

9

Nuevos diseños en investigación clínica

Desafíos y oportunidades

CARIDAD PONTES

JOSÉ RÍOS

FERRAN TORRES

SUMARIO

1. Introducción	244
2. Nuevos diseños en investigación clínica	245
2.1. Métodos secuenciales	245
2.2. Métodos adaptativos	251
2.3. Aproximaciones bayesianas	253
2.4. Modelización y simulación	254
2.5. Diseños enriquecidos	255
2.6. Alternativas a los ensayos controlados paralelos	256
3. La necesidad de cambios	259
3.1. Aumentar la eficiencia del proceso de investigación	259
3.2. Fusión de programas exploratorios y confirmatorios	260
3.3. Ensayos clínicos en poblaciones pequeñas	260
4. Importancia de validar la selección del diseño	262
5. Conclusiones	262
Referencias	264

LO QUE VD. APRENDERÁ EN ESTE CAPÍTULO

1. Los diseños flexibles, principalmente de tipo secuencial y adaptativo, destacan como alternativas al diseño clásico, de grupos paralelos con asignación aleatoria en condiciones de doble ciego.
2. Los métodos flexibles no son medios para reducir la rigurosidad de la planificación a priori de los ensayos clínicos, sino herramientas para gestionar situaciones experimentales complejas.
3. Los diseños adaptativos permiten modificar de manera preespecificada el diseño del estudio durante su ejecución.
4. Los diseños secuenciales permiten definir la finalización de un estudio basándose en la información que se va obteniendo durante su ejecución.
5. Los diseños bayesianos emplean tanto la información previa como la obtenida en el estudio para obtener información integrada de la evidencia acumulada.
6. Los modelos biológicos y estadísticos permiten establecer predicciones y refinar el diseño de los ensayos clínicos.
7. Los sistemas de información han eliminado obstáculos a la implementación práctica de las aproximaciones secuenciales y adaptativas.
8. En poblaciones pequeñas, raras o pediátricas, los métodos alternativos pueden aumentar la viabilidad y la eficiencia de los ensayos clínicos.
9. Debe mantenerse un equilibrio apropiado entre las necesidades de eficiencia estadística y las de obtener resultados clínicamente relevantes e interpretables.

1. INTRODUCCIÓN

El diseño comparativo de grupos paralelos es un diseño robusto, intuitivo y sólido, que implementa literalmente la formulación del método científico: dos o más grupos entran simultáneamente en un mismo proceso de tratamiento y seguimiento, de duración constante, manteniendo como diferencia un único factor en estudio, la intervención estudiada. La asignación aleatoria de las intervenciones en condiciones de enmascaramiento de la identidad controla el riesgo de sesgos de asignación y evaluación. El diseño establecido prospectivamente y ejecutado conforme a un protocolo detallado permite establecer una selección a priori de los márgenes de errores estadísticos aceptables, y la conducción en paralelo asegura que los grupos reciben las intervenciones en estudio y son evaluados en condiciones semejantes. Las diferencias entre los grupos al final del proceso serán atribuibles exclusivamente al factor diferenciador, es decir, a las intervenciones en estudio, permitiendo establecer relaciones de causalidad.^{1,2}

Existen multitud de diseños experimentales distintos al clásico estudio con asignación aleatoria de grupos paralelos, cuya finalidad es adaptar los estudios a ciertas particularidades de las situaciones estudiadas. Por ejemplo, en los estudios cruzados completos o incompletos los pacientes reciben más de un tratamiento, de manera que se optimiza el análisis a partir de la variabilidad intra-individual; son útiles en patologías poco frecuentes cuando los factores en estudio son independientes de la evolución temporal.³ En los ensayos clínicos con

asignación aleatoria por conglomerados (*clusters*) no se realiza la asignación de individuos a las distintas intervenciones, sino que se asignan grupos de pacientes que comparten características comunes, generalmente geográficas o de atención sanitaria (véase capítulo 8). Permiten la asignación aleatoria cuando, o bien no es posible realizarla a nivel individual, o bien, si esta se realiza, se corre el riesgo de *contaminación* entre los grupos. Como contrapartida, requieren mayor tamaño muestral porque se debe incluir la variabilidad entre conglomerados en el análisis.⁴

Otro ejemplo son los diseños flexibles, como los secuenciales o los adaptativos; ambos aprovechan la información a medida que se obtiene para aumentar la eficiencia del estudio. Los estudios secuenciales continúan incorporando información hasta que esta es lo suficientemente precisa como para poder rechazar la hipótesis nula, en el caso de un estudio positivo, o bien para poder concluir que la probabilidad de rechazarla si se continua el estudio es demasiado baja (futilidad). Estos métodos son especialmente útiles cuando el periodo de reclutamiento es prolongado y la duración del tratamiento de un paciente es relativamente corta. Los diseños adaptativos permiten el ajuste del estudio a las condiciones reales a partir de la información obtenida durante su ejecución, maximizando la eficiencia del experimento, y en general, reduciendo sus riesgos. Muchas aproximaciones alternativas al clásico ensayo clínico paralelo producen como resultado un menor número de pacientes tratados y una menor duración del estudio, lo que representa en general una ventaja

económica y estratégica cuando la investigación está dirigida al registro de nuevos medicamentos.

No obstante, la intuición y la robustez del diseño paralelo para aportar conclusiones de causalidad consistentes y fiables cuando se realiza de manera correcta, han hecho que desde hace décadas este sea el estándar metodológico de referencia en investigación clínica por delante de otros diseños. Actualmente se considera que satisface los requerimientos metodológicos más exigentes, especialmente los de las autoridades reguladoras que, desde una posición de minimización de los riesgos potenciales para la salud pública, generalmente prefieren basar sus decisiones de comercialización de nuevos tratamientos en los resultados de ensayos de grupos paralelos con asignación aleatoria en condiciones de doble ciego (o doble enmascaramiento). Esta reticencia de las autoridades reguladoras a la toma de decisiones a partir de aproximaciones metodológicas poco habituales, junto con la elevada y creciente tasa de fracaso de productos cada vez en fases más avanzadas, hace que también las compañías farmacéuticas sean extremadamente cautelosas a la hora de aceptar diseños confirmatorios distintos del clásico ensayo clínico de grupos paralelos, con el fin de reducir el «riesgo regulador» de sus proyectos.

2. NUEVOS DISEÑOS EN INVESTIGACIÓN CLÍNICA

Los llamados nuevos diseños son en realidad diseños que cuentan, en su mayoría, con una tradición de décadas. Así, en la guía *ICH E9 Principios estadísticos*

para los ensayos clínicos,⁵ publicada en el año 1998, ya podemos encontrar una detallada descripción de aproximaciones secuenciales y adaptativas simples. Los métodos bayesianos de reevaluación continua se han utilizado y se utilizan ampliamente en estudios exploratorios iniciales en oncología, y los diseños adaptativos más complejos se han utilizado en la última década extensamente en los diseños de desarrollos exploratorios.⁶

A pesar de la reticencia de las autoridades reguladoras a basar sus decisiones de autorización de comercialización en metodologías distintas del estándar de referencia comúnmente aceptado, como los estudios que enlazan fases exploratorias y confirmatorias sin solución de continuidad (*seamless designs*),^{7,8} algunos diseños adaptativos complejos ya han servido de soporte central en autorizaciones de comercialización. En la tabla 1 se muestra un pequeño resumen de los distintos diseños alternativos al de grupos paralelos; los más relevantes se comentan a continuación.

2.1. Métodos secuenciales

La metodología secuencial se basa en el concepto de aprovechar la información a medida que se genera. Así, los diseños puramente secuenciales evalúan de manera continua los resultados a medida que se obtienen, trazando una trayectoria de estudio en la que cada par de pacientes es graficado. Se establecen unas normas de interrupción del estudio a partir de regiones de continuación e interrupción, de modo que si la trayectoria del estudio cruza los límites de dichas regiones se concluye el estudio.

Tabla 1. Características principales de los diseños más frecuentes en ensayos clínicos, ventajas e inconvenientes de los mismos

Diseño	Características principales	Ventajas	Inconvenientes
Diseño clásico de grupos paralelos <i>Parallel</i>	Los sujetos se asignan aleatoriamente a uno o más grupos en paralelo y se siguen durante un periodo fijo a cabo del cual se mide la variable principal. Al finalizar la fase experimental se analizan los resultados. Los sujetos no cambian de grupo asignado	Robusto, intuitivo, ampliamente aceptado	El análisis utiliza la variabilidad interindividual (entre grupos), con lo que el tamaño muestral es mayor que en diseños cruzados
Cruzado <i>Crossover</i>	Los sujetos se asignan aleatoriamente a una o más secuencias de tratamientos consecutivos, separados por periodos sin tratamiento. Al final de cada periodo de tratamiento se mide la variable principal. Al finalizar la fase experimental se analizan los resultados ajustando por secuencia y periodo	Utiliza la variabilidad intraindividual y por este motivo optimiza el tamaño muestral	Menos robusto, y más largo en ejecución que el diseño paralelo Requiere estabilidad temporal de la enfermedad y ausencia de arrastre del efecto terapéutico (<i>carry over effect</i>) para ser aceptable En los análisis convencionales la pérdida de un sujeto supone la pérdida de todos sus periodos Los diseños incompletos requieren un análisis específico
Con asignación aleatoria en conglomerados <i>Cluster</i>	Se asignan aleatoriamente grupos de pacientes con características comunes, en vez de pacientes individuales	Permite la aleatorización cuando la asignación aleatoria individual es imposible o cuando esta implica un riesgo de <i>contaminación</i> entre los grupos estudiados	Se debe incluir la variabilidad interconglomerado en el análisis, por lo que requieren mayor tamaño muestral
Tratamiento añadido <i>Add-on</i>	Los sujetos continúan con su tratamiento habitual de base, y se asignan aleatoriamente a recibir el tratamiento experimental o un placebo	Facilita la participación de pacientes en estudios en patologías graves	Puede ser más difícil demostrar eficacia adicional a los tratamientos disponibles
N de 1 <i>N-of 1</i>	En un único sujeto se asignan aleatoriamente las estrategias terapéuticas en distintos periodos consecutivos.	Conclusiones con alta validez a nivel del paciente	Poca validez externa en la extrapolación de resultados
De modelización de dosis-respuesta <i>Modelling and simulation for dose and dose regimen selection</i>	Evaluación del efecto de las dosis en la respuesta basándose en técnicas de modelización y simulación para seleccionar la dosis más adecuada para la siguiente fase del desarrollo	Permite estimar el efecto de dosis no estudiadas, incluidas dentro del intervalo estudiado, con una precisión predeterminada	Las estimaciones fuera del intervalo de dosis estudiadas no se consideran adecuadas. Algunos autores recomiendan establecer técnicas de ajuste de multiplicitad de la modelización.
Factorial <i>Factorial</i>	Evaluación simultánea de combinaciones de dos o más intervenciones en un diseño con número de grupos igual al factorial de las intervenciones que se combinan	Evaluación del efecto comparado e interacciones dentro del mismo estudio	El número de grupos a evaluar depende de las posibles combinaciones de los fármacos y el número de dosis, por lo que este puede ser relativamente grande; el tamaño muestral global se verá muy influido por este factor

Diseño	Características principales	Ventajas	Inconvenientes
Retirada aleatoria <i>Randomised withdrawal</i>	Todos los pacientes están en tratamiento con el fármaco en estudio, y aquellos que responden se asignan aleatoriamente a continuar recibiendo el mismo tratamiento o un placebo. Una variante consiste en conmutar a otro tratamiento activo o placebo (<i>randomised switch</i>)	Facilita la aceptación a participar de pacientes con patologías graves Permite evaluar el mantenimiento del efecto una vez este se ha obtenido	Éticamente puede ser cuestionable retirar un tratamiento que es eficaz
Retirada (o abandono) temprano <i>Early-escape</i>	Los pacientes se evalúan repetidamente para valorar su respuesta individual y se retiran del estudio si no alcanzan una respuesta mínima	Limita la duración de la exposición a un tratamiento poco eficaz Útil cuando un tratamiento menos activo puede tener un impacto importante en la salud de los pacientes Mejora la aceptabilidad de participar en estudios con placebo	Se compromete la evaluación comparada a largo plazo
Análisis basado en técnicas de decisión <i>Decision-based analysis</i>	Evaluación complementaria de eficacia y seguridad combinadas. Se realizan modelizaciones y simulaciones para identificar variables y escenarios más proclives para ejecutar determinado estudio. Otros métodos que se basan en la utilización de simulaciones estadísticas, y pueden ser útiles especialmente en poblaciones pequeñas como los diseños de ordenación jerárquica y selección (<i>ranking and selection</i>) y los diseños de comparaciones múltiples (<i>multiple comparison procedures</i>)	Útil para seleccionar diseños pragmáticos Aporta información valiosa para la toma de decisiones por parte de las agencias de evaluación de tecnologías	No puede sustituir la evaluación principal de eficacia de los ensayos clínicos
Aleatorización basada en el riesgo <i>Risk-based allocation</i>	Diseño y análisis de estudios en los que la asignación está guiada por el riesgo subyacente	Permite hacer estudios cuando las características de la enfermedad y su pronóstico dificultan operativamente o hacen éticamente cuestionable la aleatorización	Los pruebas estadísticas estándares no son válidas y los resultados se deben interpretar con cautela
Enriquecimiento de la población <i>Population enrichment</i>	Selección de poblaciones especialmente sensibles al efecto de un tratamiento a partir de marcadores genéticos o de otro tipo para maximizar las posibilidades de un estudio de medir la eficacia	La inclusión selectiva de esta subpoblación más susceptible, aumenta la posibilidad del estudio para evaluar adecuadamente el efecto del tratamiento	La extrapolación de resultados queda limitada a la subpoblación en estudio Deja sin evidencia del beneficio del tratamiento a poblaciones con menor respuesta
Aleatorización del inicio (o de la retirada) <i>Staggered start and delayed withdrawal</i>	Se selecciona aleatoriamente el momento en el que se inicia (o retira) la intervención, y se evalúa cómo dicho momento afecta a la evolución de una determinada enfermedad	Útil para evaluar la modificación en la evolución de una enfermedad	Falta de información comparada con otras intervenciones

Tabla 1. Características principales de los diseños más frecuentes en ensayos clínicos, ventajas e inconvenientes de los mismos

Diseño	Características principales	Ventajas	Inconvenientes
<p>Tiempo hasta el inicio del efecto <i>(time-to-onset)</i></p> <p>Bayesiano <i>Bayesian</i></p>	<p>Evaluación del tiempo hasta el inicio del efecto <i>(time-to-onset)</i>, de manera similar a la evaluación de tiempo hasta supervivencia</p> <p>Grupo amplio de diseños que utiliza información a priori para integrarla con los resultados del estudio actual y concluir los resultados a posteriori combinados. Se utilizan actualmente sobre todo en las fases iniciales del desarrollo clínico</p>	<p>Puede ser útil para evaluar la velocidad en alcanzar un efecto cuando esta sea realmente considerada un elemento clave</p> <p>En general, ofrece mayor potencia y flexibilidad y proporciona resultados de una forma más natural e intuitiva</p>	<p>En función de lo que (no) se evalúe una vez que se obtiene el efecto, se ignorará si este se mantiene y los potenciales efectos beneficiosos y adversos posteriores</p> <p>Aunque hay sistemas para acotar el peso de la información a priori, el resultado final depende siempre de esta</p> <p>Existe debate en cuanto al uso adecuado de información a priori cuando esta procede de datos externos. Este hecho penaliza su aceptación para estudios con firmatorios de eficacia, sobre todo a nivel regulador</p>
<p>Métodos secuenciales por grupo</p> <p><i>Group sequential models</i></p>	<p>Análisis intermedios que usan datos de grupos de pacientes de tamaño predeterminado, y gestionan la multiplicidad de análisis mediante ajustes del error de tipo I. Permiten finalizar prematuramente un estudio si se rechaza la hipótesis nula a partir de los resultados de los análisis intermedios. El análisis de utilidad <i>(utility)</i> permite la finalización prematura cuando la probabilidad de rechazar la hipótesis nula al final del estudio es razonablemente baja, ajustando el error de tipo II</p>	<p>Métodos muy sencillos de aplicar y ampliamente difundidos</p> <p>Los métodos similares a la aproximación de O'Brien y Fleming apelan a requerir incremento en el tamaño de la muestra</p> <p>Los métodos similares a la aproximación de Pocock permiten la finalización prematura en inspecciones iniciales de manera más fácil</p>	<p>La normas de aplicación son bastante rígidas y los tamaños de los grupos a analizar deben estar preestablecidos.</p> <p>Los métodos similares a la aproximación de O'Brien y Fleming hacen muy difícil la finalización en inspecciones iniciales</p> <p>Los métodos similares a la aproximación de Pocock requieren un incremento notable en el tamaño de la muestra</p>
<p>Otros diseños secuenciales</p>	<p>Análisis de los datos más intensivo que los métodos secuenciales por grupos. Destacan los métodos basados en funciones de gasto de alta <i>(alpha spending function)</i> y los basados en la función de verosimilitud <i>(likelihood)</i></p>	<p>Gran flexibilidad en el diseño y en las condiciones de aplicación</p> <p>Algunos de los métodos, como los basados en la función de verosimilitud, permiten mantener el poder estadístico independientemente de las oscilaciones en la variabilidad</p>	<p>Requieren conocimientos estadísticos sólidos para su diseño y análisis</p>
<p>Adaptativo</p> <p><i>Adaptive</i></p>	<p>Grupo de diseños que utilizan los datos acumulados para decidir la modificación de aspectos del estudio durante su ejecución, de una manera pre-determinada y sin menoscabo de la validez o integridad del estudio</p>	<p>Elevado grado de flexibilidad que permite adaptarse a las condiciones específicas de cada situación</p>	<p>Cuanto más adaptaciones y cuanto más afectan estas a puntos clave (v.gr., tipo de población o cambios en la variable principal), menor es su aceptabilidad a nivel regulador</p> <p>Son complejos en el diseño, en el análisis y en los aspectos de planificación y logísticos.</p>

Diseño	Características principales	Ventajas	Inconvenientes
<p>Reestimación del tamaño muestral <i>Sample size re-estimation</i></p>	<p>Adaptación de los cálculos mediante métodos basados en estimaciones de la variabilidad, enmascaradas o con desvelamiento de la asignación. Los métodos que desvelan la asignación pueden combinarse con los secuenciales por grupos y reestimación del tamaño muestral (<i>Group sequential models combined with sample size re-estimation</i>)</p>	<p>Permite adaptarse a cambios respecto a la variabilidad esperada Se puede combinar con métodos secuenciales por grupo para establecer reglas de finalización prematura. La heterogeneidad entre las etapas es menos pronunciada que en los diseños de adaptación de respuesta, y en general no se considera un problema</p>	<p>Los métodos que desvelan el enmascaramiento son poco aceptados por las agencias reguladoras, y requieren una organización muy estricta de todos los aspectos operacionales Cualquier sospecha de desviación puede comprometer la credibilidad del estudio</p>
<p>Incremento de muestra a partir de resultados prometedores <i>Sample size increase based on promising interim results</i></p>	<p>Ampliación del tamaño muestral cuando los resultados son prometedores en un análisis intermedio, con el fin de garantizar o aumentar el poder estadístico en el análisis final</p>	<p>Permite maximizar las posibilidades de éxito</p>	<p>Requiere la aplicación de métodos estadísticos bien fundamentados pero poco difundidos. Por este motivo puede provocar conflictos en diferentes ámbitos Requiere desvelar la asignación de tratamientos</p>
<p>Modelos con adaptación de la aleatorización mediante minimización <i>Minimisation</i></p>	<p>Método de asignación dinámica de los pacientes a los grupos de tratamiento para asegurar el equilibrio de la distribución de los factores pronósticos entre los grupos</p>	<p>Mejora la estratificación y aleatorización simple en mantener el equilibrio entre grupos; maximiza la eficiencia en ensayos de tamaño pequeño a moderado</p>	<p>El rendimiento disminuye cuando se introduce más incertidumbre en el método de minimización y si existen interacciones de covariables especificadas. Las pruebas estadísticas estándares no son aplicables y hay debate en cuanto a qué pruebas son las óptimas</p>
<p>Adaptativos más complejos</p>	<p>Permiten adaptar varios elementos del estudio (v.gr., criterios de inclusión, número de grupos / dosis, esquema de aleatorización, variable principal, reglas de finalización, cambio entre no inferioridad y superioridad o viceversa) en base a resultados acumulados. El <i>II/III seamless inferential design</i> parte de un diseño de búsqueda de dosis y luego adapta la asignación en función de la respuesta, descartando grupos del estudio para convertirse sin solución de continuidad en un diseño confirmatorio, en el que se analizan conjuntamente los datos de ambas fases</p>	<p>Muy adecuados para fases iniciales del desarrollo clínico Aumentan la probabilidad de resultados positivos mediante selección de los sujetos con mayor respuesta al tratamiento, eliminación de grupos sin actividad, cambios en la hipótesis del estudio o en la variable principal, etc. Pueden ser utilizados tanto en las fases tempranas como tardías del desarrollo, o bien como combinación de ambas (<i>II/III seamless design</i>) Optimizan el tiempo en el plan de desarrollo</p>	<p>Son muy exigentes operacionalmente La realización de muchas adaptaciones críticas (v.gr., tipo de población y variable principal) en estudios confirmatorios dificulta su aceptación desde el punto de vista regulador En los <i>seamless designs</i> la ausencia de separación entre la fase exploratoria y la confirmatoria impide reaccionar ante imprevistos. Además, la heterogeneidad entre las dos etapas puede ser muy pronunciada, invalidando la combinación de las dos etapas</p>

A pesar de que la metodología estadística es relativamente antigua,⁹⁻¹¹ no se aplicó en investigación clínica hasta que se desarrolló y simplificó el cálculo de las pruebas estadísticas en la década de los 70. Pocock¹² y O'Brien y Fleming¹³ fueron los primeros que describieron métodos aplicables al análisis de datos acumulados en una serie de momentos predeterminados (métodos secuenciales por grupos, MSG). Los MSG no analizan la trayectoria paciente a paciente, sino que realizan análisis intermedios de los datos con grupos de pacientes de tamaño preestablecido, gestionando la multiplicidad de análisis mediante diversos métodos de ajuste del error de tipo I, y permiten finalizar prematuramente un estudio si se rechaza la hipótesis nula a partir de los resultados de los análisis intermedios. Existe la posibilidad de emplear valores de alfa variables en cada análisis intermedio, de manera que en los primeros análisis (cuando hay pocos pacientes) sea difícil rechazar la hipótesis nula, mientras que sea más fácil en análisis sucesivos. Este diseño, propuesto por O'Brien y Fleming, hace difícil la finalización prematura del estudio a no ser que las diferencias sean enormes, lo cual limita el interés práctico de este diseño secuencial. Por otro lado, el método de Pocock, que reparte de manera equitativa el nivel de alfa, sí que facilita la finalización en los primeros análisis intermedios. No obstante, el valor de alfa es mayor en el análisis final con el método de Pocock que con el método de O'Brien y Fleming, por lo que requiere un mayor tamaño muestral para ejecutarlo.

Una variante particular del método secuencial por grupos consiste en la rea-

lización de análisis intermedios para la finalización prematura por «futilidad». Si la probabilidad de que el estudio pueda rechazar la hipótesis nula al final del estudio es razonablemente baja en una inspección predeterminada, entonces puede tomarse la decisión de interrumpir el estudio, evitando la exposición de más sujetos a un experimento que será muy probablemente inconclusivo. Como este análisis no está encaminado a la finalización prematura por demostración de eficacia, sino justamente por la ausencia de esta, no conlleva aumento del error de tipo I; de hecho la base metodológica que subyace se basa en un ajuste del error de tipo II.

Hay numerosos métodos alternativos a los MSG que permiten el análisis de manera más intensiva y con mayor flexibilidad, entre los que destacamos dos por su difusión y aplicación práctica: a) los métodos basados en funciones de gasto de alfa,^{14,15} y b) los métodos basados en la función de verosimilitud, en línea con los desarrollos de los autores iniciales del método secuencial, ampliados por Anscombe¹⁶ y perfeccionados después por Whitehead.^{17,18} Una norma general de aplicabilidad de los estudios secuenciales es que, para ser eficientes, el tratamiento del paciente y el tiempo de observación para la determinación de la variable principal deben ser más breves que el periodo esperado de reclutamiento. De lo contrario, las decisiones se toman cuando ya no suponen ningún ahorro de muestra respecto de los diseños no secuenciales. También es importante poder obtener los resultados a medida que se producen y con el mínimo retraso, especialmente en los estudios puramente

secuenciales; este punto, junto con la incertidumbre acerca del momento exacto de finalización del ensayo y su consecuente complejidad en la planificación, han limitado durante mucho tiempo su aplicación en la práctica. No obstante, su aplicación puede estar indicada o ser útil en los ensayos clínicos con variables de evaluación basadas en el tiempo hasta la presentación, al permitir acortar el seguimiento en el grupo menos eficaz una vez se alcanza un número predeterminado de eventos. La disponibilidad de la captura de datos remota y en tiempo real, mediante aplicaciones seguras y validadas, proporciona una oportunidad excelente para que estos métodos se implementen incluso en estudios multicéntricos multinacionales e intercontinentales. En todos los casos, los métodos secuenciales requieren una estricta predeterminación de los análisis y criterios de decisión, independencia entre los análisis y la gestión y conducción del estudio, y una estricta preservación de la confidencialidad de los resultados para evitar así su influencia sobre la ejecución del ensayo. En estas condiciones, los métodos secuenciales son eficientes y robustos, y suelen requerir menos pacientes y menor tiempo para alcanzar conclusiones, por lo que existe una oportunidad real para aumentar su utilización.

2.2. *Métodos adaptativos*

El concepto clave de los diseños adaptativos es el uso de la información recabada durante la ejecución de un estudio para decidir la modificación de aspectos del mismo, sin comprometer su validez ni su integridad. La extrema flexi-

bilidad de estos métodos y su aparente capacidad para optimizar el diseño y la ejecución de los estudios han llamado la atención tanto de los promotores¹⁹ como de agencias reguladoras,²⁰⁻²³ y han motivado la publicación de buenos artículos de opinión y revisiones recientes.²⁴⁻²⁷

A menudo los parámetros determinantes del diseño de un estudio no se conocen con precisión en el momento de redactar el protocolo, y es preciso asumir una serie de premisas que posteriormente pueden resultar, o no, acertadas. Los diseños adaptativos prevén, de forma predeterminada, la modificación de uno o varios elementos del diseño (v.gr., tamaño de la muestra, razón de aleatorización, grupos comparados, criterios de inclusión y exclusión de pacientes, cambios en la variable principal, ajuste de las hipótesis estadísticas —por ejemplo de no-inferioridad o superioridad—, o combinación de ensayos o fases de tratamiento) a partir de los resultados de análisis intermedios predeterminados, con objeto de ir corrigiendo o refinando algunos aspectos del diseño sin aumentar el error de tipo I.

El término adaptativo engloba multitud de procedimientos, algunos muy simples y ampliamente utilizados y otros de mayor complejidad; la denominación de diseños adaptativos suele reservarse para estos últimos.

En relación a estos diseños, hoy en día se puede afirmar que ya hay un buen nivel de consenso, que la metodología estadística que soporta estos diseños está ampliamente fundamentada,²⁸⁻³⁰ y que la preservación del error de tipo I y la metodología inferencial subyacente están, en general, considerados resuel-

tos.³¹⁻³⁶ Por tanto, estos puntos per se no significan ninguna limitación para su implementación práctica.

Un ejemplo de las adaptaciones simples más ampliamente utilizadas es la reestimación del tamaño muestral. En el cálculo inicial se utilizan estimaciones de variabilidad para los parámetros principales del estudio basadas en datos previos o referidos en la literatura, pero estas estimaciones pueden resultar inexactas. Se pueden adaptar los cálculos mediante métodos basados en estimaciones ciegas de la variabilidad o bien empleando métodos que requieren el desvelamiento de la asignación. Para una información más detallada se pueden consultar algunas publicaciones excelentes que revisan este tema.³⁷⁻⁴⁰ La inspección de la variabilidad de una parte de los datos recogidos en el propio estudio, manteniendo el enmascaramiento de la asignación a tratamientos, puede emplearse para verificar la validez del cálculo inicial. Si la variabilidad propia del estudio resulta ser sustancialmente distinta, la nueva estimación se emplea para enmendar en consecuencia el protocolo.⁴¹ En general, los métodos que no requieren el desvelamiento de la asignación son claramente preferibles, puesto que son de implementación relativamente fácil, y no comportan riesgo de generar sesgos operacionales. Recientemente se están aplicando también métodos que proponen la ampliación del tamaño muestral cuando los resultados son prometedores en un análisis intermedio, con el fin de garantizar o aumentar el poder estadístico en el análisis final.^{42,43}

La modificación de los criterios de inclusión y exclusión del estudio duran-

te su conducción es otra modificación frecuente, que puede englobarse en el concepto de adaptación. Si este cambio es sustancial, suele ser necesario hacer una modificación adicional al protocolo para incluir la estratificación del análisis en función de la utilización de los criterios iniciales o finales. Otro cambio habitual es el de la interrupción de un grupo de tratamiento o nivel de dosis como consecuencia de nueva información sobre la seguridad de ese tratamiento o dosis; esta información puede proceder del propio estudio o de estudios llevados a cabo de manera simultánea.

Hablamos de estudios adaptativos cuando en el protocolo ya se describen de entrada distintos cambios condicionados a los resultados de análisis intermedios predeterminados: su implementación está prevista a partir de reglas de modificación, y se prevén los sistemas estadísticos y logísticos para que no se comprometan ni el enmascaramiento del estudio, ni la integridad de los resultados, ni la protección frente a los errores estadísticos. El tipo de modificaciones sustanciales del diseño de estos estudios adaptativos con frecuencia implica la interrupción de grupos de tratamiento o niveles de dosis ineficaces. Los pacientes que se encontraban en los grupos que se interrumpen pueden o no finalizar el tratamiento; si continúan en tratamiento, pueden pasar a fases de seguimiento no comparativas en condiciones abiertas (no enmascaradas), o bien derivarse a otros niveles de dosis o grupos de tratamiento. La entrada de nuevos pacientes al estudio puede requerir un cambio en los procedimientos de asignación, que puede pasar a ser desequilibrada, aumentando

la muestra en uno o varios grupos cuyos resultados son aparentemente mejores, y manteniendo otros grupos con el mínimo tamaño requerido (por ejemplo, para evitar un aumento del número de sujetos en el grupo de placebo o de control). En este último caso los aspectos prácticos pueden ser muy complejos, ya que debe realizarse una reasignación de los pacientes y una modificación de las medicaciones enmascaradas, manteniendo las garantías de trazabilidad y sin romper el enmascaramiento del ensayo. Esto puede requerir la fabricación, el acondicionamiento y la distribución de lotes de medicación adicionales en previsión de distintos escenarios, y aumentar las necesidades de medicación para el estudio. Cuando la medicación es escasa o muy costosa, puede ser necesario recuperar y reacondicionar las unidades para adaptarlas a la nueva situación, lo cual puede ser extremadamente complejo, y necesitar estructuras de soporte y controles de calidad muy importantes.

2.3. *Aproximaciones bayesianas*

El concepto clave de los diseños bayesianos es la utilización del conocimiento previo en el diseño y el análisis de la información de un nuevo estudio; podría decirse que la pregunta clave que se plantean es cómo se modifica el conocimiento previo ante la aparición de una nueva evidencia. Así, la estadística bayesiana se inicia a partir de una creencia o enunciado previo, consistente en una distribución de probabilidad, que se actualiza con los resultados observados en nuevos experimentos para integrar un enunciado posterior, consistente tam-

bién en una distribución de probabilidad, de modo que todos los resultados disponibles se tienen en cuenta. Los diferentes tipos de evidencias se pueden combinar en el modelo general de probabilidad, o se incluyen en el mismo a partir de la distribución de probabilidad previa. En la actualidad hay buenos textos de revisión de estos métodos aplicados a los ensayos clínicos.⁴⁴⁻⁴⁸

Por contraposición a los métodos bayesianos, los métodos clásicos o frecuentistas establecen inferencias basándose únicamente en la evidencia observada, sin tener en cuenta formalmente la información previa, a excepción de la estimación de la variabilidad y las diferencias esperadas para el cálculo del tamaño de la muestra. Los métodos bayesianos, mediante la aplicación de la distribución posterior a la curva de poder del estudio, permiten obtener una estimación más precisa de las diferencias entre los tratamientos, así como la distribución de probabilidad para un intervalo de diferencias posibles. Esto permite una estimación más precisa y realista del cálculo del poder de un determinado tamaño de muestra.

En ensayos clínicos de prueba de concepto los métodos bayesianos pueden maximizar la utilidad de la escasa información disponible. Estos estudios suelen incluir pocos pacientes que se observan durante un periodo relativamente corto, con objeto de obtener señales de seguridad y estimaciones preliminares de eficacia. Estas evaluaciones, a menudo, se hacen a tiempos repetidos y empleando escalas continuas, pero con frecuencia los análisis de las variables principales las categorizan en variables dicotómicas

de respuesta positiva o negativa en un momento único del tiempo, con la consiguiente pérdida de información. Asimismo, debido al pequeño número de sujetos participantes, solo pueden detectarse con un poder estadístico razonable aquellos problemas de seguridad que ocurren en un porcentaje relativamente elevado de pacientes, y con señales de eficacia relativamente intensas. Los métodos bayesianos pueden mejorar la detección de señales de seguridad y eficacia utilizando datos e información externos al ensayo. Además, la integración de los datos obtenidos en estudios de prueba de concepto en los estudios de fases más avanzadas puede maximizar la utilidad de esta información mediante métodos de modelado bayesiano.^{49,50}

Lógicamente los diseños bayesianos se han empleado y se emplean con relativa frecuencia en estudios de fases I y II, ya que estos estudios son, por su naturaleza exploratoria, ideales para métodos secuenciales y adaptativos, y estas características son los puntos fuertes de la metodología bayesiana. Así, en la fase I, los diseños, las decisiones de escalado y los incrementos de dosis se basan en las expectativas a partir de la información preclínica y de los datos observados en niveles de dosis previos. En estudios de fase II es importante poder interrumpir de manera precoz algún nivel de dosis o incluso el ensayo entero si la eficacia del producto no es la esperada. También es importante poder implementar análisis intermedios no previstos, o cambiar la proporción de pacientes asignados aleatoriamente a cada tratamiento de manera adaptativa.⁵¹ Los diseños bayesianos se han aplicado

extensamente en oncología, y son especialmente apropiados para el estudio de las enfermedades raras, en las que las dificultades de reclutamiento hacen especialmente importante la utilización de toda la información disponible para la evaluación de los tratamientos.

2.4. Modelización y simulación

La aproximación etiológica a los procesos subyacentes a la enfermedad mediante el estudio de los cambios fisiológicos, bioquímicos y genéticos en los estados patológicos permite el diseño racional de nuevas estrategias farmacoterapéuticas y el planteamiento de los estudios clínicos desde un punto de vista mecanístico. El detallado conocimiento de la farmacodinamia de los compuestos y de su farmacocinética permite establecer relaciones entre la exposición sistémica y el lugar de acción con los parámetros biológicos o biomarcadores, y construir modelos de actividad que sirven para guiar el diseño de los experimentos clínicos, tanto en la selección de poblaciones como de las variables de medida, los momentos de evaluación y la elección de las dosis a estudiar.⁵²

Una aplicación de las simulaciones en las fases exploratorias es la construcción de modelos virtuales de enfermedad a partir de la información biológica y clínica disponible (experimental o publicada) sobre la fisiopatología, la farmacodinámica y los efectos clínicos de diversos tratamientos conocidos. En los modelos virtuales resultantes se introducen parámetros de los fármacos candidatos, de tipo farmacocinético y farmacodinámico, para simular sus posibles efectos

clínicos y así guiar la selección de moléculas candidatas a desarrollar, orientar sobre su dosis y posología óptimas, y sobre la mejor manera de medir su eficacia y seguridad.^{53,54} En fases posteriores, es práctica habitual que la decisión de selección de dosis y el diseño de los estudios confirmatorios de eficacia se apoye en parte en modelos de farmacocinética y farmacodinámica construidos con los datos del desarrollo exploratorio.

En pediatría, la obtención de datos sobre dosis, exposición y respuesta en adultos suele preceder a los estudios pediátricos, de modo que el modelado a partir de estos suele emplearse para seleccionar las pautas a estudiar en niños, o los tiempos de muestreo o evaluación, optimizando así el diseño de los estudios.⁵⁵

Por último, la modelización estadística se puede utilizar para seleccionar los diseños más idóneos para un estudio determinado, las poblaciones diana que más pueden beneficiarse del nuevo tratamiento y para establecer la estrategia de desarrollo más eficiente.^{56,57}

2.5. Diseños enriquecidos

La selección de poblaciones especialmente sensibles al efecto de un tratamiento es una manera de maximizar las posibilidades de un estudio para evaluar apropiadamente la eficacia. Así, en algunas situaciones clínicas, una intervención puede ser más eficaz en pacientes con determinadas comorbilidades o con determinadas características fisiopatológicas, de manera que el estudio puede *enriquecerse* con la inclusión selectiva de esta subpoblación más susceptible, aumentando sus posibilidades

de evaluar adecuadamente el efecto del tratamiento. Además, los diseños enriquecidos pueden suponer una solución útil en situaciones en las que el reclutamiento es difícil, por ser cuestionable la equivalencia clínica (*clinical equipoise*) o cuando la ausencia de alternativas eficaces en patologías graves hace que los pacientes sean reticentes a participar en diseños con asignación aleatoria comparativos con placebo.^{58,59} Estos diseños se pueden también combinar con otros para optimizar su ejecución.⁶⁰

Una forma de llevar a cabo esta estrategia consiste en la selección de poblaciones a partir de una fase inicial de inducción de respuesta en la que todos los sujetos reciben el tratamiento en estudio, y solo los que han respondido se distribuyen aleatoriamente a una fase ciega (enmascarada) de retirada del tratamiento mediante la asignación a tratamiento activo o placebo. Aquellos que durante la fase a doble ciego pierden la respuesta al tratamiento pueden, a su vez, entrar en una fase de extensión abierta (no enmascarada) en la que reciben el tratamiento activo. Este diseño permite gestionar los conflictos éticos y las cuestiones prácticas de accesibilidad (cuando existen datos preliminares que sugieren una actividad relevante de un compuesto nuevo en indicaciones con nulas o escasas alternativas), especialmente en enfermedades graves y poco prevalentes, a la vez que proporciona elementos de causalidad suficientemente sólidos como para poder establecer la eficacia de un tratamiento, e incluso servir de soporte a una decisión reguladora sobre la utilidad del nuevo producto.^{61,62} Su aplicabilidad, no obstante,

viene condicionada por la ausencia de efectos permanentes de la medicación que se arrastren de manera posterior a su retirada (efecto *carry-over*), y a que la enfermedad de base no tenga oscilaciones importantes en su actividad que confunda las evaluaciones durante el estudio. Tampoco es un diseño sin problemas éticos, pues en su mismo concepto implica la retirada temporal de un tratamiento beneficioso en aquellos pacientes que han alcanzado una respuesta clínica durante la fase de inducción.

Otra forma de diseño enriquecido es la modificación adaptativa de los criterios de selección de pacientes, restringiendo la entrada a aquellos pacientes que con más probabilidad se beneficiarán de los tratamientos,⁶³ a medida que los datos acumulados van mostrando que un subgrupo de pacientes tiene mayor probabilidad de respuesta.

A medida que se van conociendo mejor las causas de la variabilidad interindividual en la respuesta biológica, resulta cada vez más viable la definición a priori de subgrupos de pacientes en los que los tratamientos tienen mayor o menor probabilidad de resultar eficaces. Así, la asociación de perfiles biológicos específicos con factores pronósticos de respuesta, a partir de características clínicas, biomarcadores, expresión génica o genotipos, es cada vez más habitual en los trabajos científicos de tipo exploratorio. Su aplicación al desarrollo clínico es una consecuencia obvia, de manera que la selección de poblaciones con mayores probabilidades de obtener un beneficio del tratamiento en base a criterios de inclusión específicos (v.gr., la expresión de mutaciones tumorales

en oncología) es una práctica cada vez más extendida.

Cualquier aplicación de las estrategias de enriquecimiento requiere una planificación detallada, con simulación de escenarios que permitan ponderar las ventajas del incremento en el poder del estudio y el ahorro en el tamaño de la muestra, con las mayores dificultades de reclutamiento que suponen los criterios más acotados y con las limitaciones a la aplicabilidad de los resultados que supondrá la restricción de los criterios de inclusión en el estudio.

2.6. *Alternativas a los ensayos controlados paralelos*

En general, los estudios pueden contar con controles internos o externos, que pueden ser históricos o concurrentes. Existen abundantes sesgos y problemas asociados al uso de controles externos, por lo que la situación ideal para evaluar la eficacia y la seguridad de un nuevo producto es poder contar siempre con un grupo de control interno. No obstante, pueden darse circunstancias excepcionales en las que sea inviable contar con controles internos concurrentes, como, por ejemplo, en las enfermedades por déficit enzimático en las que la utilización de un placebo puede ser éticamente inadmisibles. En estos casos puede ser aceptable utilizar controles externos históricos, por ejemplo, procedentes de registros de pacientes tratados con el estándar de tratamiento hasta la fecha, o sin tratamiento alguno. Para que estos controles puedan ser de utilidad, es preciso que se conozca bien el curso natural de la enfermedad en ausencia de trata-

miento, y que los datos de los registros sean suficientemente detallados para poder establecer comparaciones apropiadas. Por esto resulta importante que se detecte en fases precoces del desarrollo de un nuevo medicamento si va a ser necesario emplear controles históricos, de modo que se pueda prever la creación de una base de datos adecuada que esté disponible para cuando sea precisa.

La utilización de los datos en situación basal del paciente como referencia no constituye, en realidad, un grupo de control, y, a efectos prácticos, debe considerarse como una descripción de caso clínico o serie de casos no comparativa. Por último, en algunas situaciones concretas la realización de ensayos con asignación aleatoria con un único sujeto (N de 1) puede ser útil para decidir la mejor estrategia terapéutica para un paciente determinado;⁶⁴⁻⁶⁶ los datos acumulados de la experiencia de N de 1 en una serie de pacientes solo raramente y en circunstancias muy especiales puede emplearse como fuente de evidencia sobre la actividad de un tratamiento.

Los diseños factoriales permiten evaluar simultáneamente combinaciones de dos o más intervenciones o tratamientos para determinar si la combinación aporta ventajas respecto a las intervenciones individuales, obteniendo una evaluación de los efectos de dosis y superficie-respuesta. Para ello, este diseño incluye los tratamientos estudiados por separado y todas las posibles combinaciones de los mismos, resultando en un número de grupos igual al factorial del número de tratamientos a combinar; de este modo se estiman los efectos y las potenciales interacciones (sinergias, adiciones o anta-

gonismos) de los tratamientos estudiados. Se considera el diseño clave en el desarrollo de combinaciones a dosis fijas.⁴

En los diseños de tratamiento añadido (*add-on*) se estudian pacientes que ya están recibiendo tratamientos activos frente a la enfermedad en estudio, a los cuales se asigna de forma aleatoria a recibir adicionalmente un placebo o un tratamiento experimental con mecanismo de acción distinto a los que ya están recibiendo.⁶⁷ Esta alternativa facilita la participación en ensayos clínicos de pacientes con patologías graves para las que existen alternativas de tratamiento activo, aunque con una eficacia incompleta.

En los ensayos de aleatorización del abandono de la medicación en estudio (*randomised withdrawal*), todos los pacientes inician tratamiento con el tratamiento experimental, y aquellos que responden al mismo se asignan aleatoriamente a continuar recibiendo el mismo tratamiento o un placebo, para evaluar el mantenimiento del efecto una vez este se ha obtenido. Una variante de los diseños de retirada consiste en la aleatorización a conmutar tratamientos (*randomized switch designs*), es decir, a cambiar a otro tratamiento activo o bien placebo una vez se ha alcanzado un efecto determinado.⁶⁸

Se debe tener en cuenta, al evaluar los resultados de estos estudios, que son solo aplicables a la subpoblación que obtiene inicialmente el efecto buscado. Por lo tanto, si la población que inicialmente se trata y el subgrupo que responde favorablemente y finalmente puede ser seleccionado de forma aleatoria a la retirada son muy distintos, la extrapolación a la población general de los

resultados puede ser muy difícil. Para una evaluación no sesgada del efecto de mantenimiento, se recomienda que el diseño sea doble-ciego. En otros diseños en los que el abandono no se decide de manera aleatoria, la comparabilidad de los grupos generados, y por tanto las conclusiones del ensayo, están formalmente comprometidas. En el diseño de estos estudios deben considerarse no solo razonamientos de índole científica, sino también de tipo ético, que deben valorar cuidadosamente la *clinical equipoise* y la pertinencia de retirar un tratamiento que ya se ha mostrado eficaz.

Los diseños de abandono o escape temprano (*early-escape design*) permiten acotar la duración de la exposición a tratamientos que no son suficientemente eficaces o a placebo, y pueden ser útiles cuando éticamente no puede descartarse que la asignación a determinados tratamientos tenga un impacto negativo en la salud de los pacientes.

Los diseños basados en análisis de decisión (*decision analysis-based design*),⁶⁹ basados en modelación y simulación, se utilizan principalmente para la evaluación complementaria de eficacia y seguridad de manera combinada, y son útiles para las agencias de evaluación de tecnologías. También se usan para identificar las variables, poblaciones y escenarios más adecuados para ejecutar determinado estudio. Así se predefinen variables de eficacia y seguridad de acuerdo a unas probabilidades plausibles, cuyos pesos específicos se asignan de acuerdo a preferencias.⁷⁰ Otros diseños que contemplan simulaciones estadísticas para la búsqueda de las variables y/o muestras ideales son la ordenación jerárquica y se-

lección (*ranking and selection*) y los procedimientos de comparaciones múltiples (*multiple comparison procedures*).^{71,72}

En algunas situaciones, debido a determinadas características de alguna enfermedad y a su pronóstico, distribuir aleatoriamente a los pacientes puede resultar operativamente difícil o éticamente cuestionable. Existen métodos que permiten diseñar y analizar estudios en los que la asignación está guiada por el riesgo subyacente de los pacientes (*risk-based allocation*). No obstante, al asociar factores pronósticos a la asignación de las intervenciones se produce una confusión inevitable, inherente al método, de manera que las técnicas estadísticas estándar no son válidas para el análisis y los resultados obtenidos deben evaluarse con cautela.⁷³⁻⁷⁵

En algunas situaciones, en vez de asignar aleatoriamente el tipo de tratamiento, se puede asignar aleatoriamente el momento en el que se inicia la intervención (*staggered start*), para evaluar cómo el momento del inicio del tratamiento puede influir en su eficacia en una determinada enfermedad.^{76,77} Este método se ha propuesto para estudios de modificación del curso natural de enfermedades como, por ejemplo, las demencias tipo Alzheimer,⁷⁸⁻⁸¹ si bien parece más aplicable a estudios de tipo exploratorio que a estudios confirmatorios. Una variante de este método consiste en distribuir aleatoriamente el momento de la retirada del tratamiento o intervención (*delayed withdrawal*).⁸²

Finalmente, los métodos basados en la evaluación del tiempo hasta el inicio del efecto (*time-to-onset*), de manera similar a la evaluación de tiempo has-

ta supervivencia, se han propuesto por determinados autores como métodos adecuados para evaluar la velocidad en alcanzar un efecto, cuando esta sea realmente considerada como un elemento clave de la eficacia de un tratamiento.^{83,84}

3. LA NECESIDAD DE CAMBIOS

En los últimos años, la rápida evolución del entorno de la investigación clínica ha ido generando distintas circunstancias que han estimulado el desarrollo y la aplicación práctica de las distintas alternativas a los diseños clásicos.

3.1. *Aumentar la eficiencia del proceso de investigación*

La presión por disponer de nuevos medicamentos eficaces en patologías con necesidades médicas no cubiertas resulta acuciante, y cualquier método que permita acortar la espera representa una oportunidad para pacientes con enfermedades sin alternativas terapéuticas eficaces.

Por una parte, la implementación de las nuevas tecnologías de la información en el proceso de captura de datos durante los estudios ha permitido acortar drásticamente los tiempos de integración, gestión y depuración de datos, eliminando prácticamente todos los obstáculos a la implementación práctica de las aproximaciones secuenciales. El mero hecho de disponer de la información en tiempo casi real plantea de manera natural su aprovechamiento en aras de la eficiencia de la investigación.

De manera similar, la disponibilidad de sistemas remotos, que permiten

la asignación centralizada de pacientes en estudios multicéntricos a nivel mundial, abre la posibilidad de aplicar en la práctica conceptos como la asignación adaptativa, con un dinamismo que no resultaba viable antes de la era de las telecomunicaciones e Internet.

Otro aspecto a considerar es el requerimiento progresivo para acelerar la ejecución de los ensayos. Por una parte, el constante progreso en terapéutica motiva la continua aparición de nuevos tratamientos en el mercado. Durante la realización de un estudio en patologías crónicas o con periodos de observación prolongados no es infrecuente que se produzcan cambios en la práctica clínica habitual. Estos cambios pueden afectar el reclutamiento de pacientes, que puede resultar inviable, tanto desde el punto de vista ético como práctico, por ejemplo, en estudios controlados con placebo o que empleen como referencia un tratamiento que, a raíz de las novedades, no se considere ya la mejor opción disponible. Asimismo, un estudio muy dilatado en el tiempo puede obtener conclusiones basadas en estándares de tratamiento obsoletos, o puede sufrir problemas de heterogeneidad derivados de los cambios en el manejo terapéutico de los pacientes, modificaciones en el riesgo en situación basal de la población estudiada, o mejorías del pronóstico de la enfermedad a lo largo de la conducción del estudio. Estos últimos pueden afectar asimismo al poder del ensayo, puesto que la estimación del tamaño de la muestra se realizó de manera previa al inicio del estudio y basándose en datos que ya no están vigentes.

En otro orden de cosas, los costes derivados de la ejecución de un ensayo son proporcionales a la duración del mismo, de manera que conseguir diseños eficientes en cuanto a la optimización del número de pacientes necesarios y la duración de la ejecución siempre ha sido una prioridad de los financiadores de la investigación clínica, más marcada si cabe en los últimos años. Las autoridades reguladoras han ido aumentando su exigencia respecto a los requerimientos para la comercialización de nuevas entidades moleculares, y es preciso realizar programas de investigación cada vez más costosos y prolongados; mientras, los periodos de exclusividad de patente no se han modificado sustancialmente. Por ello las empresas farmacéuticas tratan de maximizar el retorno de inversión mediante estrategias que permitan reducir los costes y acelerar el desarrollo clínico de los nuevos productos, en un entorno comercial muy agresivo.

Por último, y con motivo de la globalización industrial y la armonización de requerimientos reguladores, se ha estimulado de manera importante la realización de estudios internacionales e intercontinentales (véase capítulo 3). La ejecución de un protocolo en entornos geográficos distintos y distantes supone retos de tipo financiero y logístico, pero también metodológicos en cuanto a la gestión de la heterogeneidad de los datos.

3.2. *Fusión de programas exploratorios y confirmatorios*

Los diseños adaptativos sin interrupciones o *seamless* combinan en un único

ensayo los objetivos tradicionalmente abordados en ensayos independientes, evitando los lapsos de tiempo entre fases de desarrollo debidos a la evaluación de resultados de los estudios de una fase antes de iniciar el diseño y la puesta en marcha de la siguiente.^{85,86} Estos estudios son tanto de fase I/II, en los que se unen los objetivos de tolerabilidad de la fase I a los de exploración de actividad en fase II, como de fase II/III, en los que la selección de dosis se enlaza sin solución de continuidad con la fase confirmatoria de la eficacia de las dosis seleccionadas. En este último ejemplo, la primera etapa del ensayo podría ser similar a un ensayo de fase tardía II, con un grupo de control y varios grupos de tratamiento (por ejemplo, diferentes niveles de dosis), con uno o varios análisis intermedios en los que se toman decisiones sobre la continuidad de los niveles de dosis menos eficaces. Se realiza la adaptación conforme al protocolo, y uno o más de los grupos de tratamiento activo continúan en la segunda fase junto con el grupo de control. El análisis final comparativo del o los grupos seleccionados respecto del grupo de control utiliza los datos de ambas etapas del ensayo. Esto permite una reducción en la duración del desarrollo clínico al eliminar el tiempo de transición entre fases, una mayor eficiencia en la utilización de los datos de las dos etapas del ensayo, con un menor número final de pacientes incluidos para obtener la conclusión de eficacia, y un seguimiento de seguridad prolongado que se obtiene antes en el desarrollo, puesto que se inicia el seguimiento a largo plazo ya en la primera fase.⁸⁷

3.3. *Ensayos clínicos en poblaciones pequeñas*

El diseño clásico de grupos paralelos requiere una estimación precisa del tamaño de la muestra al objeto de garantizar un poder suficiente para concluir de manera significativa. Cuando una característica determinante de la enfermedad a estudiar es su escasa incidencia o prevalencia, estas estimaciones determinan que es necesaria una muestra mayor de la disponible, incluso en aproximaciones multicéntricas; esto aboca a estudios con poder insuficiente o, directamente, a concluir que el ensayo clínico es inviable.^{88,89}

En poblaciones pequeñas, además, el riesgo de sesgos debidos a errores no sistemáticos puede ser mucho más relevante que cuando se emplean muestras amplias, en las que estos efectos se diluyen.⁹⁰ No obstante, estas poblaciones precisan disponer, tanto o más que las poblaciones prevalentes, de evidencias sobre sus a menudo escasas opciones terapéuticas. En poblaciones pequeñas, por lo tanto, es especialmente necesario aplicar metodologías que permitan maximizar la utilidad y la consistencia de la información disponible.⁹¹

Un caso particular en el que es especialmente acuciante aumentar la eficiencia de los diseños de los ensayos clínicos son las enfermedades raras, es decir, aquellas indicaciones en las que, por la rareza de la enfermedad y la escasez de pacientes, la investigación de nuevos tratamientos no representa, a priori, un interés comercial sustancial en términos convencionales. Mientras en las enfermedades muy prevalentes uno de los objetivos predominantes para

disponer de diseños eficientes es aumentar la competitividad financiera, en las enfermedades raras la necesidad de estos diseños está guiada por la dificultad para obtener pruebas de eficacia en poblaciones extremadamente escasas. En estas situaciones puede no haber tratamientos de referencia, o es posible que estos no dispongan de datos objetivos de eficacia. También es frecuente que la comparación con placebo se valore como éticamente inviable. En general, la urgencia del desarrollo clínico del nuevo medicamento viene determinada por la gravedad de la patología y el vacío terapéutico.

Los pacientes pediátricos representan otra situación especial, similar a la anterior en la escasez, pero también en que las limitaciones éticas a la participación de sujetos dependientes en estudios experimentales aconsejan maximizar la información obtenida a partir del mínimo número posible de sujetos, por lo que la eficiencia de los diseños es imprescindible.⁹²

En estas circunstancias es importante evitar que las limitaciones de acceso a la población afectada para la realización de estudios paralelos convencionales deriven en la realización de ensayos clínicos sin un mínimo de protección frente a los errores estadísticos, cuyo impacto ético podría ser aún peor que la ausencia de evidencias, especialmente cuando se trata de, por ejemplo, poblaciones pediátricas con enfermedades graves, en las que la especial vulnerabilidad de los sujetos hace más difícil la ecuanimidad de los investigadores y las familias respecto de sus expectativas sobre los tratamientos. Por estos motivos resulta especialmen-

te importante disponer de metodologías robustas frente a sesgos, y de estrategias que hagan más aceptable para los pacientes la participación en ensayos clínicos. La asignación desequilibrada a los grupos experimental y control, y la inclusión en fases de tratamiento abiertas (no enmascaradas) con las alternativas experimentales una vez finalizada la fase de comparación a doble ciego, pueden ayudar a mejorar la aceptabilidad de la participación en los estudios.^{93,94} Puede ser especialmente útil la realización de ensayos de tipo secuencial, ya que permiten que el reclutamiento cese tan pronto como se obtienen datos concluyentes, minimizando la exposición experimental de sujetos, utilizando de manera eficiente los datos disponibles y evitando la demora en la consecución de resultados. También las aproximaciones bayesianas son buenas alternativas, ya que permiten la integración de la información conocida con los datos que se van obteniendo durante el estudio, aprovechando toda la evidencia disponible.⁹¹

4. IMPORTANCIA DE VALIDAR LA SELECCIÓN DEL DISEÑO

En una revisión reciente⁹⁵ sobre los factores asociados con la decisión final de la Agencia Europea de Medicamentos (EMA), se analizaron 135 solicitudes de autorización de comercialización de nuevos medicamentos recibidas entre los años de 1995 a 1999; el factor que mejor predecía un resultado negativo de la evaluación era la ausencia de ensayos clínicos con asignación aleatoria que establecieran la eficacia clínica del producto apropiadamente. En cambio,

en las 188 solicitudes de evaluación recibidas entre los años 2004 y 2007, el cumplimiento con las indicaciones del consejo científico de la EMA fue uno de los dos factores predictivos de éxito, junto con el tamaño de la empresa promotora. Este último factor se relacionaba con la disponibilidad de recursos y experiencia en el desarrollo clínico de nuevos medicamentos, y en el proceso de obtención de autorizaciones. De dicho estudio se deriva como conclusión principal la conveniencia de consultar a las autoridades reguladoras la aceptabilidad de los diseños clínicos destinados a proporcionar soporte clave en solicitudes de autorización con medicamentos, mediante solicitudes de consejo científico en las fases iniciales del desarrollo clínico, en los momentos de decisión críticos y, especialmente, en el caso de empresas pequeñas, con menos recursos financieros y científicos y con menos experiencia en el proceso, y en el de aquellas dedicadas al desarrollo de productos en enfermedades raras. En este último caso porque las situaciones clínicas son menos conocidas y existen menos referencias para establecer la adecuación de la metodología seleccionada.

5. CONCLUSIONES

La rápida evolución del entorno de la investigación clínica ha estimulado el desarrollo y la aplicación práctica de alternativas al diseño clásico de asignación aleatoria a grupos paralelos, especialmente de métodos adaptativos y flexibles, capaces de responder a las necesidades del nuevo paradigma. La presión

por disponer de nuevos medicamentos eficaces en patologías con necesidades médicas no cubiertas resulta acuciante, y cualquier método que permita acortar la espera representa aumentar las oportunidades para pacientes con enfermedades sin alternativas terapéuticas eficaces.

El auge de las tecnologías de la información en el proceso de captura de datos durante los estudios ha permitido acortar drásticamente los tiempos de integración, gestión y depuración de datos, eliminando prácticamente todos los obstáculos a la implementación práctica de las aproximaciones secuenciales, de modo que el hecho de disponer de la información en tiempo real plantea de manera natural su aprovechamiento, en aras de la eficiencia de la investigación. Igualmente, las telecomunicaciones han hecho posible la implantación de sistemas de asignación y aleatorización centralizados y automatizados, que permiten llevar a cabo adaptaciones de las pautas de asignación durante el estudio manteniendo la integridad del mismo.

La necesidad de modificar el diseño de un ensayo clínico de fase III durante su ejecución podría percibirse como una evidente contradicción con su ca-

rácter confirmatorio dentro del plan de desarrollo de un nuevo producto, raramente aceptable sin más justificación. Pero los diseños adaptativos no deben considerarse a priori como medios para reducir la rigurosidad de la planificación de los ensayos clínicos, sino como herramientas para gestionar situaciones experimentales complejas. En todos los casos el tipo de modificación (cambio de tamaño de la muestra, interrupción de los grupos de tratamiento, etc) debe ser previsto, descrito y justificado en el protocolo del estudio.

Aunque no existen métodos específicos para el diseño, la realización o el análisis de los ensayos clínicos en poblaciones pequeñas, los métodos alternativos pueden aumentar su viabilidad y eficiencia. Por último, independientemente de la metodología empleada, debe mantenerse un equilibrio apropiado entre las necesidades de eficiencia estadística y las de obtener resultados clínicamente relevantes e interpretables, siendo estas últimas las más importantes.

REFERENCIAS

1. Bakke OM, Carné X, García Alonso F. Ensayos clínicos con medicamentos. Fundamentos básicos, metodología y práctica. Barcelona: Doyma; 1994.
2. Pocock SJ. Clinical trials - A practical approach. West Sussex: John Wiley & Sons; 1983.
3. Senn S. Cross-over Trials in clinical research. Chichester: John Wiley; 1993.
4. Chow SC, Liu JP. Design and analysis of clinical trials: concepts and methodologies. 2.^a ed. John Wiley & Sons, Inc: Hoboken, New Jersey; 2004.
5. ICH E 9 Statistical principles for clinical trials. CPMP/ICH/363/96. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002928.pdf. Consultado el 16 de julio de 2013.
6. Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. *J Clin Oncol*. 2007; 25: 4982-6.
7. Koch A. Confirmatory clinical trials with an adaptive design. *Biomet J*. 2006; 48: 574-85.
8. Hung HMJ, O'Neil RT, Wang SJ, Lawrence J. A regulatory view on adaptive/flexible clinical trial design. *Biomet J*. 2006; 48: 565-73.
9. Barnard GA. Sequential test in industrial statistics. *J Roy Statist Soc*. 1946; supl. 8: S1-26.
10. Wald A. Sequential analysis. Nueva York: Wiley; 1947.
11. Armitage P. Sequential methods in clinical trials. *Am J Public Health*. 1958; 48: 1395-402.
12. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika*. 1977; 64: 191-9.
13. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979; 35: 549-56
14. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika*. 1983; 70: 659-63.
15. Kim K, DeMets DL. Design and analysis of group sequential tests based on the type I error spending rate function. *Biometrika*. 1987; 74: 149-54.
16. Anscombe FJ. Sequential medical trials. *J Amer Statist Assoc*. 1963; 58: 365-83.
17. Whitehead J, Jones D. The analysis of sequential clinical trials. *Biometrika*. 1979; 66: 443-52.
18. Whitehead J. The design and analysis of sequential clinical trials. Chichester: Ellis Horwood; 1992.
19. Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, Pinheiro J. PhRMA Working Group. Adaptive designs in clinical drug development. An executive summary of the PhRMA working group. *J Biopharm Stat*. 2006; 16(3): 275-83; discusión 285-91, 293-8, 311-2.
20. Gallo P, Anderson K, Chuang-Stein C, et al. Viewpoints on the FDA draft adaptive designs guidance from the pharma working group. *J Biopharm Stat*. 2010; 20: 1115-24.
21. European Medicines Agency, 2009. Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf. Consultado el 1 de octubre de 2012.
22. Food and Drug Administration, 2010. Adaptive design clinical trials for drugs and biologics (draft guidance). Disponi-

- ble en: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm201790.pdf>. Consultado el 1 de octubre de 2012.
23. Wang SJ, Hung HM, O'Neill R. Regulatory perspectives on multiplicity in adaptive design clinical trials throughout a drug development program. *J Biopharm Stat.* 2011; 21: 846-59.
 24. Chang M, Chow SC, Pong A. Adaptive design in clinical research: issues, opportunities, and recommendations. *J Biopharm Stat.* 2006; 16: 299-309; discusión 311-2.
 25. Berry DA. Adaptive clinical trials: the promise and the caution. *J Clin Oncol.* 2011; 29: 606-9.
 26. Bauer P, Brannath W. The advantages and disadvantages of adaptive designs for clinical trials. *Drug Discov. Today.* 2004; 9: 351-7.
 27. van der Graaf R, Roes KC, van Delden JJ. Adaptive trials in clinical research: scientific and ethical issues to consider. *JAMA.* 2012; 307: 2379-80.
 28. Liu Q, Proschan MA, Pledger GW. A unified theory of two-stage adaptive designs. *J Am Stat Assoc.* 2002; 97: 1034-41.
 29. Brannath W, Posch M, Bauer P. Recursive combination tests. *J Am Stat Assoc.* 2002; 97: 236-44.
 30. Berry DA. Bayesian statistics and the efficiency and ethics of clinical trials. *Stat Sci.* 2004; 19: 175-87.
 31. Bauer P, Köhne K. Evaluation of experiments with adaptive interim analyses. *Biometrics.* 1994; 50: 1029-41. Corrección: 1996; 52: 380.
 32. Proschan MA, Hunsberger SA. Designed extension of studies based on conditional power. *Biometrics.* 1995; 51: 1315-24.
 33. Cui L, Hung HMJ, Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics.* 1999; 55: 853-7.
 34. Lehmacher W, Wassmer G. Adaptive sample size calculations in group sequential trials. *Biometrics.* 1999; 55: 1286-90.
 35. Müller HH, Schäfer H. Adaptive group sequential designs for clinical trials: combining the advantages of adaptive and of classical group sequential approaches. *Biometrics.* 2001; 57: 886-91.
 36. Brannath W, König, F, Bauer P. Improved repeated confidence bounds for one sided flexible group sequential designs. *Biom J.* 2003; 45: 311-24.
 37. Gould AL. Sample size re-estimation: recent developments and practical considerations. *Stat Med.* 2001; 20: 2625-43.
 38. Chuong-Stein C, Anderson K, Gollo P, Collins S. Sample size reestimation: a review and recommendations. *Drug Information Journal* 2006; 40: 475-84.
 39. Proschan MA. Sample size re-estimation in clinical trials. *Biom J.* 2009; 51: 348-57.
 40. Gao P, Ware JH, Mehta CR. Sample size re-estimation for adaptive sequential design in clinical trials. *J Biopharm Stat.* 2008; 18: 1184-96.
 41. Gould AL. Interim analyses for monitoring clinical trials that do not materially affect the type I error rate. *Stat Med.* 1992; 11: 55-66.
 42. Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Stat Med.* 2012; 30: 3267-84.
 43. Levin GP, Emerson SC, Emerson SS. Adaptive clinical trial designs with

- pre-specified rules for modifying the sample size: understanding efficient types of adaptation. *Stat Med.* 2013; 32: 1259-75; discusión 1280-2.
44. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. John Wiley & Sons Ltd, Chichester; 2004.
 45. Berry DA. Bayesian clinical trials. *Nature Rev Drug Discov.* 2006; 5: 27-36.
 46. Geller NL. Advances in clinical trial biostatistics. Marcel Dekker: Nueva York; 2004.
 47. Berry DA. Introduction to Bayesian methods. III. Use and interpretation of Bayesian tools in design and analysis. *Clin Trials.* 2005; 2: 295-300.
 48. Berry DA. Adaptive trial design. *Clin Adv Hematol Oncol.* 2007; 5: 522-4.
 49. Sheiner LB, Steimer JL. Pharmacokinetic/pharmacodynamic modeling in drug development. *Ann Rev Pharmacol Toxicol.* 2000; 40: 67-95
 50. Lalonde RL et al. Model-based drug development. *Clin Pharmacol Ther.* 2007; 82: 21-32.
 51. Orloff JJ, Stanski D. Innovative approaches to clinical development and trial design. *Ann Ist Super Sanità.* 2011; 47: 8-13.
 52. Danhof M, Alvan G, Dahl SG, Kuhlmann J, Paintaud G. Mechanism-based pharmacokinetic/pharmacodynamic modeling- a new classification of biomarkers. *Pharm Res.* 2005; 22: 1432-7.
 53. Benson N, Cucurull-Sanchez L, Demin O, Smirnov S, van der Graaf P. Reducing systems biology to practice in pharmaceutical company research; selected case studies. *Adv Exp Med Biol.* 2012; 736: 607-15.
 54. Rullmann JA, Struemper H, Defranoux NA, Ramanujan S, Meeuwisse CM, van Elsas A. Systems biology for battling rheumatoid arthritis: application of the Entelos PhysioLab platform. *Syst Biol (Stevenage).* 2005; 152: 256-62.
 55. Bellanti F, Della Pasqua O. Modelling and simulation as research tools in paediatric drug development. *Eur J Clin Pharmacol.* 2011; 67: 75-86.
 56. Holford NHG, Kimko HC, Monteleone JPR, Peck CC. Simulation of clinical trials. *Annu Rev Pharmacol Toxicol.* 2000; 40: 209-34.
 57. Wang Y, Bhattaram VA, Jadhav PR, et al. Leveraging prior quantitative knowledge to guide drug development decisions and regulatory science recommendations: impact of FDA Pharmacometrics during 2004-2006. *J Clin Pharmacol.* 2008; 48: 146-56.
 58. Temple R. Special study designs: Early escape, enrichment, studies in non-responders. *Commun Stat Theory Methods.* 1994; 23: 499-531.
 59. Temple R. Enrichment designs: efficiency in development of cancer treatments. *J Clin Oncol.* 2005; 23: 4838-9.
 60. Mehta C, Gao P, Bhatt DL, Harrington RA, Skerjanec S, Ware JH. Optimizing trial design: sequential, adaptive, and enrichment strategies. *Circulation.* 2009; 119: 597-605.
 61. European Medicines Agency. Committee on Human Medicinal Products. Assessment report for Ilaris (canakinumab). Procedure n.º EMEA/H/C/001109. Ref.: EMEA/503722/2009. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/001109/WC500031679.pdf. Consultado el 16 de julio de 2013.
 62. European Medicines Agency. Human Medicines Development and Evalua-

- tion. Paediatric Rheumatology Expert Group Meeting - 17 November 2010. Ref: EMA/836276/2010. Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2011/03/WC500103514.pdf. Consultado el 16 de julio de 2013.
63. Simon N, Simon R. Adaptive enrichment designs for clinical trials. *Bios-statistics*. 21 de marzo de 2013. [Epub, en prensa]
64. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy-randomized trials in individual patients. *N Engl J Med*. 1986; 314: 889-92.
65. Guyatt GH, Sackett D, Adachi J, et al. A clinician's guide for conducting randomized trials in individual patients. *CMAJ*. 1988; 139: 497-503.
66. Lewis JA. Controlled trials in single subjects 2. Limitations of use. *BMJ*. 1991; 303: 175-6.
67. Temple R. Problems in interpreting active control equivalence trials. *Account Res*. 1996; 4: 267-75.
68. Rubenfire M, McLaughlin VV, Allen RP, et al. Transition from IV epo-prostenol to subcutaneous treprostinil in pulmonary arterial hypertension: A controlled trial. *Chest*. 2007; 132: 757-63.
69. Pauker SG. Decision analysis and small clinical trials. Presentation to the Institute of Medicine Committee on Strategies for Small-Number-Participant Clinical Research Trials. Washington, D.C.; 2000.
70. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*, 2.^a ed. Oxford: Oxford University Press; 1997.
71. Gibbons JD, Olkin I, Sobel M. An introduction to ranking and selection. *The American Statistician*. 1979; 33: 185-95.
72. Nelson BL, Matejcik FJ. Using common random numbers for indifference-zone selection and multiple comparisons in simulation. *Management Science*. 1995; 41: 1935-45.
73. Robbins, H. Comparing two treatments under biased allocation. *La Gazette des Sciences Mathématique du Quebec*. 1993; 15: 35-41.
74. Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new treatment for those at greater risk. Part I. Introduction. *Am J Pub Health*. 1996; 86: 691-5.
75. Finkelstein MO, Levin B, Robbins H. Clinical and prophylactic trials with assured new treatment for those at greater risk. Part II. Examples. *Am J Pub Health*. 1996; 86: 696-705.
76. Leber P. Observations and suggestions on antedementia drug development. *Alzheimer Dis Assoc Disord*. 1996; 10 (Suppl 1): 31-5.
77. Bodick N, Forette F, Hadler D, et al. Protocols to demonstrate slowing of Alzheimer disease progression. Position paper from the International Working Group on Harmonization of Dementia Drug Guidelines. The Disease Progression Sub-Group. *Alzheimer Dis Assoc Disord*. 1997; 11 (Suppl 3): 50-3.
78. Cummings J, Gould H, Zhong K. Advances in designs for Alzheimer's disease clinical trials. *Am J Neurodegener Dis*. 2012; 1: 205-16.
79. Olanow CW, Rascol O, Hauser R, et al. ADAGIO Study Investigators. A double-blind, delayed-start trial of rasagiline in Parkinson's disease. *N Engl J Med*. 2009; 361: 1268-78.

80. Zhang RY, Leon AC, Chuang-Stein C, Romano SJ. A new proposal for randomized start design to investigate disease-modifying therapies for Alzheimer disease. *Clin Trials*. 2011; 8: 5-14.
81. McDermott MP, Hall WJ, Oakes D, Eberly S. Design and analysis of two-period studies of potentially disease-modifying treatments. *Control Clin Trials*. 2002; 23: 635-49.
82. Cummings JL. Defining and labeling disease-modifying treatments for Alzheimer's disease. *Alzheimers Dement*. 2009; 5: 406-18.
83. Sano M, Ernesto C, Thomas RG, et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *N Engl J Med*. 1997; 336: 1216-22.
84. Petersen RC, Thomas RG, Grundman M, et al. Alzheimer's Disease Cooperative Study Group. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med*. 2005; 352: 2379-88.
85. Krams M, Burman CF, Dragalin V, et al. Adaptive designs in clinical drug development: opportunities, challenges, and scope reflections following PhRMA's November 2006 workshop. *J Biopharm Stat*. 2007; 17: 957-64.
86. Bretz F, Schmidli H, König F, Racine A, Maurer W. Confirmatory seamless Phase II/III clinical trials with hypotheses selection at interim: general concepts. *Biom J*. 2006; 48: 623-34.
87. Bornkamp B, Bretz F, Dmitrienko A, et al. Innovative approaches for designing and analyzing adaptive dose-ranging trials. *J Biopharm Stat*. 2007; 17: 965-95.
88. Evans CH, Ildstad ST, Editors, Committee on Strategies for Small-Number-Participant Clinical Research Trials. *Small clinical trials: issues and challenges*. Washington, D. C.: Institute of Medicine, National Academy Press; 2004.
89. Korn EL, McShane LM, Freidlin B. Statistical challenges in the evaluation of treatments for small patient populations. *Sci Transl Med*. 2013; 5: 178sr3.
90. Committee for Medicinal Products for Human Use-CHMP (2006) Guideline on clinical trials in small populations. Doc. Ref. CHMP/EWP/83561/2005.) Disponible en: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003615.pdf. Consultado el 16 de junio de 2013.
91. Palmer CR, Rosenberger WF. Ethics and practice: alternative designs for phase III randomized clinical trials. *Control Clin Trials*. 1999; 20: 172-86.
92. Goodman SN. Ethics and evidence in clinical trials. *Clin Trials*. 2005; 2: 195-6.
93. Morris AD, Zaritsky AL, LeFever G. Evaluation of ethical conflicts associated with randomized, controlled trials in critically ill children. *Crit Care Med*. 2000; 28: 1152-6.
94. Baiardi P, Giaquinto C, Girotto S, Manfredi C, Ceci A. Innovative study design for paediatric clinical trials. *Eur J Clin Pharmacol*. 2011; 67 (Suppl 1): S109-S115
95. Regnstrom J, Koenig F, Aronsson B, et al. Factors associated with success of market authorisation applications for pharmaceutical drugs submitted to the European Medicines Agency. *Eur J Clin Pharmacol*. 2010; 66: 39-48.