

Primera versión: 30 de mayo de 2006
 Versión definitiva: 9 de junio de 2006
 Aceptado: 15 de junio de 2006

Ángel
 Guerra-Márquez,¹
 María Guadalupe
 Carrillo-Montes²

Conceptos básicos y premisas para calcular el tamaño de muestra

¹ Hematólogo. Jefatura de Departamento Clínico.
² Patóloga Clínica.

Banco Central de Sangre
 Centro Médico Nacional
 "La Raza", Instituto Mexicano del Seguro Social.

Comunicación con:
 Ángel Guerra-Márquez.
 Tel.: y fax: 5583 5047.
 Dirección electrónica:
angel_guerr@terra.com.mx

RESUMEN

El tamaño de muestra es un aspecto crucial en un proyecto de investigación, ya que permite valorar el efecto real de una intervención u observación. Su cálculo depende de una serie de factores como: pregunta de investigación, diseño del estudio, hipótesis, variables, rango de error, tamaño del efecto y significancia estadística. Cada uno de ellos debe ser conocido y establecido por el investigador. Los paquetes estadísticos en sistemas informáticos facilitan la tarea del cálculo de tamaño de muestra, pero invariabilmente requieren de la introducción de información que el investigador puede obtener de estudios publicados, de su propia experiencia o de estudios piloto.

SUMMARY

Sample size is a crucial aspect in a research project, since it allows the assessment of the real effect of an intervention or observation. Its calculation depends on a series of factors, such as the research question, study design, hypothesis, variables, rank of error, effect size and statistical significance. Each of them must be known and established by the researcher. Statistical packages in computer systems make the calculation of the sample size easier, but they invariably require the introduction of information, which the researcher can get from published studies, his own experience or pilot studies.

Introducción

La determinación del tamaño de muestra es una parte fundamental en los diseños de investigación clínica, ya que de ella depende, en buena medida, la solidez de los resultados del estudio. Para realizar un cálculo adecuado del tamaño de muestra es necesario que el investigador conozca algunos conceptos básicos de metodología de la investigación, ya que si bien en la actualidad se cuenta con paquetes estadísticos que facilitan la tarea, éstos requieren de la introducción de datos que el investigador debe conocer.

El presente documento pretende ser una guía general para realizar el cálculo de tamaño de muestra acorde con los objetivos del proyecto de investigación.

Una vez que el investigador decide cuál es la pregunta de investigación y con ello a quién

y qué va a estudiar, debe entonces calcular el número de individuos necesarios para su estudio, un número que resulte suficiente para responder a la pregunta planteada, acorde al tipo de diseño de la investigación y que evite costos, dificultades y una estimación inadecuada del fenómeno.

Calcular el tamaño de muestra está indicado siempre que se quiere probar una hipótesis causal, la relación entre un factor de riesgo y una enfermedad o comparar si un tratamiento es mejor que otro. En aquellos estudios donde sólo se desea identificar la frecuencia de ciertos fenómenos y se conoce el total de la población a estudiar no es necesario calcular un tamaño de muestra. Sin embargo, si se conoce el total de la población a estudiar pero el tiempo de reclutamiento es largo o el número de sujetos muy elevado sí es necesario hacerlo. La mayoría de los autores concuerdan en que las

Palabras clave:

- ✓ estadística
- ✓ tamaño de muestra

Key words:

- ✓ statistics
- ✓ sample size

series de casos no necesitan un tamaño de muestra.

La representatividad de una muestra depende por lo general del tamaño y la selección de la misma.^{1,2} Para realizar el cálculo adecuado del número de sujetos necesarios para demostrar un efecto real se necesitan establecer algunos principios: el investigador debe plantear adecuadamente la pregunta de investigación, establecer el diseño del estudio, la hipótesis, identificar las variables predictora y de desenlace, establecer el rango del error tipo I y II, el tamaño de efecto y el nivel de significancia estadística.^{3,4}

Pregunta de investigación: se define como la incertidumbre acerca de algún fenómeno relativo a una población. Se considera parte fundamental del proceso de investigación. Debe expresar una relación entre variables, formularse en forma de pregunta, posibilitar la prueba empírica de las variables y expresarse en una dimensión temporal y espacial.³

Diseño del estudio: depende de la pregunta de investigación y del objetivo propuesto. Puede tratarse de estudios observacionales donde el investigador sólo observa y registra los datos para su análisis o estudios experimentales donde el investigador realiza una intervención, generalmente controlada, y registra los efectos de la maniobra para su análisis. En la investigación clínica se utilizan principalmente cinco tipos de diseño: serie de casos, transversal, casos y controles, cohortes (observacionales) y el ensayo clínico controlado (experimentales).

Hipótesis de estudio: es una afirmación que establece una respuesta a la pregunta de investigación, así como las bases para las pruebas de significación estadística. Se recomienda que sea simple y específica. Debe tener magnitud (tamaño del efecto esperado) y dirección (que identifique el sentido de la asociación entre las variables).^{3,5}

Hipótesis nula: establece que no existe asociación entre las variables predictora y de desenlace en la población donde se extrajo la muestra. Se considera que la hipótesis nula es la base para establecer la significación estadística partiendo del hecho de que cualquier asociación observada se deba al azar.^{3,5}

Variables: son las características que el investigador mide en un estudio. De acuerdo con

su relación causal se dividen en: dependientes (desenlace) e independientes (predictora).

Una variable dependiente o de desenlace representa el resultado de los valores de una consecuencia o antecedente. Por lo tanto, es el resultado de interés de cualquier tipo de estudio. La variable independiente o predictora es una característica que condiciona o determina la presencia de un fenómeno (generalmente un factor de riesgo) y define las condiciones bajo las cuales se realizará la medición de la variable dependiente.⁵

Error tipo I (falso positivo): en ocasiones sólo por el azar una muestra puede no ser representativa de la población y se genera el error tipo I que significa establecer un resultado como cierto cuando en realidad es falso en la población. Para minimizar este error se acepta de manera general establecer un valor de α de 0.05 %.^{3,5}

Error tipo II (falso negativo): el investigador concluye que un resultado es falso en la población cuando en realidad es verdadero. Se acepta de manera general establecer el valor de β en 20 %; con este valor puede estimarse que el tamaño de muestra tiene un poder $(1-\beta)$ de 80 %.^{3,5}

Tamaño del efecto: es la probabilidad de que en el estudio se pueda detectar una asociación entre la variable predictora y la de desenlace. Generalmente el investigador no conoce este dato y debe elegir el tamaño de la asociación que a él le gustaría ser capaz de detectar en la muestra.¹

Nivel de significancia estadística: se expresa como valor de p ($p =$). Es la probabilidad de que los resultados obtenidos en un estudio sean debidos al azar. Generalmente se utiliza un valor de p menor a 0.05.^{3,5}

En la actualidad existen paquetes estadísticos informáticos que permiten calcular el tamaño de muestra para un estudio determinado, tomando en consideración algunos datos como la frecuencia del fenómeno a analizar, el número de la población objeto del estudio y el margen de error aceptable. En ocasiones es posible que el investigador desconozca alguno de estos datos, por lo que puede apoyarse en los datos publicados en las referencias bibliográficas o en datos obtenidos de su práctica diaria. De la confiabilidad de los valores estimados depende la solidez de los resultados del

estudio. Si el investigador encuentra un estudio similar al que planea realizar, los resultados obtenidos pueden utilizarse para llevar a cabo un cálculo adecuado. Cuando no se tiene ningún dato que permita realizar un estimado del tamaño de muestra se recomienda utilizar la prevalencia de 50 % ya que con este valor se estima el tamaño de muestra más grande o en su defecto es recomendable realizar un estudio piloto.²

Consideraciones generales

Para estimar un tamaño adecuado de muestra se recomienda tener en cuenta lo siguiente:

1. Conocer la prevalencia de la enfermedad: a mayor frecuencia del fenómeno menor tamaño de muestra.
2. Conocer la variabilidad de los datos: a mayor variabilidad mayor tamaño de muestra.
3. El tipo de estudio: los estudios de cohorte requieren mayor número de muestra que los casos y controles.
4. Cuando la asociación entre variables es débil o el efecto esperado es pequeño, el tamaño de muestra aumenta.
5. Una muestra muy grande puede producir un error tipo I.

6. Una muestra muy pequeña puede producir un error tipo II.
7. Cuando se dice que se encontró una asociación estadísticamente significativa, quiere decir que tenemos una probabilidad menor a 5 % de cometer un error tipo I.
8. Cuando se establece una hipótesis bidireccional, el tamaño de muestra aumenta.

Ángel
Guerra-Márquez et al.
Cálculo de tamaño de muestra

Referencias

1. Reynaga Obregón J, Gómez Gómez M. Estadística básica en ciencias de la salud. Mex. DF; 2000.
2. Mejía-Arangué J, Fajardo-Gutiérrez A, Gómez-Delgado A, Cuevas-Urióstegui M, Hernández-Hernández D, Garduño-Espinoza J y cols. El Tamaño de la muestra: un enfoque práctico en la investigación clínica pediátrica. Bol Med Hosp Infant Mex 1995;52(6):381-391.
3. Browner W, Newman T, Cummings S, Hulley S. Preparación para estimar el tamaño de la muestra: hipótesis y principios subyacentes. En: Hulley B. Diseño de la investigación clínica. Segunda edición; 1997. p. 143-151.
4. Gómez-Gómez M, Danglot-Banck C, Velásquez-Jones L. Bases para la revisión crítica de artículos médicos. Rev Mex Pediatr 2001;68(4):152-59.
5. Fajardo Rodriguez A, González Ávila G. Cómo preparar un resumen estructurado. Nutr Clin 1998;1(1):28-39. 

