

I C B

Investigación Clínica y Bioética

Boletín de la Sociedad Española de Farmacología Clínica dirigido a los Comités Éticos de Investigación Clínica

El método secuencial en el ensayo clínico

Cuando para una determinada enfermedad se dispone ya de un tratamiento eficaz, si se desea probar un nuevo fármaco, éste debe ser comparado con el tratamiento activo establecido. Salvo excepciones, las diferencias en eficacia que se pueden hallar son de pequeña magnitud y eso implica que el tamaño de la muestra necesaria es grande. La inclusión de pacientes en un ensayo clínico se realiza de manera gradual y por lo general lenta, según el número de pacientes con la patología de interés. Estas razones hacen que la duración de los ensayos pueda ser larga y que sea necesario, por motivos éticos, realizar análisis intermedios antes de completar la inclusión del número total previsto de pacientes. Estos análisis deben supervisar el cumplimiento del protocolo y la calidad de la información recogida y deben evaluar las posibles reacciones adversas, pero también pueden servir, si se emplean los métodos adecuados, para valorar la eficacia del nuevo tratamiento y detener el estudio en caso de que ya se pueda obtener una respuesta concluyente a las preguntas iniciales que motivaron el ensayo.

Las pruebas estadísticas clásicas que se emplean para analizar ensayos clínicos sólo permiten realizar un único análisis al final, cuando se dispone de los datos sobre todos los pacientes. Si se realizan observaciones intermedias y se finaliza el estudio porque se hayan hallado diferencias importantes, los valores de significación calculados con métodos clásicos no son válidos, pues cuantos más análisis se realicen, más probable será que alguno de ellos, por azar, muestre diferencias significativas, cuando en realidad los dos tratamientos pueden ser equivalentes. En este caso se comete un error de tipo I con probabilidad mayor de la estable-

cida en el diseño. Sin embargo, hay muchas razones para finalizar un estudio con el número de pacientes mínimo posible: económicas –en tiempo y dinero– y éticas, pues se reduce el número de pacientes expuestos al tratamiento menos eficaz.

Los **métodos estadísticos secuenciales** permiten realizar un tratamiento adecuado del problema de los análisis múltiples. Estas técnicas se desarrollaron como secretos militares durante la Segunda Guerra Mundial para realizar el control de calidad del armamento, y se empezaron a aplicar a los ensayos clínicos en los años setenta. A pesar de ser métodos antiguos, no han tenido demasiada aceptación. Su principal ventaja es que, si se realiza un diseño correcto, pueden ahorrar un gran número de pacientes, pues el estudio finaliza justo cuando se puede asegurar una conclusión a la vez que se mantienen las probabilidades estadísticas y en los valores predeterminados. Sus detractores exageran los inconvenientes que supone una organización del ensayo más meticulosa, pues para obtener provecho del diseño secuencial, se debe definir un único objetivo claro y se deben conocer los resultados de la intervención poco tiempo después de la inclusión del paciente. El hecho de que no se pueda saber con exactitud el número de pacientes que se incluirán puede dificultar la planificación de recursos. También se les achaca una mayor complicación matemática en los métodos estadísticos de diseño y análisis, pero este inconveniente es cada día menos importante gracias al desarrollo de programas de ordenador que realizan los cálculos tediosos. Los métodos secuenciales pueden adaptarse a cualquier tipo de variable de interés, incluso el tiempo de supervivencia. También existe la posibilidad de em-

plear análisis multivariante para ajustar factores de riesgo que queden mal equilibrados entre los grupos a comparar. Existen en la actualidad varias estrategias posibles para diseñar un ensayo clínico secuencial. Cada una tiene sus ventajas e inconvenientes, y la elección debe basarse en las características y los objetivos del ensayo. Todas ellas permiten tomar la decisión acerca de la superioridad de un nuevo tratamiento, sobre la base de unas especificaciones estadísticas determinadas (el nivel de significación, α y el poder estadístico para detectar una diferencia deseada, $1 - \beta$). A continuación se comentan brevemente las características más importantes de cada método.

Métodos secuenciales por grupos

Esta estrategia, también llamada pruebas de significación repetidas, fue introducida por Armitage¹ y revisada por Pocock.² Consiste en definir *a priori* un número determinado máximo, K , de análisis intermedios. Éstos se realizarán a medida que se cuente con la información referente a un grupo de pacientes de tamaño preestablecido, según las especificaciones del ensayo. Se emplea en cada análisis intermedio un nivel de significación α' menor que el habitual, de 0,05. Si se encuentran diferencias significativas a ese nivel, el estudio finaliza; en caso contrario, se incluye otro grupo de pacientes de igual tamaño y se repite la prueba estadística con todos los pacientes acumulados. Si al finalizar el análisis con K grupos no se encuentran diferencias, el estudio finaliza con resultado negativo y se acepta la hipótesis nula con riesgo β . El valor de α' se calcula de manera que, en función del número de análisis intermedios a realizar (K) se mantenga en conjunto el nivel de significación en 0,05. Dado que α' es menor que α , el número total de pacientes necesario para finalizar el ensayo puede ser mayor que el de un estudio con tamaño fijo si la hipótesis nula es cierta.

Existe la posibilidad de emplear valores de α' variables en cada análisis intermedio, de manera que en los primeros, cuando hay pocos pacientes, sea difícil rechazar la hipótesis nula, mientras que en análisis sucesivos ello sea más fácil. Este diseño, propuesto por O'Brien y Fleming,³ hace difícil la finalización prematura del estudio a no ser que las diferencias sean enormes, lo cual pone en entredicho el interés de este diseño secuencial. Por otro lado, el método de Pocock facilita mucho la finalización en los primeros análisis intermedios y el pequeño tamaño de muestra empleado, aunque permite concluir que los tratamientos son diferentes, pro-

porciona estimaciones de las diferencias con unos intervalos de confianza excesivamente amplios. Aunque los métodos secuenciales por grupos son sencillos de emplear, imponen una rigidez en su diseño más estricta que el diseño clásico con el tamaño fijo, pues los análisis deben realizarse exactamente al obtener los resultados de un grupo de pacientes de tamaño determinado. Algunos autores han estudiado este problema⁴ y en general no se modifican los resultados por pequeñas desviaciones del plan, como realizar análisis en función del calendario en lugar de hacerlos según el número de pacientes, siempre que la tasa de inclusión sea más o menos constante.

Métodos basados en límites continuos

Estos métodos, descritos inicialmente por Armitage⁵ y Anscombe,⁶ han sido elaborados en profundidad por Whitehead y Jones⁷ y permiten plantear una amplia variedad de diseños secuenciales. Consisten en definir dos fórmulas estadísticas a partir de los datos que se van acumulando. Una, llamada Z , valora la diferencia acumulada entre los dos tratamientos. Otra, llamada V , calcula la información acumulada y es proporcional al tamaño de la muestra. Cada vez que se desea realizar un análisis intermedio se calculan los valores de Z y V y se dibujan en una gráfica (véase la figura 1) en la que se ha definido una *región de continuación*. Así, mientras las coordenadas de (V, Z) se encuentren en el interior de esa región, el estudio debe continuar. En el momento en que cruce cualquiera de los límites de esta región, el estudio finaliza. En los diseños más sencillos, como el de la figura, si se cruza el límite superior, se debe a que existen diferencias a favor del nuevo tratamiento. Si se cruza el límite inferior, el estudio finaliza con resultado negativo. Las ventajas de estos diseños

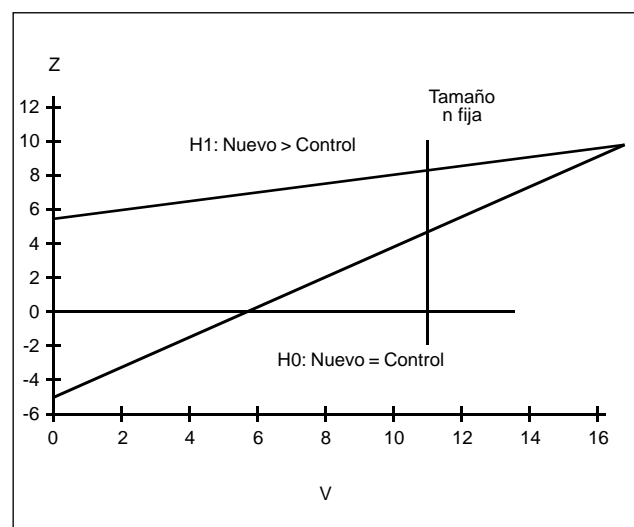


Figura 1. Región de continuación del método triangular secuencial.

respecto a los secuenciales por grupos radican en su flexibilidad. Pueden realizarse tantos análisis intermedios como se desee, de manera regular o irregular. Incluso se puede calcular el valor de Z y V tras la finalización de la participación de cada paciente, pero es importante que la frecuencia de las inspecciones no dependa de que el camino del ensayo se esté acercando a un límite. Además, este tipo de diseño permite finalizar pronto un estudio que tiene resultados negativos, mientras que el diseño secuencial por grupos debe emplear la totalidad del tamaño de muestra calculado para llegar a esta conclusión. Se pueden emplear regiones de continuación con formas variadas. Las más adecuadas son las que se definen por el *método de la razón de probabilidades secuencial*, con límites paralelos y abiertos, y por el *método triangular*, dibujado en la figura. El cálculo de la posición de los límites se basa en métodos matemáticos laboriosos, pero existe un programa de ordenador⁸ que realiza estos cálculos y que proporciona datos teóricos sobre el tamaño de muestra esperado en función del diseño y las características del ensayo. En promedio, los métodos secuenciales con límites continuos permiten un ahorro de alrededor de un 40% de los pacientes necesarios en un diseño fijo. Existe siempre, sin embargo, una pequeña proporción (menor del 10%) de estudios que precisan más pacientes que los necesarios en un diseño clásico.

Métodos basados en una función de gasto de

Puesto que el principal problema en los análisis múltiples es que se incrementa el nivel de significación por encima de lo deseado, se han propuesto métodos para compensar el “gasto” de a medida que se realizan análisis, de manera que se mantenga el valor global deseado.⁹ Estos métodos también son flexibles, en el sentido de que permiten realizar los análisis que se desee y cuando se desee. En la práctica, consisten en definir una función matemática que describe cómo se modifica el nivel de significación que debe emplearse en los análisis en función de la cantidad de información acumulada y del número de análisis previos. Al igual que en los diseños secuenciales por grupos, si se llega a un máximo de tamaño de muestra preestablecido y no se ha rechazado la hipótesis nula, el estudio finaliza con resultado negativo. Existen múltiples propuestas de funciones de gasto posibles, pero no se han desarrollado programas de ordenador que faciliten los cálculos, por lo que estos métodos son de difícil aplicación.

Intervalos de confianza repetidos

Cuando el interés de un ensayo consiste en estimar con gran precisión el efecto de un tratamiento, puede emplearse la técnica propuesta por Jennison y por Turnbull.¹⁰ Está basada en el método secuencial por grupos, según en cual cada vez que se acumula un grupo de pacientes de tamaño preestablecido se calcula el intervalo de confianza (en general al 95%) de la diferencia entre tratamientos. Si el intervalo excluye el cero, el estudio puede finalizar con un resultado positivo. Si se alcanza el tamaño de muestra final, el intervalo de confianza calculado entonces tendrá la propiedad de contener el valor real de la diferencia entre tratamientos con una probabilidad del 95%.

Otros métodos de análisis secuencial

Se pueden emplear otras técnicas de análisis secuencial, como el método de la restricción estocástica (*stochastic curtailment*) propuesto por Lan, Simon y Halperin,¹¹ el cual, en condiciones adecuadas, permite detener prematuramente un estudio diseñado con tamaño fijo. También se han propuesto métodos bayesianos, que se basan en la teoría *subjetiva* de la probabilidad. Aunque hay una gran controversia sobre su validez, estas técnicas estadísticas pueden dar respuesta a algunos problemas que no podrían ser abordados con los métodos clásicos.

Tabla 1. Criterios para considerar adecuado un diseño secuencial.

Cuándo puede considerarse adecuado:

- Gran tamaño de muestra necesario.
- El objetivo del estudio (*end point*) es único y está bien definido.
- Se prevé una larga duración del estudio.
- Se conoce el resultado de un paciente pronto, mucho antes de que finalice la fase de inclusión.

Cuándo no interesa un diseño secuencial:

- El número de pacientes necesario es reducido.
- El estudio tiene objetivos múltiples.
- El resultado de los pacientes se conoce muy tarde, una vez finalizado el período de inclusión.
- La logística del estudio no permite realizar análisis intermedios.

Referencias

1. Armitage P. Sequential methods in clinical trials. *Am J Pub Health* 1958;48:1395-402.
2. Pocock SJ. Group sequential methods in the design and analysis of clinical trials *Biometrika* 1977;64:191-9.
3. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1980;35:549-56.
4. Jennison C. Efficient group sequential tests with unpredictable group sizes. *Biometrika* 1987;74:155-65.
5. Armitage P. Sequential tests in prophylactic and therapeutic trials. *QJ Med* 1954;23:255-74.
6. Anscombe FJ. Sequential medical trials. *J Am Statist Assoc* 1963;58:365-83.
7. Whitehead J, Jones DR. The analysis of sequential clinical trials. *Biometrika* 1979;66:443-52.
8. Whitehead J, Brunier H. *PEST3.0 Operating Manual*. Reading: Reading University, 1993.
9. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983;70:659-63.
10. Jennison C, Turnbull BW. Repeated confidence intervals for group sequential clinical trials. *Controlled Clin Trials* 1984;5: 33-45.
11. Lan KKG, Simon R, Halperin M. Stochastically curtailed tests in long-term clinical trials. *Sequential Analysis* 1982;1: 207-19.

Víctor Moreno Aguado. *Laboratorio de Bioestadística y Epidemiología. Facultad de Medicina. Universitat Autònoma de Barcelona.*

Consejo editorial: Diego Gracia (Director), Joan-Ramon Laporte (Jefe de Redacción), Carles Vallvé (Secretario), Antonio Carcas, Xavier Carné, Carlos Romeo, Joan Costa, Francisco de Abajo, José Antonio Durán, Magí Farré, Javier Galiana, Fernando García Alonso, Angela Idoipe, Antonio Portolés, Emilio Sanz, M^a Antonia Serrano, Joan-Antoni Vallès.

Comité de redacción: Xavier Carné, Fernando García Alonso, Joan-Ramon Laporte, M^a Antonia Serrano, Carles Vallvé.

© Sociedad Española de Farmacología Clínica
ISSN 1131-8910
Dep. Legal B-4772-1995

Servicio de Farmacología Clínica
CSU Vall d'Hebron
08035 Barcelona

Tel.: 934 28 3029
Fax: 934 89 4109
Suscripción: 1.300 ptas. al año

El Boletín *ICB* es editado trimestralmente por la Sociedad Española de Farmacología Clínica gracias a una ayuda del Ministerio de Sanidad y Consumo. Además, las siguientes instituciones contribuyen a su publicación: Bayer, Boehringer Ingelheim, Ciba Geigy, Esteve, Ferrer, Glaxo, Grifols, ICN Ibérica, Janssen, Merk Sharp & Dohme, Menarini, Prodesfarma, Rhône-Poulenc Rorer, Rovi, Sandoz, Searle, Syntex, Vita y Zeneca. Los artículos y notas publicados en *ICB* no pueden ser utilizados para anuncios, publicidad u otra promoción de ventas, ni pueden ser reproducidos sin permiso escrito.