleta).

El Barrio Mujica está integrado por 2 barrios: 31 y 31 Bis, que tienen en total 129 cuadras, denominadas manzanas, que se agrupan en 10 sectores: Bajo Autopista, Comunicaciones, Cristo Obrero, Ferroviario, Güemes, Inmigrantes, Playón Este, Playón Oeste, San Martín e YPF. En la figura 1 se observa el mapa del Barrio Mujica con los 10 sectores que lo conforman.

Cada sector del barrio tiene una determinada densidad poblacional, medida por la cantidad de viviendas, que está disponible en el censo del año 2017 realizado por el Departamento de Estadísticas y Censos del Gobierno de la Ciudad de Buenos Aires (DGEyC)⁴, del cual se obtuvo la cartografía necesaria para la planificación de diferentes

Recibido: 29/7/20

Aceptado: 11/9/20

Departamento de Investigación (V.P., D.H.G., S.F.). Sección Epidemiología (M.S.). Hospital Italiano de Buenos Aires. Argentina. Correspondencia: vanina.pagotto@hospitalitaliano.org.ar

esquemas de muestreo. La información de la cantidad de viviendas es la que disponía la DGEyC; es posible que, en el transcurso de tres años, el Barrio haya presentado un crecimiento poblacional, lo que podría considerarse como una debilidad de la estrategia de muestreo.

Para llevar a cabo las encuestas se realizó la agrupación de las manzanas en áreas de unidades primarias de muestreo (Áreas-UP). Por definición, estas áreas son un conjunto de casas contiguas de un tamaño aproximado de 350 casas. En total se construyeron 36 Áreas-UP.

Se diseñó un muestreo en dos etapas. El primer nivel fue cada uno de los 10 sectores que conforman el Barrio 31 y 31 Bis, los cuales fueron considerados estratos. En esta primera etapa se seleccionaron Áreas-UP en cada uno de los estratos con probabilidad proporcional al tamaño; la medida del tamaño fue el número total de viviendas. En la figura 2 muestra la estrategia de muestreo probabilístico. El gráfico superior es el plano del Barrio Mujica; los colores representan los 10 sectores en que está dividido. El segundo gráfico muestra la imagen aumentada de unos de los sectores llamado Ferroviario y, dentro de este sector, una manzana correspondiente a un Área-UP. La tabla superior es la cantidad de manzanas de ese sector y la tabla inferior las casas seleccionadas. De cada casa seleccionada se invitó a participar a todos los habitantes de 14 años o más. El tamaño de la muestra de la primera etapa fue 20 Áreas-UP. En la segunda etapa se seleccionaron 30 viviendas de cada Área-UP mediante un muestreo sistemático, el cual también es un muestreo probabilístico. La estrategia fue la siguiente: se seleccionó en forma aleatoria la vivienda de arranque, la ruta establecida y el intervalo de muestreo a partir del cual se seleccionaron las 29 casas restantes. En la figura 3 se observa un ejemplo de un Área-UP. Dentro de cada vivienda se incluyeron todos los habitantes de 14 años o más que

aceptaron participar del estudio. A fin de mejorar la respuesta, la vivienda seleccionada fue visitada en dos días distintos en horarios diferentes y se utilizó un material informativo impreso con un número telefónico y mencionando la posibilidad de realizar el test y la encuesta en el Polo Educativo María Elena Walsh, lugar de referencia Dispositivo Estratégico de Testeo para Coronavirus en Territorio de Argentina (DETECTAr) en el Barrio Mujica. Los colores representan los diferentes lados para acceder a las casas. El triángulo morado muestra la primera casa y el punto negro es la última casa de la manzana. El intervalo de selección en esta manzana fue cada 15 casas. Las flechas negras indican la dirección para realizar la selección de casas. Las flechas rojas indican las casas seleccionadas. Los números azules son la cantidad de pisos en cada casa y la leyenda roja UE muestra si hay alguna actividad económica (como kiosco, peluquería u otro negocio). Los números negros son los números asignados a cada casa; cuando no hay un número, se muestra S/N.

Se estimó un tamaño muestral de 406 hogares, para una seroprevalencia estimada de 7% con una precisión del 3,5%; considerando la falta de respuesta y la agrupación potencial de seropositividad por hogar se estimó el efecto de diseño de 2.

$$n = Z^2 \frac{p(1-p)}{\delta^2} EDF$$

donde p es la proporción para estimar (7%), Z es el cuartil 95% de una distribución normal estándar²; es la semiamplitud del intervalo de confianza (3,5%) y EDF es el efecto de diseño² de muestra⁵.

El EDF puede interpretarse como el número de unidades colectadas por un procedimiento de muestreo específico que son equivalentes, para propósitos de estimación, con una unidad colectada mediante muestreo aleatorio simple.



Figura 1. Mapa del Barrio Mujica. Los 10 sectores están representados en diferentes colores en el mapa, las áreas en verde corresponden a parques o plazas, o canchas de fútbol. Imagen extraída de la Secretaría de Integración Social y Urbana del Gobierno de la Ciudad de Buenos Aires³.

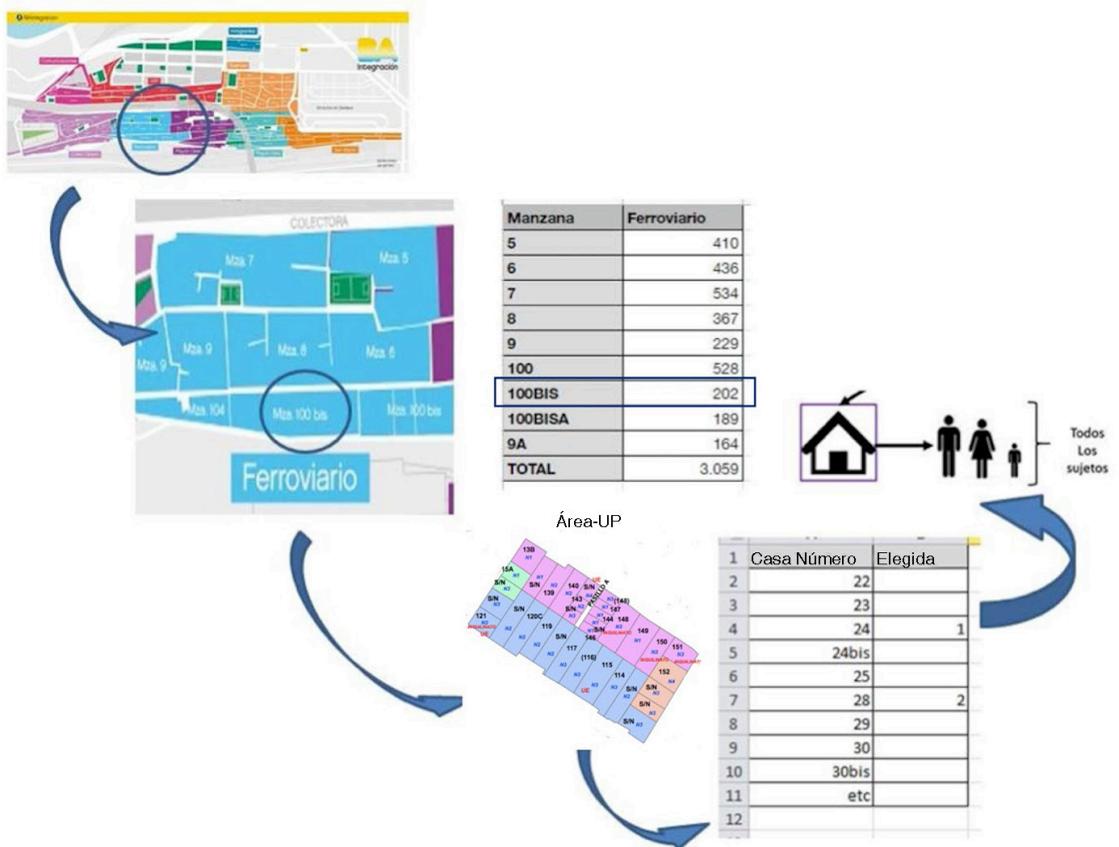


Figura 2. Estrategia de muestreo empleada en el estudio “Encuesta de infección por coronavirus tipo 2 del síndrome respiratorio agudo grave (SARS-CoV-2), a nivel comunitario en habitantes de un barrio vulnerable urbano de la ciudad de Buenos Aires”

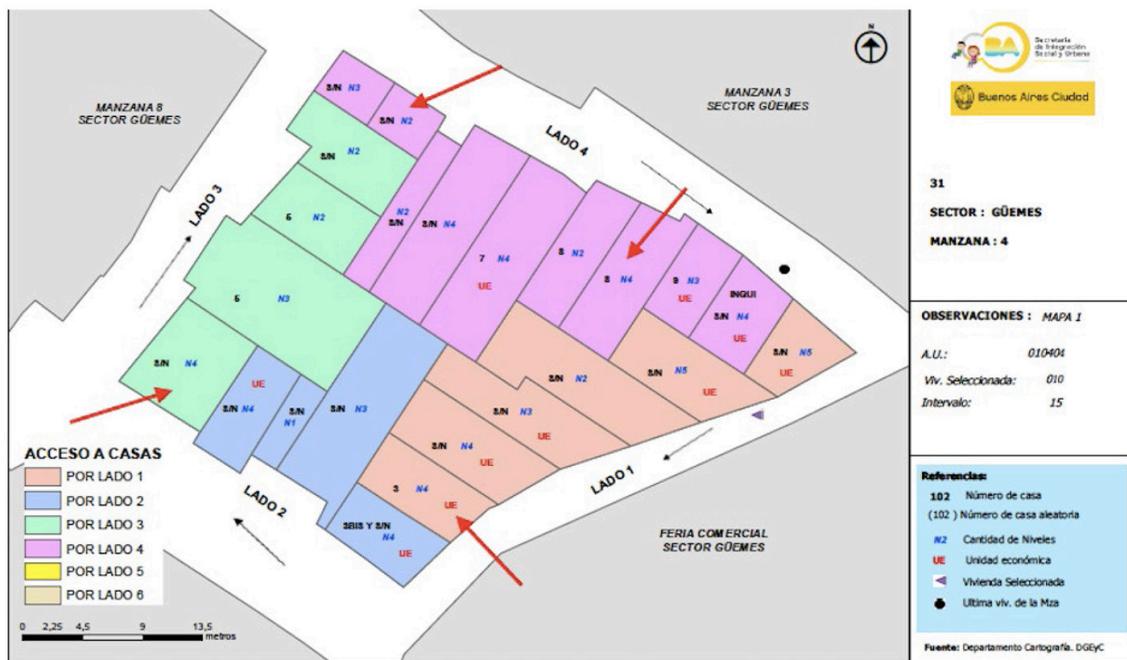


Figura 3. Ejemplo de un Área-UP utilizada en el muestreo del estudio “Encuesta de infección por coronavirus tipo 2 del síndrome respiratorio agudo grave (SARS-CoV-2), a nivel comunitario en habitantes de un barrio vulnerable urbano de la ciudad de Buenos Aires”.

Con el muestreo se obtuvo una prevalencia de positividad de anticuerpos para SARS-CoV 2, la cual fue expandida al censo del barrio por 3 factores: a nivel del barrio, a nivel hogar y a nivel de cada individuo. Este procedimiento permitió obtener la prevalencia ponderada (W) de 54,3%, calculando W como la inversa de la probabilidad de selección de la última unidad de muestreo (hogar)^{6,7}. Los 3 factores de la expansión mencionados al nivel del hogar implicaron 3 tipos de ajustes: el primero relacionado con la no respuesta, dado que algunos hogares no quisieron participar en la encuesta; el segundo relacionado con la proyección de la muestra a toda la población y el tercero con la técnica

de calibración utilizando grupos de edad y sexo, por medio información de la población del censo. De esta manera, las variables de calibración fueron las personas de 14 años o más, agrupadas por sexo y por intervalos de edad: 14-30; 31-45; 46-59; 60 y más.

Como conclusión, en este estudio, la utilización de un muestreo probabilístico multietápico, ya que la muestra se seleccionó en 2 etapas sucesivas, permitió una estrategia de muestreo factible y eficaz para los recursos disponibles en el estudio global, que posibilitó estimar la seroprevalencia con un margen preciso en un estudio realizado en 16 días que incluyó una muestra de 398 viviendas, 577 familias y 872 habitantes del Barrio Mujica.

Conflictos de interés: los autores declaran no tener conflictos de interés.

REFERENCIAS

1. Pimienta Lastra R. Encuestas probabilísticas vs. no probabilísticas. *Política y Cultura*. 2000;13:263-276.
2. Statistics Canada. Business, Provincial and Municipal Relations Division. Catalogue. Ontario: Statistics Canada; 1985.
3. Buenos Aires Ciudad. Barrio 31: obras de integración social y urbana en el Barrio 31 [Internet]. Buenos Aires: Gobierno de la Ciudad Autónoma de Buenos Aires; 2020? [citado 2020 jul 28]. Disponible en: <https://www.buenosaires.gob.ar/baobras/barrio-31>.
4. Buenos Aires Ciudad. Estadística y censos [Internet]. Buenos Aires: Gobierno de la Ciudad Autónoma de Buenos Aires; 2020? [citado 2020 jul 27]. Disponible en: <https://www.estadisticaciudad.gob.ar/eyc/>.
5. Dean AG, Sullivan KM, Soe MM, et al. OpenEpi: estadísticas epidemiológicas de código abierto para salud pública [Internet]. Versión 3.0, actualizado 2013/04/06 [citado 2020 jul 27]. Disponible en: https://www.openepi.com/Menu/OE_Menu.htm.
6. Romero-Martínez M, Shamah-Levy T, Cuevas-Nasu L, et al. Diseño metodológico de la Encuesta Nacional de Salud y Nutrición de Medio Camino 2016. *Salud Pública Méx*. 2017;59(3):299-305.
7. R Core Team. R: a language and environment for statistical computing [internet]. Vienna: R Foundation for Statistical Computing; 2018 [citado 2020 jul 27]. Disponible en: <http://www.R-project.org/>.