

Metodologías cuantitativas: Cálculo del tamaño de muestra con STATA y R

Quantitative methods: Sample size calculation with STATA and R

Antonio M Quispe^{1,a}, Diego F Pinto^{2,b}, Mariella R Huaman^{2,b}, Gilda M Bueno^{3,c}, Andree Valle-Campos^{4,d}

RESUMEN

El cálculo de tamaño de muestra es un aspecto esencial del diseño de estudios cuantitativos. Un adecuado tamaño de muestra nos permite determinar cuál es la mínima cantidad de participantes necesarios para probar nuestra hipótesis de interés. De esta manera, podemos reducir costos, maximizar el uso de nuestros recursos de investigación y garantizar la factibilidad del estudio. Contradictoriamente, a pesar de su relevancia muy pocos investigadores dominan esta habilidad. Esta revisión tiene por objeto revisar los conceptos básicos para realizar un cálculo de tamaño de muestra y compartir códigos de Stata y R específicamente diseñados para facilitar estos cálculos

Palabras clave: Tamaño de Muestra; Diseño de Investigaciones Epidemiológicas; tamaño de efecto; Stata; R (Fuente: DeCS-BIREME).

ABSTRACT

The calculation of sample size is an essential aspect of the design of quantitative studies. An adequate sample size allows us to determine the minimum number of participants necessary to test our hypothesis of interest. Hence, we can reduce costs, maximize the use of our research resources and guarantee the feasibility of the study. Contradictorily, despite its relevance, very few researchers dominate this skill. This review aims to review the basics of sample size calculation and share Stata and R codes specifically designed to facilitate these calculations.

Keywords: Sample Size, Epidemiologic Research Design; effect size; Stata; R (Source: DeCS-BIREME).

INTRODUCCIÓN

El cálculo de tamaño de muestra es un aspecto esencial del diseño de estudios cuantitativos⁽¹⁾. Un adecuado cálculo del tamaño de muestra no solo nos permite determinar cuál es la mínima cantidad de participantes necesarios para probar nuestra hipótesis de interés sino determinar la factibilidad del mismo y ajustar nuestros presupuestos⁽²⁾. Esto es muy importante porque los

recursos para realizar una investigación casi siempre son escasos y trabajar con toda la población de estudio no siempre es factible. De ahí que la alternativa es analizar una muestra representativa de la población de estudio utilizando la estadística inferencial para extrapolar nuestros hallazgos a toda la población objetivo⁽³⁾. Además del aspecto metodológico, un óptimo tamaño de muestra permite minimizar el riesgo de eventos adversos limitando la cantidad de participantes al mínimo y estrictamente necesario⁽⁴⁾. Si bien existe una gran variedad de escenarios posibles de necesidad de cálculo de tamaño de muestra, lo cierto es que unos pocos escenarios representan la mayoría de los cálculos de tamaño de muestra utilizados en la literatura⁽⁵⁾.

En la práctica, todos los supuestos necesarios para realizar un cálculo del tamaño de muestra deben ser prefijados durante la fase de planeamiento del estudio. Y si por algún motivo se requiriese un ajuste del tamaño de muestra después de haber iniciado la recolección de datos del estudio este debe ser comunicado al Comité de Ética Institucional que aprobó el protocolo del estudio en primera instancia⁽⁶⁾. Sin embargo, pese a la relevancia de realizar un correcto cálculo de tamaño de muestra muchos estudios omiten reportar el mismo o lo reportan mal⁽⁷⁾. Estudios previos han estimado en aproximadamente un 54% de los estudios publicados por investigadores de países desarrollados cumplió con reportar su cálculo de tamaño de muestra, pero solo una tercera parte lo hizo de manera correcta y reproducible⁽⁸⁾. De la misma manera, se ha reportado

1. Universidad Continental, Huancayo, Perú.
2. Sociedad Científica de San Fernando, Universidad Nacional Mayor de San Marcos, Lima, Perú.
3. Facultad de Medicina Humana, Universidad Nacional Mayor de San Marcos, Lima, Perú.
4. Universidad Nacional Mayor de San Marcos, Lima, Perú.
a. Médico Epidemiólogo.
b. Estudiante de Medicina Humana.
c. Médico cirujano.
d. Biólogo Epidemiólogo.

que las principales causas de los errores en el cálculo de tamaño de muestra es el uso de ecuaciones obsoletas que no toman en cuenta todos los parámetros necesarios para dicho cálculo según diseño del estudio⁽⁹⁾. Esto se hace evidente con el creciente número de cartas al editor que se publican criticando estudios con un incorrecto cálculo del tamaño de muestra, una mala interpretación de los resultados y un reporte inadecuada del dicho cálculo^(10,11). Probablemente debido a las fórmulas complejas o conceptos abstractos que conllevan calcularlo, se ha descrito la dificultad en el aprendizaje e interiorización de tópicos estadísticos, incluido para médicos y futuros médicos^(12,13).

El cálculo del tamaño de muestra es una habilidad esencial para investigar y publicar estudios cuantitativos porque facilita el análisis de factibilidad y probar hipótesis de una manera más rentable y sostenible⁽¹⁴⁾. Debido a todo lo anteriormente expuesto, el objetivo del presente artículo es ofrecer una revisión de los conceptos básicos para realizar un cálculo de tamaño de muestra y compartir códigos de Stata y R específicamente diseñados para facilitar estos cálculos. Para ello primero revisaremos los conceptos básicos necesarios para realizar un cálculo de tamaño de muestra conservador y reproducible. Segundo, revisaremos los ocho escenarios más comunes en los cuales se requiere de un cálculo de tamaño de muestra en estudios cuantitativos. Y finalmente, compartiremos códigos de Stata y R específicamente diseñados para facilitar estos cálculos.

¿POR QUÉ CALCULAR UN TAMAÑO DE MUESTRA?

Todo trabajo de investigación en el que se opta por utilizar una técnica de muestreo estadístico tiene por objeto utilizar la inferencia estadística para extrapolar los resultados de la muestra a la población objetivo del estudio. El problema es que durante este proceso se puede introducir una serie de sesgos que atentan contra la validez interna y externa del estudio. Entre estos, uno de los errores más comunes es un mal cálculo del tamaño de muestra necesario para probar la hipótesis de interés del estudio, ya sea porque no escogió la fórmula correcta o porque, aún escogiéndola no se penalizó conservadoramente la muestra⁽¹⁵⁾. Claramente, cuanto más grande la muestra mayor la representatividad de la misma, menor la variabilidad del outcome y más precisas serán los estimados del estudio. Sin embargo, dado que los recursos siempre son finitos, el objetivo del muestreo es estimar el mínimo tamaño de muestra que garantice que al término del estudio se podrá probar la hipótesis principal del estudio.

CONCEPTOS BÁSICOS PARA REALIZAR UN CÁLCULO DE TAMAÑO DE MUESTRA

El alfa y la significancia estadística (valor p)

Teóricamente, el alfa es el error tipo I, es decir, aquel error que se comete al rechazar la hipótesis nula cuando esta es verdadera. Mientras que la probabilidad de alfa o significancia estadística (valor p) es la probabilidad de rechazar la hipótesis nula (p. e. que la intervención de interés no tiene ningún efecto sobre el outcome de interés) cuando esta es verdadera. De manera práctica la significancia estadística significa que el resultado obtenido es poco probable que se deba al azar. Dicho esto, es importante precisar que no encontrar significancia estadística (por convención, un valor $p < 0,05$), no significa que podamos aceptar la hipótesis nula o que hayamos demostrado que el efecto de interés no existe⁽¹⁶⁾.

Esto se debe porque al no rechazar la hipótesis nula pueden darse dos escenarios, uno que efectivamente que la hipótesis nula sea cierta o que no tuvimos el poder de estudio suficiente para detectar la magnitud del efecto de interés. De ahí que si no encontramos un valor p significativo no podemos saltar a la conclusión de que el estudio ha probado que el efecto de interés no existe⁽¹⁷⁾. Por otro lado, y de manera similar, si el valor p resulta ser significativo no podemos saltar a la conclusión de que el efecto encontrado además de estadísticamente significativo también es clínicamente significativo⁽¹⁸⁾. Ambos conceptos son por definición diferentes, siendo la significancia clínica una determinación práctica de qué efecto es el mínimo que se puede aceptar como clínicamente relevante para los pacientes.

La beta y el poder de estudio

Teóricamente, la beta es el error tipo II o error que se comete al no rechazar la hipótesis nula cuando esta es falsa. Mientras que el poder de estudio (que se calcula como $1-\beta$) es la probabilidad de que rechazar la hipótesis nula cuando esta es efectivamente falsa. De manera práctica, el poder de estudio es la probabilidad de encontrar el efecto de interés cuando este existe en la realidad. De ahí que mientras mayor el poder de estudio (y menor el β), mayor será la probabilidad de probar la hipótesis primaria de un estudio⁽¹⁹⁾.

Para calcular el tamaño de muestra en un estudio comparativo necesitamos conocer el valor de los siguientes cuatro elementos. Los dos primeros son el nivel de significancia estadística (alfa) y el poder del estudio ($1-\beta$), de los cuales hemos hablado en párrafos anteriores. Los otros dos elementos importantes para realizar esta tarea son el tamaño de efecto que se espera encontrar y la variabilidad del outcome en la población. Al momento de seleccionar los valores para dichos elementos debemos basarnos en referencias previamente publicadas; o de lo contrario, en el caso del tamaño de efecto que se espera encontrar, podemos elegir el tamaño del efecto que deseamos ver (por debajo del cual, digamos, el coste de un tratamiento no vale la pena). Por supuesto, se trata de estimaciones y no se sabe si la variación de la muestra será la misma, lo

que puede significar que, aunque se alcance el tamaño estimado de la muestra y se vea un efecto, puede no alcanzar la significación estadística⁽²⁰⁾.

La magnitud del efecto de interés o la diferencia esperada

Cuando planteamos una hipótesis estamos obligados a precisar la magnitud esperada del efecto de interés o la diferencia esperada⁽²¹⁾. Esto se debe porque precisamente este valor es necesario para calcular el tamaño de muestra del estudio. Este estimado suele tomarse de los antecedentes del estudio, pero cuando estos no están disponibles o no son confiables la recomendación es realizar un estudio piloto con una muestra pequeña. Independientemente de la fuente, estos son solo estimados por lo que la recomendación general es no realizar un solo cálculo de tamaño de muestra sino varios, y eventualmente escoger el cálculo más conservador. Esto se debe a que es mejor escoger el cálculo de tamaño de muestra más conservador, porque de escoger cualquiera de los cálculos más pequeños corremos el riesgo de a pesar de alcanzar el tamaño de muestra estimado no encontrar la significación estadística para probar la hipótesis del estudio⁽⁹⁾.

Es recomendable que el cálculo de la muestra se realice en función del tamaño de efecto que el investigador desea observar. Este caso podría dar lugar a un tamaño de muestra muy grande. Sin embargo, se puede optar por modificar el nivel de significancia o el poder de estudio. Puede ser tentador, pero probablemente mala idea, modificar el tamaño del efecto⁽²²⁾. Puede no parecer obvio al principio, pero buscar efectos de tratamiento más grandes generalmente significa que se necesitan tamaños de muestra más pequeños. Si el efecto de un tratamiento es muy grande, es probable que se observe muy pronto en el experimento, y esto sugeriría que no se necesita una muestra muy grande. Puede ser bastante tentador diseñar un estudio de esta manera para reducir el tiempo de reclutamiento y los costos. No obstante, pocos tratamientos tienen un tamaño de efecto tan grande en la vida real. Por lo que, si un investigador decide usar un tamaño de efecto demasiado grande, no se podría detectar una diferencia "estadísticamente significativa" con el tamaño de muestra seleccionado. Incluso, si existe una diferencia clínicamente importante⁽²⁰⁾.

La variabilidad del outcome

La variabilidad de la respuesta al tratamiento es un elemento importante en el cálculo del tamaño de muestra. Nos ayuda a determinar si el efecto observado es real o puede estar influenciado por el azar. Si la variabilidad del outcome es grande, la muestra necesaria también deberá serlo y viceversa, si la variabilidad es pequeña, la muestra requerida también será menor⁽²³⁾.

La variabilidad de las variables continuas es cuantificada mediante la desviación estándar (σ). La desviación estándar mide la variación o dispersión de los datos. Puede conceptualizarse como el promedio de las desviaciones de la media (μ). El valor de σ puede ser difícil de obtener, pero el investigador puede optar por realizar un estudio piloto previo o basarse en el σ reportado en anteriores estudios similares⁽²⁴⁾. Para variables dicotómicas, la variabilidad depende de la proporción de sujetos con el resultado de interés. Así, para la comparación de dos proporciones, el tamaño de la muestra se calcula utilizando los valores de las dos proporciones de interés: la proporción de control (tasa de eventos de referencia) puede obtenerse de la literatura publicada, mientras que la proporción de tratamiento (tasa de eventos) se deriva una vez elegido el tamaño del efecto. El cálculo del tamaño de la muestra se basa en la variación de la población y el tamaño del efecto de las dos proporciones de interés.

ESCENARIOS MÁS COMUNES DE CÁLCULO DE TAMAÑO DE MUESTRA

En medicina se pueden listar por lo menos ocho escenarios como los más comunes entre aquellos que se requiere realizar un cálculo de tamaño de muestra (Tabla N°1). Estos escenarios son los siguientes:

- Cuando se requiere estimar la media (escenario 1) o la prevalencia (escenario 2) de un outcome en una misma población
- Cuando se quiere comparar la media (escenario 3) o la prevalencia (escenario 4) de un outcome entre dos poblaciones muestreadas de manera independiente
- Cuando se quiere comparar la media (escenario 5) o la prevalencia (escenario 6) de un outcome entre dos poblaciones que al ser muestreadas fueron pareadas por alguna característica o características en particular produciendo muestras por definición dependientes.
- Y cuando se quiere comparar la media (escenario 7) o la prevalencia (escenario 8) de un outcome entre dos o más poblaciones muestreadas de manera independiente.

Para calcular el tamaño de muestra mínimo necesario en cada uno de estos escenarios se requiere definir a priori cuáles son los supuestos bajo el cual se va a realizar dicho cálculo, incluyendo la hipótesis de estudio y la direccionalidad de la misma (una cola o dos colas), la magnitud aceptada de los errores alfa y beta, el valor esperado de la magnitud del efecto o la magnitud de la diferencia esperada, variabilidad del outcome, cuánto se espera de rechazos y pérdidas en el seguimiento^(15,16).

Estos supuestos son muy importantes porque tienen un impacto directo en el tamaño de la muestra estimado. A menor los errores alfa y beta mayor el tamaño de

Tabla N°1. Escenarios más comunes de cálculo de tamaño de muestra

Escenarios	H0	Ha
Una muestra		
-Escenario 1. ¿Cuál es la media (μ) del O1 en la población de estudio?	$\mu = \mu_0$	$\mu \neq \mu_0$
-Escenario 2. ¿Cuál es la prevalencia (p) del O2 en la población de estudio?	$p = p_0$	$p \neq p_0$
Dos muestras independientes		
-Escenario 3. ¿La media del O1 en la población 2 (μ_2) es diferente en la población 1 (μ_1)?	$\mu_2 = \mu_1$	$\mu_2 \neq \mu_1$
-Escenario 4. ¿La prevalencia del O2 en la población 2 (p_2) es diferente en la población 1 (p_1)?	$p_2 = p_1$	$p_2 \neq p_1$
Dos muestras pareadas		
-Escenario 5. ¿La media del O1 es diferente (d) en los pares divergentes de dos muestras pareadas?	$d = d_0$	$d \neq d_0$
-Escenario 6. ¿La prevalencia del O2 es diferente en los pares divergentes de dos muestras pareadas?	$p_{21} = p_{12}$	$p_{21} \neq p_{12}$
Más de dos muestras independientes		
-Escenario 7. ¿La media del O1 es diferente en alguna de las muestras independientes?	$\mu_1 = \mu_2 = \dots = \mu_k$	$\delta \neq 0$
-Escenario 8. ¿La prevalencia del O2 es diferente en alguna de las muestras independientes?	$p_1 = p_2 = \dots = p_k$	$\delta \neq 0$

O1, Outcome 1; O2, Outcome 2; H0, hipótesis nula; Ha, hipótesis alterna

Tabla N°2. Comandos de Stata para el cálculo de tamaño de muestra de los escenarios propuestos

Escenarios	H0	Ha
Una muestra		
-Escenario 1: power onemean 20 30, sd(10)	$\mu = 30$	$\mu \neq 30$
-Escenario 2: power oneproportion 0.1 0.2	$p = 20\%$	$p \neq 20\%$
Dos muestras independientes		
-Escenario 3: power twomeans 12 16, sd1(5.5) sd2(5)	$16 = 12$	$16 \neq 12$
-Escenario 4: power twoproportions 0.1 0.5	$50\% = 10\%$	$50\% \neq 10\%$
Dos muestras pareadas		
-Escenario 5: power pairedmeans 12 16, sd(5.5) corr(0.4)	$d = 4$	$d \neq 4$
-Escenario 6: power pairedproportions 0.1 0.4	$p_{21} = p_{12}$	$p_{21} \neq p_{12}$
Más de dos muestras independientes		
-Escenario 7: power oneway 50 60 70, varerror(900)	$50 = 60 = 70$	$\delta \neq 0$
-Escenario 8: artbin,pr(.5 .6 .7) ngroups(3) aratios(1 1 1) distant(0) alpha(0.05) power(.80)	$50\% = 60\% = 70\%$	$\delta \neq 0$

H0, hipótesis nula; Ha, hipótesis alterna

muestra. A menor la variabilidad del outcome menor el tamaño de muestra. Cuanto menor la magnitud del efecto o diferencia esperada mayor el tamaño de muestra. Y una vez hecho dicho cálculo es altamente recomendable penalizar la misma por la cantidad esperada de rechazos y pérdidas en el seguimiento⁽²⁵⁾.

SOFTWARE PARA CÁLCULO DE TAMAÑO DE MUESTRA

Existe un gran número de programas informáticos para realizar cálculos del tamaño de la muestra. Las capacidades de cada uno de ellos varían desde los cálculos básicos hasta los complejos, pero en la mayoría se integran diseños de ensayo comunes. Hay muchas calculadoras de muestras en línea e incluso algunas aplicaciones para teléfonos inteligentes que pueden

realizar cálculos del tamaño de la muestra⁽²⁶⁾. Si bien esas calculadoras pueden ser útiles, cabe señalar que es improbable que hayan sido validadas, y en general es aconsejable utilizar un programa informático validado⁽²⁷⁾.

CÁLCULO DE TAMAÑO DE MUESTRA CON STATA Y R

En el programa estadístico STATA, el comando "power" permite ejecutar el cálculo de tamaño de muestra, poder de estudio y magnitud del efecto o diferencia esperada para cada uno de los ocho escenarios propuestos en el presente manuscrito. Adicionalmente este comando ofrece la posibilidad de calcular varios cálculos a la vez y elaborar tablas y gráficos con los resultados de estos cálculos. Sin embargo, estos

últimos tienen como desventaja que solo produce resultados con textos en inglés. De ahí que para el presente manuscrito hemos escrito un código que no solo permite calcular el tamaño de muestra, poder de estudio y magnitud del efecto o diferencia esperada para cada uno de los ocho escenarios descritos en la Tabla 1 sino que también permite producir gráficos con textos traducidos al español. Este código podrá ser descargado utilizando el siguiente link: http://tiny.cc/crac_muestra. Brevemente, para cada escenario se tiene que predefinir los supuestos y correr el comando específico para cada caso tal como figura en la Tabla N°2.

CALCULO DE TAMAÑO DE MUESTRA CON R

En el lenguaje estadístico R, el paquete "pwr" permite ejecutar el cálculo de poder o tamaño de muestra dado un grupo fijo de parámetros o supuestos. Similar a Stata este paquete también permite generar gráficos con diferentes cálculos en simultáneo. Sin embargo, sus funciones no permiten evaluar rangos de valores por parámetro de poder, tamaño de efecto, media, proporción o desviación estándar. Para facilitar estos cálculos hemos creado el paquete "powder" el cual permite generar tablas y gráficos similares a los que se pueden obtener con Stata variando los parámetros y corriendo varios cálculos de tamaño de muestra a la vez. Este paquete es público y gratuito por lo que para descargarlo solo tiene que usar el siguiente link: <https://github.com/avallecam/powder>. Para crear este paquete hemos adaptado un grupo de funciones del paquete "pwr" de manera tal que ahora acepte secuencias de valores usando la función "seq(to,from,by)" y generar un objeto de clase "tibble". Adicionalmente, con las funciones "pwr_tidy" podemos construir las tablas especificando los distintos parámetros de ingreso como filas. Y con la función "pwr_plot" podremos construir gráficos especificando la hipótesis de interés, el método de cálculo y los parámetros de ingreso. Para su desarrollo se el paquete "broom" para facilitar la generación de tablas resumen ya que genera "tibbles" a partir de salidas estadísticas de tipo "power.htest" y el paquete "purrr" para facilitar el cálculo de múltiples escenarios al optimizar la programación funcional reemplazando el uso de loop por la familia de funciones "map". Finalmente, el paquete permite configurar el idioma de interés y así poder generar gráficos en español.

REPORTE DEL CÁLCULO DE TAMAÑO DE MUESTRA EN LOS MANUSCRITOS

Es esencial reportar los detalles que dieron origen al cálculo de la muestra en un estudio. Esto se realiza mencionando los supuestos y elementos utilizados. Los elementos más importantes a especificar corresponden al nivel de significancia, el poder de estudio, la variabilidad del desenlace y el tamaño de efecto

esperado. Adicionalmente deberá incluirse factores de otro tipo que hayan influido en el cálculo final, por ejemplo, la inflación realizada debido a un número estimado de pérdidas durante el seguimiento. La consideración de estos elementos en el reporte permite la reproducibilidad del estudio y denota un buen diseño de estudio. Vale la pena señalar que el reporte del procedimiento que dio origen al número de pacientes en el estudio es un requisito de la lista de verificación de las declaraciones CONSORT Y STROBE, utilizadas para el reporte de ensayos clínicos aleatorizados y estudio observacionales, respectivamente⁽²⁸⁾.

CONCLUSIONES

El cálculo de tamaño de muestra es un aspecto esencial del diseño de estudios cuantitativos. Un adecuado tamaño de muestra nos permite determinar cuál es la mínima cantidad de participantes necesarios para probar nuestra hipótesis de interés. De esta manera, podemos reducir costos, maximizar el uso de nuestros recursos de investigación y garantizar la factibilidad del estudio. Para facilitar que los investigadores aprendan y dominen esta importante habilidad hemos preparado códigos de Stata y R específicamente diseñados para facilitar cálculos de tamaño de muestra, los mismos que estarán disponibles de manera pública y gratuita. Dicho esto, es muy importante resaltar la importancia de realizar el cálculo del tamaño de muestra siempre a priori y no post hoc, los cuales no tienen ningún valor y pueden inducir a violaciones de tipo ética. Asimismo, se recomienda no realizar un solo cálculo de tamaño de muestra sino uno para cada una de las hipótesis de interés del estudio y a escoger el escenario más conservador, sacrificando precisión si es necesario. Finalmente, se recomienda tener especial cuidado en penalizar la muestra ya que el cálculo de tamaño de muestra sirve para calcular el tamaño de muestra con el que deberíamos terminar el estudio, esto es luego de los rechazos y pérdidas en el seguimiento. De ahí que se recomienda siempre ser conservador con sus cálculos e interpretarlos con mucha cautela.

Conflictos de interés: Los autores niegan conflictos de interés.

Financiamiento: Autofinanciado.

REFERENCIAS BIBLIOGRÁFICAS

1. Sormani MP. The most frequently asked question to a statistician: The sample size. *Mult Scler.* 2017;23(5):644-6. doi:10.1177/1352458517695471
2. Burmeister E, Aitken LM. Sample size: how many is enough? *Aust Crit Care.* 2012;25(4):271-4. doi:10.1016/j.aucc.2012.07.002
3. Rao UK. Concepts in sample size determination. *Indian J Dent Res.* 2012;23(5):660-4.

- doi:10.4103/0970-9290.107385
4. Faber J, Fonseca LM. How sample size influences research outcomes. *Dental Press J Orthod*. 2014;19(4):27-9. doi:10.1590/2176-9451.19.4.027-029.ebo
 5. Chiuzan C, West EA, Duong J, Cheung KY, Einstein AJ. Sample size considerations for clinical research studies in nuclear cardiology. *J Nucl Cardiol*. 2015;22(6):1300-13. doi:10.1007/s12350-015-0256-7
 6. Cesana BM, Antonelli P. Sample size calculations in clinical research should also be based on ethical principles. *Trials*. 2016;17(1):149. doi:10.1186/s13063-016-1277-5
 7. Cook DA, Hatala R. Got power? A systematic review of sample size adequacy in health professions education research. *Adv Health Sci Educ Theory Pract*. 2015;20(1):73-83. doi:10.1007/s10459-014-9509-5
 8. Clark T, Berger U, Mansmann U. Sample size determinations in original research protocols for randomised clinical trials submitted to UK research ethics committees: review. *BMJ*. 2013;346:f1135. doi:10.1136/bmj.f1135
 9. Loria A. Tamano de muestra de protocolos clinicos. *Gac Med Mex*. 2018;154(3):409-12. doi:10.24875/GMM.18003627
 10. Kumar R. Sample size calculation. *Indian J Ophthalmol*. 2012;60(6):582; discussion doi:10.4103/0301-4738.103809
 11. Maggard MA, O'Connell JB, Liu JH, Etzioni DA, Ko CY. Sample size calculations in surgery: are they done correctly? *Surgery*. 2003;134(2):275-9. doi:10.1067/msy.2003.235
 12. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA*. 2007;298(9):1010-22. doi:10.1001/jama.298.9.1010
 13. Freeman JV, Collier S, Staniforth D, Smith KJ. Innovations in curriculum design: a multi-disciplinary approach to teaching statistics to undergraduate medical students. *BMC Med Educ*. 2008;8:28. doi:10.1186/1472-6920-8-28
 14. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H, Jr. Publication bias and clinical trials. *Control Clin Trials*. 1987;8(4):343-53. doi:10.1016/0197-2456(87)90155-3
 15. Morse JM. Analytic Strategies and Sample Size. *Qual Health Res*. 2015;25(10):1317-8. doi:10.1177/1049732315602867
 16. Stang A, Poole C, Kuss O. The ongoing tyranny of statistical significance testing in biomedical research. *Eur J Epidemiol*. 2010;25(4):225-30. doi:10.1007/s10654-010-9440-x
 17. Lu Y, Belitskaya-Levy I. The debate about p-values. *Shanghai Arch Psychiatry*. 2015;27(6):381-5. doi:10.11919/j.issn.1002-0829.216027
 18. Gagnier JJ, Morgenstern H. Misconceptions, Misuses, and Misinterpretations of P Values and Significance Testing. *J Bone Joint Surg Am*. 2017;99(18):1598-603. doi:10.2106/JBJS.16.01314
 19. Mascha EJ, Vetter TR. Significance, Errors, Power, and Sample Size: The Blocking and Tackling of Statistics. *Anesth Analg*. 2018;126(2):691-8. doi:10.1213/ANE.0000000000002741
 20. Weber EJ, Hoo ZH. Why sample size estimates? *Emerg Med J*. 2018;35(12):755-6. doi:10.1136/emered-2018-207763
 21. Malone HE, Nicholl H, Coyne I. Fundamentals of estimating sample size. *Nurse Res*. 2016;23(5):21-5. doi:10.7748/nr.23.5.21.s5
 22. Rothman KJ, Greenland S. Planning Study Size Based on Precision Rather Than Power. *Epidemiology*. 2018;29(5):599-603. doi:10.1097/EDE.0000000000000876
 23. Billings M, Lopez Mitnik G, Dye BA. Sample size for clinical trials. *Oral Dis*. 2017;23(8):1013-8. doi:10.1111/odi.12611
 24. Gogtay NJ. Principles of sample size calculation. *Indian J Ophthalmol*. 2010;58(6):517-8. doi:10.4103/0301-4738.71692
 25. Martinez-Mesa J, Gonzalez-Chica DA, Bastos JL, Bonamigo RR, Duquia RP. Sample size: how many participants do I need in my research? *An Bras Dermatol*. 2014;89(4):609-15. doi:10.1590/abd1806-4841.20143705
 26. Johnston KM, Lakzadeh P, Donato BMK, Szabo SM. Methods of sample size calculation in descriptive retrospective burden of illness studies. *BMC Med Res Methodol*. 2019;19(1):9. doi:10.1186/s12874-018-0657-9
 27. Black CS. Understanding primary outcomes, sample size calculation, and the perils of online calculators. *Paediatr Anaesth*. 2018;28(3):307-8. doi:10.1111/pan.13330
 28. Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: sample size and power calculations-why, when and how? *Eur J Cardiothorac Surg*. 2018;54(1):4-9. doi:10.1093/ejcts/ezy169.

Correspondencia

Antonio M Quispe MD, MSc, CPH, PhD.
 Dirección: General Borgoño 355, Miraflores.
 Teléfono: 962169519
 Correo: drantonioquispe@gmail.com

Revisión de pares

Recibido: 15/01/2020
 Aceptado: 20/01/2020