



Lectura crítica en pequeñas dosis

Diseños secuenciales: interrupción precoz del ensayo clínico

Manuel Molina Arias

Publicado en Internet:
04-enero-2017

Manuel Molina Arias:
mma1961@gmail.com

Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid. Comité de Pediatría Basada en la Evidencia de la AEP y la AEPap. Editor de www.cienciasinseso.com

Palabras clave:

- Ensayos secuenciales
- Análisis intermedios

Resumen

Al contrario que en el ensayo en paralelo habitual, en los diseños secuenciales el tamaño muestral no es fijo, sino que depende de los resultados de análisis intermedios que se realizarán a lo largo del estudio. Son los resultados de estos análisis intermedios los que determinarán si el ensayo puede finalizar o debe continuar hasta alcanzar un tamaño muestral máximo previamente fijado.

Sequential designs: early interruption of clinical trial

Key words:

- Sequential trials
- Interim analysis

Abstract

Unlike in the usual parallel clinical trials, in sequential designs the sample size is not fixed, but depends on the results of interim analysis that will be carried out throughout the study. The results of these interim analyzes will determine if the trial can be interrupted or must continue until reaching a previously set maximum sample size.

Uno de los requisitos de buena calidad metodológica de los ensayos clínicos en paralelo habituales es el cálculo previo del tamaño muestral¹. Este pretende saber el número mínimo de participantes que será necesario para considerar como estadísticamente significativa una diferencia de efecto clínicamente importante entre las dos ramas del ensayo, la de intervención y la de control. De esta manera, el ensayo comienza al abrirse el periodo de inclusión y termina cuando se obtienen los resultados del número de participantes previamente calculado.

Sin embargo, en ocasiones los autores encuentran motivos para la interrupción y finalización precoz del ensayo. Estos motivos pueden ser los siguientes²:

- Eficacia: una de las dos ramas se muestra claramente más eficaz que la otra.
- Seguridad: puede ocurrir que se observen un gran número de efectos adversos o que estos sean muy graves.
- Éticos: si se observa un beneficio claro en una de las ramas se plantea un conflicto ético, ya que se está privando a los participantes de la otra rama de la intervención más eficaz.

Cómo citar este artículo: Molina Arias M. Diseños secuenciales: interrupción precoz del ensayo clínico. Rev Pediatr Aten Primaria. 2017;19:87-90.

- Económicos: no cabe duda que finalizar antes un ensayo clínico supone un ahorro en tiempo y recursos.
- Nueva información: puede ocurrir que a lo largo del ensayo se disponga de información nueva que haga que no tenga sentido hacer la comparación entre las dos ramas, ya sea por motivos de eficacia o de seguridad.
- Otros motivos diversos, como dificultades para el reclutamiento de los participantes.

Existe un tipo de ensayos clínicos en los que esta interrupción precoz del estudio está prevista de antemano durante la fase de diseño del estudio. Hablamos de los ensayos clínicos secuenciales².

DISEÑOS SECUENCIALES

En el ensayo clínico secuencial se realiza cada cierto tiempo un análisis intermedio de los resultados para detectar precozmente las situaciones en las que una de las dos ramas es claramente superior a la otra. Si esto es así, el ensayo se da por finalizado. En caso contrario, prosigue la inclusión de pacientes hasta el siguiente análisis intermedio o hasta alcanzar un tamaño muestral máximo previamente definido, momento en el que se acepta la hipótesis nula de igualdad de efectos si el análisis intermedio no muestra diferencias estadísticamente significativas. El beneficio claro es la posibilidad de finalizar el estudio con un tamaño muestral menor, aunque hay que pagar el precio de una mayor complejidad en el análisis de los resultados³.

Este tipo de diseños son típicos de los ensayos de fase III en los que se evalúa la eficacia o la toxicidad de los tratamientos. Como es lógico, los criterios de interrupción son más conservadores en caso de falta de eficacia que de toxicidad, pero, en cualquier caso, deben estar previamente definidos en el protocolo del ensayo.

Los requerimientos para poder realizar un diseño secuencial son los siguientes²:

- El tamaño muestral debe ser grande, para permitir el dividir la muestra en bloques y realizar los sucesivos análisis intermedios.
- El objetivo del ensayo debe ser único y estar muy bien definido.
- El ensayo debe tener una duración prevista prolongada de, al menos, dos años.
- Tiene que haber un tiempo corto entre inclusión, aleatorización y obtención de la respuesta al tratamiento. En caso contrario no sería posible realizar los análisis intermedios antes de reclutar nuevos participantes.
- Por último, los análisis intermedios deben realizarse por un grupo independiente, ya que si lo hiciesen los propios investigadores se rompería el enmascaramiento. El resultado del análisis debe comunicarse como la finalización del estudio o su continuación, pero no debe proporcionarse a los investigadores el valor de la significación estadística conseguida.

INCONVENIENTES DE LOS DISEÑOS SECUENCIALES

Los ensayos secuenciales tienen tres inconvenientes principales³:

- Al hacerse múltiples comparaciones aumenta el error de tipo I en cada comparación, con el consiguiente riesgo de falsos positivos (dar por bueno un efecto que en realidad no existe)⁴. Por ejemplo, si hacemos cinco comparaciones, la probabilidad de cometer un error de tipo I sube hasta 0,22, y este valor será más alto cuantos más análisis intermedios realicemos. Por este motivo, es imprescindible realizar los ajustes estadísticos precisos para mantener el valor α global en 0,05, utilizando para ello algunos de los métodos que trataremos a continuación.
- La repetición de análisis intermedios favorece la sobreestimación del efecto. Si repetimos un ensayo varias veces, la medida de efecto que utilicemos oscilará alrededor de un valor determinado simplemente por la existencia de un error aleatorio. Así, si repetimos el análisis, las fluctuaciones que tiendan a exagerar el efecto llevarán con más frecuencia a la interrupción del ensayo. Si interrumpimos un ensayo por un

valor de p especialmente bajo, es posible que, si hubiésemos continuado, el siguiente valor de p fuese más alto, simplemente por el fenómeno de regresión a la media.

- Favorecen la adopción de tratamientos ineficaces. Es una conclusión que se desprende de los dos puntos anteriores.

Por todo lo anterior, es fácil comprender la necesidad de que exista un comité de monitorización independiente que tenga conocimiento de los resultados y capacidad para decidir cuándo debe finalizarse el ensayo y cuándo debe continuar.

Además, las reglas de decisión y los análisis intermedios deben estar establecidos en el protocolo del ensayo.

MÉTODOS DE ANÁLISIS SECUENCIALES

De manera somera, describiremos los tres métodos más utilizados para la realización de ensayos secuenciales²:

- Método de análisis secuencial por grupos con regla de finalización. Consiste en dividir el tamaño total de la muestra en un número fijo de bloques tras los que se realizan los análisis intermedios. En cada análisis se hace el contraste estadístico con la técnica habitual, pero ajustando en cada uno el valor de la significación estadística para mantener el α global en 0,05. A este método corresponden las reglas de Pocock, de Peto y de O'Brien-Fleming. Lo habitual es proporcionar unas tablas con los valores ajustados de p en función de los resultados y el número de participantes².
- Métodos basados en límites continuos. Son matemáticamente más complejos. Básicamente, se definen dos funciones denominadas de diferencia acumulada entre tratamientos y de información acumulada. Con los resultados de estas funciones se obtienen unas coordenadas que se representan gráficamente para definir su relación con una región de continuación. Si el valor de la coordenada está en el interior de esta región, el estudio debe continuar; en caso

contrario, se procede a su interrupción. En la práctica, este método se utiliza empleando programas estadísticos disponibles a tal fin.

- Métodos basados en una función de gasto de α . Estos métodos, como el de DeMets⁵, son muy populares porque permiten realizar todos los análisis intermedios que se desee sin necesidad de establecerlos previamente. Se basa en disminuir el valor de significación estadística para evitar la posibilidad de error por las comparaciones múltiples. Su base matemática consiste en crear una función de gasto de α que especifica como varía su nivel en cada análisis intermedio en función del resultado y de los datos previos, con objeto de mantener el α global constante en 0,05.

Una vez más, lo habitual es emplear programas estadísticos que implementan estos métodos. Decir, por último, que las reglas de Pocock y de O'Brien-Fleming pueden convertirse en funciones de gasto de α mediante fórmulas matemáticas disponibles para tal fin.

LECTURA CRÍTICA DE ENSAYOS SECUENCIALES

Los ensayos clínicos secuenciales son cada vez más populares en la literatura médica, sobre todo por el ahorro que supone demostrar el efecto de una intervención con un tamaño muestral menor. Sin embargo, lejos de considerarlo como un signo de calidad metodológica, debemos ser escépticos con los ensayos interrumpidos precozmente y comprobar que se cumplen una serie de requisitos metodológicos para poder dar validez a sus resultados^{3,6}.

Tal como especifica la declaración CONSORT⁷, los autores del ensayo deben informar del tamaño muestral planeado, del análisis intermedio en el que se toma la decisión de interrumpir el ensayo y de la regla de decisión que se ha empleado.

Esta regla de decisión debe estar especificada con anterioridad en el protocolo del ensayo. Toda interrupción no planificada debe figurar como una

enmienda al protocolo del ensayo y estar debidamente justificada.

Como ya hemos dicho, es fundamental prestar atención al número de análisis intermedios realizados y al método utilizado para evitar el problema de los falsos positivos con las comparaciones múltiples. Debemos desconfiar de los resultados del estudio si el valor de p no se especifica o si este es ligeramente inferior a 0,05, ya que puede corresponder con una fluctuación aleatoria. Hay que tener en cuenta que con métodos como el de Peto, son habituales valores de p menores de 0,001 para alcanzarse significación estadística.

Para finalizar, en los casos en los que el estudio se interrumpe por observar un beneficio claro de una

de las dos ramas, puede ser útil comparar la muestra calculada al inicio con la empleada finalmente. Si existe mucha desproporción o el efecto es exageradamente beneficioso, podemos estar ante una sobreestimación del efecto, especialmente en los casos en los que la frecuencia del evento en estudio sea bajo, en los que el error aleatorio será mayor.

CONFLICTO DE INTERESES

El autor declara no presentar conflictos de intereses en relación con la preparación y publicación de este artículo.

BIBLIOGRAFÍA

1. Wang D, Bakhai A. Sample size and power. En: Duolao W, Bakhai A (eds.). *Clinical trials. A practical guide to design, analysis and reporting*. Londres: Remedica; 2006. p. 81-7.
2. Monleón-Getino T, Barnadas-Molins A, Roset-Gamisans M. Diseños secuenciales y análisis intermedio en la investigación clínica: tamaño frente a dificultad. *Med Clin (Barc)*. 2009;132:437-42.
3. Latour-Pérez J, Cabello-López JB. Interrupción precoz de los ensayos clínicos. ¿Demasiado bueno para ser cierto? *Med Intensiva*. 2007;31:518-20.
4. Molina Arias M. El problema de las comparaciones múltiples. *Rev Pediatr Aten Primaria*. 2014;16:367-70.
5. DeMets DL, Lan KKG. Interim analysis: the alpha spending function approach. *Statistics Med*. 1994;13:1341-52.
6. Pocock SJ. When (not) to stop a clinical trial for benefit. *JAMA*. 2005;294:2228-30.
7. González de Dios J, Buñuel Álvarez JC, González Rodríguez MP. Listas guía de comprobación de ensayos clínicos: declaración CONSORT. *Evid Pediatr*. 2011;7:72.