

Obstetricia

Artículo original

El tamaño muestral en la investigación clínica: ¿cuántos pacientes es necesario incluir en un estudio?

DRES. ROBERTO LEDE, GUILLERMO CARROLI Y LIC. MARTA ALARCÓN¹

El presente artículo está dirigido al médico asistencial, interesado en desarrollar investigaciones clínicas y persigue los siguientes objetivos:

- a) justificar por qué las investigaciones deben incluir una cantidad mínima necesaria de individuos y
- b) proveer de los elementos básicos que permitan el cálculo de esa cantidad mínima necesaria de individuos.

¿Qué es el tamaño muestral?

Se denomina tamaño muestral (TM) a la cantidad mínima necesaria de sujetos pertenecientes a la porción o muestra de la población total que debe ser incluida en un estudio, para permitir que sus conclusiones estén confiablemente protegidas de errores producidos por el azar.

Es imprescindible tener en cuenta que en medicina no existe la certeza absoluta y que siempre está presente la probabilidad de que los resultados obtenidos durante la investigación no expresen la realidad, ya sea por dar por cierto hechos falsos o por falso hechos ciertos.

¿Para qué sirve el tamaño muestral?

El objetivo de calcular el TM adecuado es permitir que las conclusiones que se obtengan de una investigación se deban exclusivamente al factor en estudio y no al azar.

En los estudios analíticos y experimentales, alcanzar el TM exigido permite que las conclusiones acerca de los efectos que alguna exposición o intervención produce sobre un resultado estén confiablemente protegidas de los errores introducidos por el azar. En los estudios descriptivos, un TM adecuado permite aproximarse confiablemente a los valores reales de los parámetros estudiados.

El TM debe ser calculado en fases tempranas del proyecto, pues influye decisivamente sobre la factibilidad y el diseño del estudio.

La factibilidad de la realización de un estudio es un punto importante a ser evaluado, ya que si para alcanzar conclusiones confiables se requiere una cantidad de sujetos que excede las posibilidades de los investigadores, éstos

¹ Centro Rosarino de Estudios Perinatales, Rosario, República Argentina.

deberán replantear su idea o abandonarla. Es preferible evitar los esfuerzos y gastos que implican una investigación si no se estiman razonables las posibilidades de llegar a realizarla.

Otro punto de importancia a considerar para la factibilidad es evaluar el índice de reclutamiento de individuos aptos para la investigación, ya que generalmente se peca por exceso de optimismo, debiéndose considerar que una cierta proporción de los pacientes a los que se puede acceder no serán elegibles por diversas razones, así como también que algunos de ellos no aceptarán ingresar al estudio.

Ocurre frecuentemente que, al ser parte de un estudio, las prácticas asistenciales habituales involucradas en el mismo se efectúan con mayor rigurosidad; por este motivo, puede disminuir la ocurrencia esperada de eventos adversos, lo que en definitiva requerirá un TM mayor.

Con respecto al diseño del estudio, existen ciertas estrategias aplicables cuando los requerimientos del TM exceden las posibilidades de la fuente de pacientes disponible. Recurrir a la participación de varios centros de reclutamiento es una de ellas, casi imprescindible en estudios que requieren un gran número de pacientes, hecho muy frecuente en los diseños experimentales.

Si habiendo practicado la estrategia anteriormente comentada, el TM necesario permanece aún inalcanzable, puede ampliarse el espectro de los individuos elegibles mediante la reducción de las exigencias de inclusión. Esto redundará sobre la representatividad de las conclusiones, pero en los estudios analíticos y experimentales la magnitud del efecto de la exposición o intervención tendrá menor impacto, ya que se incluirán individuos con características menos apropiadas para perjudicarse o beneficiarse con las mismas.

Como sugerencia general se estima que, si a pesar de todas las estrategias aplicadas, el período de reclutamiento estimado es mayor de 3 años, debe considerarse seriamente la realización o continuación del estudio. Períodos prolongados de reclutamiento demandan presupuestos elevados y atentan contra el entusiasmo de los participantes, agregándose la probabilidad de la aparición de modificaciones en las prácticas médicas, incluso sugeridas inadvertidamente por el propio protocolo en prueba. Se debe destacar que los estudios de seguimiento longitudinal, por sus propias exigencias, frecuentemente exceden este período recomendado, lo que los convierte en verdaderos desafíos gerenciales.

Es preciso tener en cuenta que toda conclusión no avalada por la metodología científica adecuada no será *invariablemente falsa*, pero sí *insalvablemente incierta*.

¿Qué relación existe entre el TM y las significaciones biológica y estadística?

El investigador clínico se interesa por comprobar si, entre las muestras estudiadas, existen diferencias que entrañen significado para la práctica clínica. Es decir, debe interesarse en revelar diferencias que sean de utilidad para la mejor atención de los pacientes.

Para demostrar que las diferencias biológicas observadas no se deben exclusivamente al azar sino a un efecto real, el investigador se vale de métodos estadísticos; por este motivo se calcula el TM. En el contexto del análisis, es muy importante que los conceptos de *estadísticamente*

significativo y *biológicamente significativo* sean considerados cada uno en su real dimensión y en forma conjunta.

En ciertas ocasiones, una diferencia puede resultar estadísticamente significativa pero carente de valor biológico; o por el contrario, ser biológicamente significativa pero no serlo estadísticamente.

Esta situación puede ejemplificarse de la siguiente manera: un investigador evalúa el peso corporal de dos grupos de 120 recién nacidos cada uno y los expresa como promedio \pm DS (desvío estándar):

Grupo I : 2.700 \pm 590 gramos
Grupo II : 2.800 \pm 580 gramos

En estas muestras, la diferencia entre 2.700 y 2.800 gramos no resulta estadísticamente significativa, pero, biológicamente, una diferencia de 100 g en el peso de los recién nacidos puede tener relevancia.

Por el contrario, si cada grupo incluyera 150 recién nacidos y sus pesos medios fueran 2.950 \pm 58 gramos y 2.931 \pm 60 gramos respectivamente, se puede comprobar que esta diferencia resulta estadísticamente significativa; sin embargo, una diferencia de 19 g en los pesos medios no reclama atención biológica.

¿Qué conclusión se obtiene a partir de las situaciones expuestas? En el marco de una prueba de hipótesis, establecidos *a priori* los márgenes de error y las diferencias que se deseen observar, es posible alcanzar la significación estadística a expensas de aumentar el tamaño muestral; pero estas diferencias *estadísticamente significativas* logradas, no siempre son *biológicamente representativas*. Esto lleva a destacar que el investigador debe recordar en todo momento cuál es la diferencia biológica que, con la prueba de hipótesis, pretende averiguar si es *estadísticamente significativa*, para lo cual debe calcular el tamaño muestral requerido.

El término *significativo*, aplicado habitualmente para calificar una diferencia, no resulta una elección afortunada, ya que promueve la idea de que dicha diferencia resulta importante en el marco del tema de estudio. En realidad, esta calificación sólo expresa que la probabilidad de que el resultado obtenido haya sido por azar es menor que la probabilidad de error que el investigador se permitía cometer. Mucho ganaría la comprensión de los informes científicos si se abandonara el uso del término *significativo*.

¿Cuál es el tamaño muestral apropiado?

El TM apropiado es aquél que protege las conclusiones de los errores por azar, para los niveles de error admitidos por los investigadores.

El TM apropiado se basa en el concepto de que repitiendo un experimento u observación una cierta cantidad de veces, da ocasión a mostrar cuál de los eventos que ocurren a consecuencia de los mismos lo hace con mayor frecuencia.

Esa reiteración de observaciones depende de la frecuencia con que se estima que ocurrirá el evento de interés, con cuánta precisión se desea establecer las mismas y con qué probabilidad máxima de concluir erróneamente se admite trabajar.

Por ende, el TM depende principalmente de las pretensiones del investigador, ya que es éste quien establece las pautas que lo definirán.

Cuanto más pequeñas sean las diferencias que desee comprobar, cuanto más preciso pretenda ser en las apreciaciones, cuanto mayor seguridad quiera para sus conclusiones, mayor será el TM requerido.

El TM que se determina mediante los cálculos estadísticos es el mínimo que se necesita alcanzar; se prefiere siempre excederlo, lo que otorgará mayor confiabilidad a las conclusiones

El trabajo científico se efectúa con el fin de conocer la realidad. Evitar que ocurran errores en la interpretación de la misma resulta prácticamente imposible, aunque es factible acotarlos al respetar las normas de diseño y análisis establecidos.

¿Qué son y cómo se fijan los niveles de error?

Admitir que existe una diferencia entre las muestras comparadas, cuando en realidad ésta se ha manifestado solo por azar, es cometer el error llamado tipo I. Dicho mediante un simple ejemplo, es afirmar que un tratamiento es mejor que otro cuando en realidad no lo es.

Concluir que no hay diferencias entre las muestras estudiadas, cuando en realidad éstas existen, es cometer el error denominado tipo II. Equivale a aseverar que dos tratamientos tienen iguales resultados cuando en realidad uno es más efectivo que otro.

Se denomina α al máximo error tipo I (medido en términos de probabilidad) que el investigador está dispuesto a tolerar y se denomina β a la probabilidad máxima tolerable de cometer el error tipo II.

El investigador puede otorgarles a α y β el nivel que considere oportuno, pero con un criterio de orden se ha convenido internacionalmente que para los estudios biomédicos el nivel máximo tolerable para α es el 5% (raramente 10%) y para β 20%.

Así, por ejemplo, y adoptando las pautas recién expresadas, un estudio que presente un TM requerido, expresa que posee una probabilidad menor del 5% de admitir como real una diferencia inexistente y que, además, la probabilidad de negar una diferencia realmente existente es menor al 20%.

¿Cómo se calcula el tamaño muestral?

Tanto en los diseños analíticos como en los experimentales, los elementos intervinientes en el cálculo del TM son sustancialmente los mismos: estimación de la frecuencia de ocurrencia del evento de interés, selección de la diferencia mínima a detectar y fijación de los niveles de error α y β que está dispuesto a aceptar.

Los diseños descriptivos, al no poseer hipótesis a probar, requieren menos definiciones que los recién mencionados. Estas son: la estimación de la probabilidad de ocurrencia del evento o de una medida de posición y dispersión de la variable a describir, la precisión con la que se efectuará y la fijación del nivel de α .

El médico asistencial con inclinación a la actividad investigativa no siempre cuenta con la formación académica necesaria para la interpretación de los conceptos estadísticos aplicados al cálculo del TM, respondiendo en general con una actitud de rechazo hacia estas disciplinas. En un intento de facilitarle su acceso a la elaboración de la respuesta a la inquietante pregunta sobre el TM requerido para su estudio en ciernes, se exponen los procedimientos algebraicos necesarios para calcularlo, los cuales pueden

ejecutarse mediante los métodos tradicionales o con una simple calculadora de mano.

Independientemente de que el interesado sea capaz de efectuar los cálculos necesarios, es recomendable que, antes de iniciar la marcha del estudio, consulte con alguien especializado en diseño de investigación. Todo el tiempo y precauciones insumidos en la etapa preparatoria de los estudios es tiempo y desilusiones futuras que se ahorran.

CÁLCULO EN LOS ESTUDIOS ANALÍTICOS

Cohorte

Debe disponerse de la siguiente información:

p_0 = la proporción de participantes del grupo no expuesto, en la que se espera observar el efecto de interés durante el período de evaluación de la cohorte y del cual estaban libres al incorporarse al estudio.

p_1 = la proporción de participantes del grupo expuesto, en la que se espera observar el efecto de interés durante el período de evaluación de la cohorte y del cual estaban libres al incorporarse al estudio.

La diferencia entre estas proporciones es la que los investigadores estiman como una diferencia de importancia clínica y es fijada por ellos de acuerdo a la evidencia disponible sobre el tema. Ejemplo: se estima que el grupo expuesto tendrá el doble de probabilidades de padecer el problema que el grupo no expuesto. Si en este último el efecto ocurre en un 10% (p_0) de los individuos, se consideraría de importancia clínica que al exponerse aumentara al 20% (p_1). Si se utiliza el concepto de veces en que se incrementa o reduce el riesgo de adquirir el efecto (Riesgo Relativo, RR), p_1 se calcula como se muestra en la fórmula 2.

α = nivel de probabilidad admitido de cometer el error tipo I.

β = nivel de probabilidad admitido de cometer el error tipo II.

Los valores probabilísticos de α y β ($Z\alpha$ y $Z\beta$), para los niveles fijados, se buscan en las tablas preparadas al efecto. La tabla de valores Z incluye los más requeridos. Es recomendable aplicar siempre los valores para pruebas bilaterales, ya que la dirección del efecto que suponemos puede no cumplirse.

Usando toda esta información calculamos el TM de la siguiente manera:

(p , p_1 y p_0 deben utilizarse como expresiones de probabilidad, por lo que se anotarán como valores entre 0 y 1. Ejemplo: una p de 10% se inscribirá en la fórmula como 0,10; una de 5% como 0,05, etc.).

$$n = [2 (Z\alpha + Z\beta)^2 \times p (1-p)] / (p_0 - p_1)^2 \quad (1)$$

donde
$$p_1 = (p_0 \times RR) / [1 + p_0 \times (RR - 1)] \quad (2)$$

$$p = (p_0 + p_1) / 2 \quad (3)$$

y “n” significa la cantidad mínima de individuos que deben ser incluidos en cada uno de los grupos (grupo con el factor y grupo sin el factor de riesgo en consideración).

Ejemplo

Interrogantes: ¿Cuántos individuos es necesario involucrar en un estudio para investigar si fumar durante el embarazo (>2 cigarrillos/día) es factor de riesgo para el nacimiento de recién nacidos de bajo peso?

Desarrollo: 1) Establecer la frecuencia de bajo peso entre las no fumadoras. Según estudios previos realizados en la población a estudiar es del 7% ($p_0=0,07$).

2) Establecer la diferencia que se aprecie como importante detectar. Para este ejemplo se estima que la exposición al cigarrillo (>2 cigarrillos/día) incrementa en 1,5 veces la frecuencia de bajo peso al nacer (RR = 1,5).

3) Definir el nivel de α . Habitualmente 0,05. Dado que no se espera que el fumar tenga un efecto protector sobre el peso al nacer, el sentido de la respuesta será único, por lo que se deberá observar el valor de $Z\alpha$ en una tabla unilateral ($Z_{0,05}=1,64$).

4) Se define la potencia del estudio. Ejemplo, $\beta = 0,10$ ($Z_{0,10} = 1,28$).

5) Calcular la proporción de recién nacidos de bajo peso que se espera encontrar en la cohorte de fumadoras, p_1

$$p_1 = (0,07 \times 1,5) / [1 + 0,07 \times (1,5-1)] = 0,101$$

6) Calcular p:

$$p = (p_0 + p_1) / 2 = 0,0855$$

7) Aplicando lo anterior, n será:

$$n = [2 \times (1,64 + 1,28)^2 \times 0,0855 \times 0,9145] / 0,0009 = 1,481$$

Conclusión: si los hechos son como los investigadores suponen, se requiere incluir a 1.481 embarazadas fumadoras y otras tantas no fumadoras para establecer, con una confianza del 95%, que existe un incremento de por lo menos 1,5 veces en el riesgo de producir un recién nacido de bajo peso al fumar más de 2 cigarrillos/día.

Caso control

Con cantidad igual de casos y controles. La fórmula es la misma que la utilizada en los diseños de cohorte, puesto que en ambos casos existe una hipótesis que probar.

El cálculo dependerá de la siguiente información:

p_0 = la proporción en la que se estima se halla el factor de riesgo en estudio en la población sin el daño (*controles*).

p_1 = la proporción en la que se estima se halla el factor de riesgo en estudio en la población con el daño (*casos*). Se calcula con la fórmula (2), en la que se reemplaza RR por RO.

RO = es la razón de ocurrencia *odds-ratio* mínimo que se estima de interés clínico detectar. Se requiere para calcular p_1 .

p = se calcula aplicando la fórmula (3).

Fijación de α y β con iguales valores y propósitos que en el caso anterior. Con todos estos datos a disposición, el TM surge al aplicar la fórmula (1) y significa la cantidad de pares de casos y controles que deben ser incluidos en el estudio.

Ejemplo

Interrogante: para demostrar que la colonización placentaria por micoplasmas se asocia a un riesgo 3 veces mayor de muerte fetal, ¿qué TM se requiere?

Desarrollo: 1) Determinación de la frecuencia en que se espera hallar colonización placentaria por micoplasmas entre los controles. Por estudios previos es posible admitir que la misma está presente en el 14% de los controles (recién nacidos vivos) ($p_0 = 0,14$).

2) Determinación de la RO mínima que se estima de interés demostrar: los investigadores deciden que resulta clínicamente importante si la presencia de colonización placentaria entre los nacidos muertos (casos) es 3 veces mayor (fijación arbitraria por el investigador) que entre los controles (RO = 3).

3) Calcular la hipotética frecuencia a hallarse entre los casos (p_1)

$$p_1 = (0,14 \times 3) / [1 + 0,14 \times (3 - 1)] = 0,328$$

4) Definir el nivel de α . Dado que no se espera (no hay fundamentos que así lo sugieran) que la colonización placentaria por micoplasmas tenga un efecto protector sobre la vida fetal, el sentido de la respuesta será único, por lo que se deberá observar el valor de $Z_{0,05}$ en una tabla unilateral, 1,64.

5) Se elige la potencia del estudio. En este ejemplo, $\beta = 0,10$. ($Z_{0,10} = 1,28$)

6) Se calcula el valor de p .

$$P = (p_1 + p_0) / 2 = 0,234$$

7) Se aplica la fórmula (1):

$$n = [2 \times (1,64 + 1,28)^2 \times 0,234 \times 0,766] / 0,035 = 86$$

Conclusión: si la colonización placentaria por micoplasmas resulta, efectivamente, un mínimo de 3 veces mayor entre los casos que los controles, se requieren 86 pares de casos-control para afirmarlo con un 95% de confianza.

Con cantidad desigual de casos y controles. Dado que existe un desbalance en los grupos, se debe introducir un factor de corrección. Este se identifica como "r" y representa la cantidad de individuos-control por cada individuo-

caso. Ej.: 3:1, entonces r será igual a 3. Establecido este factor, se calcula la ecuación de ajuste, de la siguiente manera:

$$(r + 1) / (2 \times r)$$

Al TM calculado mediante el procedimiento expuesto en el ítem precedente, se lo multiplica por el factor de corrección calculado.

Cálculo en los estudios experimentales

Investigación clínica aleatorizada. En estos estudios se compara la proporción de participantes que presentarían el fenómeno de interés habiendo recibido una intervención, versus la proporción de participantes que presentarían el fenómeno habiendo recibido una intervención alternativa. Esgrime un criterio similar al estudio de cohorte, pues analiza la ocurrencia de un efecto ante la presencia de una intervención, es decir, de una exposición controlada por el investigador. La diferencia reside en que en las cohortes se observa lo que sucede espontáneamente, mientras que en las investigaciones clínicas aleatorizadas el investigador produce una intervención. El cálculo de TM depende de la siguiente información:

p_0 = proporción de los participantes del grupo control en los cuales se espera que se desarrolle el punto final de interés.

p_1 = proporción de los participantes del grupo experimental en los cuales se espera que se desarrolle el punto final de interés.

p = se calcula mediante la fórmula (3).

Fijación de α y β , con los valores y propósitos ya explicitados.

Determinados estos elementos, se procede a calcular el TM, aplicando la fórmula (1). Informa la cantidad de individuos que deben ser incluidos en cada uno de los grupos.

Ejemplo

Interrogante: una investigación que pretende averiguar si la administración de bajas dosis de aspirina reduce la frecuencia de retardo del crecimiento intrauterino (RCI) ¿cuántos participantes debe incluir?

Desarrollo: 1) Establecer la frecuencia de RCI procedente de embarazadas que no son sometidas a la administración de aspirinas en bajas dosis. Por datos previos de la población que se estudiará se conoce que es del 8% ($p_0=0,08$).

2) Establecer la reducción en dicha frecuencia que se estima de interés alcanzar. Ejemplo: se decide que una reducción mínima del 30% en la frecuencia de RCI es de importancia clínica (criterio propio de los investigadores). $p_1 = 70\%$ de $p_0 = 0,056$ (es decir, p_0 menos el 30%).

3) Definir el nivel de α . Habitualmente 0,05. Aun cuando no se espera que la administración de aspirina tenga un efecto adverso sobre la frecuencia de RCI, es preferible utilizar siempre valores de $Z\alpha$ para bilateralidad ($Z_{0,05}=1,96$).

4) Se define la potencia del estudio. Ejemplo, $\beta = 0,10$ ($Z_{0,10} = 1,28$).

5) Se estima el valor de p

$$p = (p_0 + p_1) / 2 = 0,068$$

6) Se calcula el TM según la fórmula (1):

$$n = (2 \times 10,4976 \times 0,068 \times 0,932) / 0,00057 = 2.334$$

Conclusión: se requiere incluir un mínimo de 2.334 embarazadas a las cuales se le administren bajas dosis de aspirina y 2.334 a las que no se le administren para detectar, con una confianza del 95%, si dicha administración reduce la frecuencia de RCI en un mínimo del 30%.

Cálculo en los estudios descriptivos

Los estudios descriptivos se efectúan con la intención de describir lo que sucede en la población, sin confrontar alternativas.

El objetivo de estos estudios es determinar el valor de la variable de interés. Si la variable se expresa en forma cualitativa, interesará determinar la frecuencia con que se presenta en la población. Si se la expresa como una variable continua, se utilizará una medida de posición y dispersión (media, mediana, percentilos, etc.).

Los únicos elementos que el investigador debe fijar para el cálculo del TM son la precisión con que se desea establecer en la estimación a realizar y el nivel de confianza con el cual se admitirá el resultado. Al carecer de hipótesis, no cabe fijar el nivel de error β .

Cálculo del TM al analizar variables dicotómicas. Son estudios que se efectúan para estimar la frecuencia aproximada con la que ocurre un determinado evento en la población.

Se necesita fijar el nivel de confianza en los resultados, es decir, determinar los límites dentro de los cuales estará incluida la verdadera frecuencia con la que se presenta el punto final en la población estudiada. Habitualmente se establece un nivel de confianza del 95%.

La expresión α denota la probabilidad de cometer el error de aceptar un intervalo de confianza que en realidad no contenga la frecuencia real. Cuanto más exigente sea el nivel de confianza, mayor será el TM requerido. Dado que el intervalo de confianza se extiende hacia ambos lados del valor estimado de la variable, se debe utilizar el valor de $Z\alpha$ para pruebas bilaterales que se corresponda con el nivel α elegido.

También debe fijarse el nivel de precisión; esto es, que la diferencia entre el valor estimado y el valor real de la medida de estudio no exceda ciertos límites. Esta precisión se simboliza con d , y habitualmente se la fija en el 5 ó 10% ($d = 0,05$ ó $0,10$). Serán $0,005$ ó $0,010$, si la expresión de proporción es por mil, etc.

Estimar la frecuencia (p) en la que se halla la variable de interés en la población a estudiar. Si no se tiene ningún dato ni es posible practicar un estudio exploratorio, se adopta 50% ($p = 0,5$).

Usando toda esta información, se determina el TM mediante la aplicación de la siguiente fórmula:

$$n = p \times (1 - p) \times (Z_{\alpha} / d)^2 \quad (4)$$

Ejemplo

Interrogante: si se estima que el 6% de la población de embarazadas desarrolla preeclampsia ¿cuántas embarazadas se deben supervisar para corroborar que ésta es la verdadera frecuencia en que el evento ocurre?

Desarrollo: 1) Fijación de la precisión con la cual se desea establecer el dato. Es habitual exigir un 0,05 (es decir, que la proporción que se determine estará, como máximo, a un $\pm 5\%$ de la verdadera).

2) Se fija el valor de α en el 5%; $Z_{0,05} = 1,96$ (prueba bilateral).

3) Estimación del valor probable: por estudios anteriores, se admite que la misma se presenta en el 6% de las embarazadas ($p = 0,06$).

4) Se aplica la fórmula (4):

$$n = 0,06 \times (1 - 0,06) \times (1,96 / 0,05)^2 = 86$$

Conclusión: si realmente la frecuencia de preeclampsia es 6%, se requiere incorporar a 86 embarazadas al estudio descriptivo para afirmar la misma con una precisión de $\pm 5\%$, existiendo un 95% de probabilidades de que la verdadera frecuencia se halle entre 1% y 11% (valores que resultan de restar y sumar, respectivamente, 5 –precisión admitida- a la frecuencia observada, 6%).

Si la precisión deseada hubiese sido 1%, manteniendo iguales los restantes criterios, el TM necesario para corroborar la frecuencia (6%) con dicha tolerancia (1%) se eleva a 2.167. El mensaje diría que la frecuencia de preeclampsia en la población es un valor que se halla entre 7% y 8% (restar y sumar 1 a 6), con una probabilidad del 95% de estar acertado.

Cálculo del TM en el análisis de variables continuas. Se llevan a cabo para conocer el valor de la media de una variable. Es necesario estimar el desvío estándar (DE) de la variable a estudiar.

Seleccionar la precisión con la que se desea establecer el dato, expresada en unidades similares a las que definen la media (ej: gramos, centímetros, etc.). Dado que la precisión es hacia ambos lados de la media (\pm), el total de la dispersión (TD) admisible se obtiene al multiplicarlo por 2.

Establecer el nivel de confianza con el que se determinará el dato. Habitualmente se adopta el 95%. Dado que éste se extenderá hacia ambos lados (\pm), el valor de $Z_{0,05}$ es 1,96.

Utilizando esta información se aplica la siguiente fórmula:

$$n = 4 \times Z_{\alpha}^2 \times DE^2 / TD^2 \quad (5)$$

Es posible carecer de los datos necesarios para efectuar este cálculo, a pesar de haber realizado extensas búsquedas bibliográficas. En tal caso, conviene transformar la variable continua en una discreta, adjudicándole un punto de corte que demarque dos opciones.

Ejemplo

Interrogante: si se estima que, en una determinada ciudad, el desvío estándar del peso medio al nacer de los varones a término es de 400 gramos, ¿cuántos recién nacidos varones deben pesarse para establecer con exactitud el valor medio de esta variable?

Desarrollo: 1) Fijación de la precisión con la cual se desea establecer el dato. En este ejemplo, los investigadores pretenden establecer el dato con una precisión de 100 gramos. Dado que éste debe multiplicarse por 2 TD = 200.

2) El valor de α es siempre bilateral ($Z_{0,05} = 1,96$).

3) Estimación del desvío estándar de la variable peso al nacer en la población a estudiar (en este ejemplo, DS = 400).

4) Con los datos obtenidos, se aplica la fórmula (5).

$$n = (4 \times 1,96^2 \times 400^2) / 200^2 = 61$$

Conclusión: se requiere pesar a 61 recién nacidos varones de término, para establecer con precisión de ± 100 gramos la media del peso al nacer, admitiendo una probabilidad de errar en la estimación menor al 5%.

Si el tamaño de la precisión deseado hubiese sido ± 20 gramos, el TM requerido se transformaría en 1.537.

Si se hubiese carecido de los datos necesarios, se podría haber optado por clasificar a los recién nacidos en mayores de un determinado peso e iguales o menores al mismo. Ello es proceder a transformar una variable continua en una discreta, lo que facilitará el cálculo del TM, aunque esto deba acompañarse de la adecuación de la pregunta a contestar.

En los ejercicios presentados se observa claramente cómo las decisiones del investigador (ej: precisión con la que desea establecer el dato), tienen un profundo efecto en el TM. La precisión mínima con que un dato mantiene su utilidad es una determinación exclusiva del investigador, y para ello debe aplicar el más depurado criterio médico, ya que no existen reglas ni cálculos al respecto.

¿Qué sucede si no se alcanza el tamaño muestral apropiado?

Las conclusiones del estudio estarán potencialmente viciadas por el error del azar con una probabilidad de ocurrir mayor que la preespecificada, por lo cual la incertidumbre persistirá sobre los resultados. Esta situación torna al estudio, que insumió esfuerzos y recursos, en potencialmente erróneo y aun nocivo, pues no percatarse de su invalidez puede sugerir conductas equívocas.

Bibliografia

1. Monographs in Epidemiology and Biostatistics, Vol. 10 Kelsey J, Douglas Thompson W, Evans A. Methods in Observational Epidemiology. Oxford University Press. 1986
2. An Epidemiologic Approach to Reproductive Health. Ed. Phyllis Wingo, Higgins J. Rubin G. Zahniser. Centers for Disease Control, Atlanta Georgia USA.
3. Rothman K: Modern Epidemiology. Eds. Little Brown & Co. USA. 1986
4. Pocock SJ: Clinical Trials: a practical approach. Eds. John Wiley & Sons, UK 1983.
5. Hulley SB y Cummings SR: Designing Clinical Research: an epidemiological approach. Ed. Williams & Wilkins, Baltimore, 1988.
6. Monographs in Epidemiology and Biostatistics. Vol 2 Scjlesselman J Case-Control Studies: Design, Conduct, Analysis. Oxford University Press. 1982.
7. Gardner MJ y Altman DG: Statistics with confidence: Confidence intervals and statistical guidelines. Ed. Brits Medical Journal London UK 1988.
8. Effective Care in Pregnancy and Childbirth. Eds. I Chalmers, M Enkin y MJNC Keirse. Oxford University Press. 1989.