

ICH E9: principios estadísticos en ensayos clínicos

... y también en este número:
Las normas de buena práctica clínica y los CEIC

Hasta no hace mucho tiempo, las recomendaciones en materia de estadística en ensayos clínicos (EC) han sido independientes en los tres grandes mercados (Europa, EEUU y Japón). En Europa, un documento elaborado en 1994 por el Committee for Proprietary Medicinal Products (CPMP) fue el primero dedicado a tratar de forma exclusiva aspectos estadísticos.¹ En febrero de 1998 el «ICH Steering Committee» hizo oficial la recomendación de adoptar el documento definitivo titulado «Principios estadísticos en ensayos clínicos». Tanto la EMEA como la FDA adoptaron el documento en el curso de 1998.

En este artículo se ofrece un resumen de dicha guía, conservando su estructura en siete capítulos.

Introducción

En la introducción se citan las fuentes en que se ha inspirado esta guía, en especial el documento análogo del CPMP¹ aunque también los correspondientes documentos operantes en EEUU² y Japón.³ A lo largo de la guía se hace referencia a otros documentos que incluyen aspectos relacionados con los principios o la metodología estadística.

Asimismo, se hace explícito que el propósito del documento es armonizar los *principios estadísticos* aplicados a ensayos clínicos para solicitudes de registro de medicamentos en Europa, Japón y EEUU. Por tanto, se pretende establecer *principios generales* más que recomendar procedimientos concretos, y se responsabiliza al promotor del ensayo clínico de velar por la adecuada implementación de estos principios generales.

Para ello, se asume que la responsabilidad de ejecutar los procedimientos estadísticos específicos recaerá sobre un *estadístico debidamente cualificado y experimentado*, tal como ya se establecía en las recomendaciones de Buenas Prácticas Clínicas. Aquí se añade que «el estadístico debe tener una combinación de formación y experiencia *suficiente* para implementar los principios articulados en esta guía». Lamentablemente, el significado preciso de los términos «estadístico debidamente cualificado y experimentado» no es evidente, ni siquiera después de haber intentado aclararlo.⁴

Aunque los principios recogidos en esta guía son de interés fundamental en ensayos de fases avanzadas del desarrollo, muchos de los cuales son de naturaleza confirmatoria y tienen como objetivo principal la demostración de eficacia, también son relevantes cuando el objetivo principal se refiere a aspectos de seguridad o incluso en ensayos de fases más tempranas cuya naturaleza es exploratoria. Muchos de estos principios tratan de minimizar el sesgo y de maximizar la precisión de las estimaciones del efecto terapéutico. Por lo tanto, deberían tenerse en cuenta, en la medida de lo posible, en todas las fases del desarrollo clínico.

En gran medida, se hace referencia al uso de métodos frecuentistas cuando se discuten pruebas de contraste de hipótesis e intervalos de confianza. Sin embargo, se advierte que se pueden adoptar otras aproximaciones, como la bayesiana, siempre y cuando existan claras razones para ello y las conclusiones alcanzadas mediante dicha aproximación sean suficientemente robustas. Los métodos bayesianos y los frecuentistas configuran las dos grandes escuelas de la estadística que difieren en la forma de entender y definir el concepto de probabilidad.

El desarrollo clínico global

En este capítulo se abordan los aspectos de los EC que deben considerarse en el contexto del plan de desarrollo de un nuevo medicamento, cuyo objetivo fundamental habría de ser dilucidar si existe un intervalo de dosificación que sea a la vez seguro y efectivo, e identificar los individuos que pueden beneficiarse del tratamiento así como las indicaciones específicas para su uso.

En primer lugar, se destaca la necesidad de disponer de un *programa ordenado* de ensayos clínicos en el que se especifique el objetivo concreto de cada uno de ellos. Dentro del programa, deben distinguirse los EC *confirmatorios* de aquellos cuyos objetivos son eminentemente *exploratorios*. Los primeros están destinados a someter a prueba una hipótesis *preespecificada*, directamente derivada del objetivo principal del EC, y son imprescindibles para demostrar los pretendidos beneficios del tratamiento. La adhesión al protocolo preestablecido en lo relativo al análisis de los datos y al criterio de decisión utilizado para rechazar o no la hipótesis sometida a prueba, es especialmente importante. Cualquier desviación habrá de ser documentada y justificada con todo cuidado. En los EC exploratorios, más propios de fases tempranas del desarrollo, el objetivo fundamental no es someter a prueba una hipótesis preexistente, sino más bien generar hipótesis que habrán de ser contrastadas en ulteriores EC confirmatorios. Por tanto, la preespecificación del análisis, aun siendo necesaria, puede ser algo más vaga, en consonancia con el propósito de exploración. Naturalmente, objetivos confirmatorios y exploratorios pueden convivir en un mismo EC, pero es imprescindible que el protocolo distinga de forma clara el objetivo confirmatorio de los exploratorios.

En lo referente a la población de inferencia, se considera que en EC de fases tempranas del desarrollo puede ser deseable restringirla para maximizar la probabilidad de detectar el efecto del tratamiento. Sin embargo, cuando se abordan los ensayos confirmatorios, la población de inferencia debería asemejarse en lo posible a la población potencialmente beneficiaria.

Se dedica cierta extensión a las variables que se utilizarán como criterio de evaluación, en particular, la variable principal que, a diferencia de las variables secundarias, se utilizará en el análisis confirmatorio y en el cálculo de tamaño muestral. De preferencia, la variable principal debe ser única para evitar problemas de multiplicidad. También se advierte la necesidad de definir de forma inequívoca la variable principal para evitar problemas de multiplicidad. Ello implica especificar, no sólo los procedimientos de medida (definición operacional) sino el tiempo de observación (por ejem-

plo, a los 12 meses de tratamiento) y las estrategias de sustitución en caso de que la observación no esté disponible en ese tiempo (por ejemplo, última observación válida disponible, especificando lo que se entiende por válida). Se discute el uso de variables compuestas (combinación de variables observadas, de acuerdo con un algoritmo predefinido), variables de evaluación global (que integran eficacia y seguridad, o utilidad global), variables sustitutivas («surrogates») y variables categorizadas, comentando las ventajas, alertando acerca de sus limitaciones y ofreciendo criterios para su uso.

En este capítulo se mencionan brevemente las dos técnicas más importantes para evitar el sesgo en la estimación del efecto terapéutico: la asignación aleatoria de los tratamientos y el enmascaramiento. De la primera, se exige que esté documentada y que sea reproducible, aludiéndose a las ventajas de la aleatorización por bloques frente a la no restringida, y recomendándose el uso de bloques de tamaño aleatorio. Cuando se desea equilibrar los grupos de tratamiento en cuanto a un número limitado de factores pronósticos, se propone la aleatorización estratificada recordándose la necesidad de incluir dichos factores en el análisis. Otros procedimientos de asignación dinámicos no deterministas (en los que la probabilidad de asignación a los tratamientos se va modulando para tender a equilibrar los grupos en cuanto a factores pronósticos), se consideran más adecuados si el número de factores pronósticos es elevado. En cuanto al enmascaramiento, se menciona la necesidad de mantenerlo, no solo durante la fase clínica del ensayo, sino durante la revisión de los datos que, como es normal, tiene lugar inmediatamente antes de su análisis. Deben documentarse los casos en que, por cualquier motivo, se haya desvelado el tratamiento.

El diseño del ensayo

Este capítulo trata aspectos del diseño y comienza discutiendo las ventajas y las limitaciones de los tres diseños experimentales más frecuentes en investigación clínica: el diseño paralelo, el cruzado y el factorial. El diseño de dos o más grupos paralelos se presenta como el más frecuente y menos comprometido desde el punto de vista de las suposiciones en que descansa su análisis. Del diseño cruzado 2x2 se menciona su mayor eficiencia pero se advierten algunas limitaciones conocidas. En cualquier caso, se considera que este diseño es satisfactorio cuando se desea demostrar la bioequivalencia de dos formulaciones de un mismo principio activo. Los diseños factoriales se aconsejan para estudiar el efecto conjunto de dos fármacos o las relaciones entre dosis y respuesta para hallar una combinación

óptima de dosis cuando ambos deban de administrarse al mismo tiempo.

Se presta atención a los estudios multicéntricos, de los que se reconoce tanto la importancia de un diseño equilibrado en efectivos (esto es, con igual número de pacientes por centro) como la de una gestión eficaz, durante el reclutamiento, para mantener este equilibrio. Se aconseja también definir en el protocolo el efecto centro, tal como se utilizará en el análisis, si existe la menor posibilidad de duda (por ejemplo, efecto investigador en lugar de centro), incluyendo los criterios que se utilizarán para combinar centros. Se llama la atención sobre la necesidad de definir en el protocolo el modelo concreto que se utilizará para estimar el efecto terapéutico. Dicho modelo puede ser de efectos fijos o aleatorios y contener o no un término de interacción tratamiento-centro.

Las implicaciones de perseguir una demostración de *equivalencia* en lugar de una demostración de superioridad se tratan con cierto detalle en este capítulo. En primer lugar, se distinguen dos categorías de EC según se persiga una demostración de *equivalencia estricta* (o bilateral) o bien una demostración de *no-inferioridad* (o equivalencia unilateral). Los ensayos de *bioequivalencia* caen en la primera de estas categorías mientras que los ensayos de *equivalencia terapéutica* de un nuevo producto frente a otro de referencia suelen caer en la segunda. Estos últimos diseños son mucho más seguros si se incorpora un control adicional con placebo que permita establecer la superioridad respecto a placebo, es decir, la eficacia estricta. Los EC que pretenden demostrar equivalencia terapéutica frente a un control activo tienden a producir estimaciones del efecto sesgadas en favor del objetivo (equivalencia) ante ciertas transgresiones del protocolo. En consecuencia, al contrario de lo que ocurre cuando el objetivo es demostrar superioridad, el análisis por intención de tratar es menos conservador que el análisis de los casos que se han ceñido al protocolo. En cuanto a las técnicas de análisis, no puede establecerse equivalencia en base a la no demostración de diferencias entre los tratamientos, sino que es necesario utilizar técnicas de análisis que sean consistentes con el objetivo de demostrar equivalencia. Estas técnicas pueden estar basadas en el cálculo de intervalos de confianza bilaterales (equivalencia estricta) o unilaterales (no inferioridad), o en pruebas de hipótesis, siempre y cuando las hipótesis nula y alternativa estén planteadas de acuerdo con el objetivo de demostrar equivalencia.

En la línea de las implicaciones que tienen los objetivos particulares de un EC sobre el análisis, se discute brevemente el caso de los ensayos destinados a estudiar la relación existente entre la dosis administrada y la respuesta. Estos ensayos permiten establecer la

eficacia, determinar la forma y situación de la curva dosis-respuesta o estimar la dosis inicial adecuada. En ellos se promueve el cálculo de intervalos de confianza y el uso de métodos gráficos por encima de las pruebas de hipótesis o, en todo caso, la proposición de pruebas de hipótesis adaptadas a los aspectos de la curva dosis-respuesta que se desea probar.

El cálculo del tamaño muestral necesario para proporcionar una potencia adecuada es uno de los puntos de capital importancia en el diseño de EC y, como tal, se le presta la merecida atención en este capítulo. En principio, dicho cálculo debe basarse en el *análisis principal* de la *variable principal*, por lo que es necesario especificar previamente tanto la variable como las hipótesis nula y alternativa involucradas en la prueba de hipótesis y el estadístico en que se basará dicha prueba. También es necesario especificar el riesgo (o probabilidad de rechazar la hipótesis nula en caso de que sea cierta), la potencia (o probabilidad de no rechazar la hipótesis nula en caso de que sea cierta una hipótesis alternativa concreta) y la hipótesis alternativa concreta frente a la que se desea garantizar esta potencia. Esta hipótesis alternativa concreta vendrá determinada por la magnitud del efecto terapéutico que se considera de interés mínimo (relevancia clínica) o que se espera obtener, en caso de que este último sea mayor. Además, habrá que tener en cuenta el impacto de las pérdidas, las retiradas y las transgresiones del protocolo. El método de cálculo del tamaño muestral debe detallarse en el protocolo, así como las estimaciones utilizadas en el cálculo y sus fuentes. Para la selección del riesgo se debe tener en cuenta la posible necesidad de ajuste inducida por problemas de multiplicidad, la plausibilidad *a priori* de las hipótesis sometidas a prueba y la credibilidad que tendrán los resultados. En cuanto a la potencia, el promotor es quien habría de tener mayor interés en incrementarla en la medida de lo posible.

En EC de equivalencia, el cálculo de tamaño muestral debe basarse en los procedimientos de análisis adecuados para este tipo de diseños. Aunque es práctica común adoptar como la alternativa de interés aquella que supone una diferencia de efectos nula, es importante tener presente que esta suposición puede infraestimar el tamaño muestral necesario para alcanzar la potencia deseada en caso de que dichas diferencias sean no nulas. La definición del intervalo de equivalencia debe ser razonada y su límite o sus límites deben ser inferiores a lo que se considere una diferencia mínima clínicamente relevante.

En ensayos con diseño de grupos secuenciales, el tamaño muestral no se determina de antemano, pero debe considerarse su distribución (valor esperado, valor máximo) al adoptar una determinada regla de decisión para la finalización prematura del ensayo.

La ejecución del ensayo

En este capítulo se discuten las cuestiones propias de la fase clínica del EC que tienen importantes implicaciones estadísticas, como las modificaciones en los criterios de inclusión o exclusión, el reclutamiento, la reestimación del tamaño muestral y, en especial, los análisis intermedios y el papel de los comités independientes de monitorización de datos.

A este respecto, se hace distinción entre dos tipos de monitorización según su posible impacto en los resultados del ensayo o su interpretación. Ciertos procesos de monitorización, como los destinados a controlar la calidad de los datos recogidos, la adhesión al protocolo por parte de los investigadores, el ritmo de inclusión de pacientes, o las reevaluaciones del tamaño muestral necesario, no implican desvelar el código de tratamiento ni involucran análisis comparativos, por lo que no representan una amenaza. Por el contrario, los análisis intermedios implican ambas cosas y tienen dos consecuencias importantes: consumen riesgo y desvelan el tratamiento. Este tipo de análisis requiere una planificación previa en el protocolo donde se especifique la regla de decisión que se utilizará para decidir la continuidad del ensayo y se aborde el problema de multiplicidad implícito, asegurándose la contención del riesgo global mediante métodos apropiados. También es necesario garantizar la confidencialidad de la información que se maneja para llevar a cabo el análisis intermedio, incluyendo sus resultados, pues de lo contrario podrían producirse cambios de actitud en los responsables del ensayo que produzcan alteraciones en el futuro curso del mismo. Para ello, los análisis intermedios deben realizarse bajo procedimientos normalizados de trabajo que garanticen la confidencialidad del proceso, e idealmente deben ser ejecutados por un comité de monitorización independiente con objeto de no diseminar la información entre el personal responsable del ensayo.

El análisis de los datos

En lo que se refiere al análisis de los datos, quizás sea una contribución notable del documento la forma en que se trata el plan de análisis. Desde luego, el protocolo debe contener una sección destinada a planificar el análisis de los datos. En los EC confirmatorios, el análisis principal (confirmatorio) debe especificarse con detalle. En el caso de análisis exploratorios la planificación del análisis puede ser menos detallada. En todos los casos, se recomienda que, con posterioridad a la finalización del protocolo, se redacte un *plan de análisis* en el que se elaboren todos los aspectos relativos al

análisis de datos que no se hayan detallado en el protocolo. Este documento, puede revisarse o actualizarse como resultado de revisiones de los datos llevadas a cabo antes de desenmascarar los tratamientos. De esta forma, se puede juzgar objetivamente hasta qué punto los análisis finales realizados difieren de los previstos. Este juicio es muy importante porque sólo los análisis preespecificados pueden considerarse confirmatorios.

Otra contribución afortunada de esta guía es la forma en que se contemplan los distintos *conjuntos de análisis*, su definición, y el distinto papel que pueden jugar según los objetivos del ensayo. En primer lugar, se reconoce el carácter utópico del análisis por intención de tratar que, en estado puro, habría de incluir a todos los sujetos aleatorizados. Se introduce el término de conjunto de análisis *completo* («full analysis set»), que debe ser lo más cercano posible al de todos los sujetos aleatorizados. Sin embargo, se reconoce que además de las limitaciones impuestas por la falta de datos post-aleatorización, es razonable excluir de este conjunto aquellos casos que no pertenezcan a la población objetivo, siempre y cuando dicha pertenencia pueda establecerse con seguridad, según el estado previo a la aleatorización y con el mismo rigor para todos los casos. También se permite la exclusión de los casos que no iniciaron el tratamiento asignado si puede considerarse que la decisión de no tomar la medicación fue independiente del tratamiento asignado.

Las transgresiones del protocolo tras la asignación del tratamiento, pueden estar relacionadas con las características del propio tratamiento y, en general, deben mantenerse en el conjunto de análisis completo. La carencia de datos debidos a abandonos y retiradas puede tratarse mediante diversas técnicas de sustitución como la traslación de la última observación disponible («last observation carried forward»), la suposición del peor escenario posible en el caso de variables dicotómicas u ordinales, o el ajuste de modelos matemáticos complejos. El protocolo debe prever la utilización de cualquiera de estas estrategias y, en caso de optarse por la última de las mencionadas, las suposiciones del modelo deben explicarse cuidadosamente.

Un segundo conjunto de análisis utilizado con frecuencia es el de los casos que se consideran *válidos* porque han seguido fielmente el protocolo o ciertos aspectos fundamentales del mismo como, por ejemplo, la administración del tratamiento durante un tiempo mínimo, la disponibilidad de la variable principal o la ausencia de transgresiones mayores. Este conjunto de análisis maximiza las oportunidades de detectar un efecto terapéutico pero puede introducir sesgos si se invalidan casos por transgresiones del protocolo que, a su vez, dependen del tratamiento (por ejemplo, interrupción del tratamiento por acontecimientos adversos). Por ello, con-

viene comparar los grupos de tratamiento en cuanto a la frecuencia de exclusiones por distintos motivos.

Es deseable que el protocolo establezca y defina los conjuntos de análisis. Sin embargo, muy a menudo surgen problemas no anticipados y puede ser preferible diferir ciertas decisiones hasta que se efectúa la revisión ciega de los datos. Las exclusiones de casos de los distintos conjuntos de análisis deben documentarse en el informe final, junto con los motivos que las originaron.

Es aconsejable explorar la sensibilidad de los resultados del análisis frente a variaciones en el conjunto empleado y, de hecho, puede convenir no limitarse a los dos conjuntos mencionados, que representan dos extremos. Cualquier diferencia en los resultados debe ser discutida en el informe final.

Por último, se llama la atención sobre el distinto papel que juegan los conjuntos completos y de casos válidos según cuál sea el objetivo del ensayo. En los ensayos destinados a demostrar superioridad, el conjunto completo suele proporcionar resultados más conservadores en cuanto al efecto terapéutico y por ello el análisis principal tendría que realizarse sobre este conjunto. Sin embargo, en los ensayos que pretenden demostrar equivalencia o no inferioridad, el conjunto completo no produce resultados conservadores y su papel en estos ensayos debería considerarse con prudencia.

Otros temas abordados en este capítulo hacen referencia al manejo de valores ausentes y de valores extremos y al uso de transformaciones. Para todos ellos, se aconseja establecer la estrategia al respecto durante la revisión ciega de los datos.

En lo referente a la estimación del efecto terapéutico y al uso de pruebas de hipótesis, se exige que los procedimientos utilizados se estipulen en el protocolo. El modelo estadístico utilizado en el análisis debe reflejar el estado actual de conocimientos médicos y estadísticos en el área de investigación así como el diseño del ensayo. Es necesario especificar el modelo por completo, mencionando todos sus términos, incluyendo los introducidos para compensar desequilibrios en factores pronósticos o para reducir la variabilidad residual, como las variables de ajuste («covariates») utilizadas en los análisis de la covarianza. La selección de métodos de análisis debe hacerse teniendo en cuenta la distribución de las variables pero también la necesidad de ofrecer estimaciones puntuales y por intervalo del efecto terapéutico.

El análisis principal de la variable principal debe distinguirse claramente de otros análisis secundarios realizados sobre la variable principal o las variables secundarias.

Los problemas de multiplicidad derivados de utilizar

más de una variable principal, realizar comparaciones múltiples (en caso de diseños con más de dos tratamientos) o de repetir las medidas de una variable en distintos momentos, habrán de resolverse mediante ajustes del riesgo. Sin embargo, es preferible evitar o reducir la multiplicidad ocasionada por las situaciones anteriores definiendo una única variable principal, identificando un contraste de interés crítico o resumiendo las medidas repetidas mediante índices apropiados, como el área bajo la curva, la pendiente u otras. En los análisis confirmatorios, cualquier problema de multiplicidad que persista tras adoptar este tipo de soluciones habrá de resolverse mediante ajustes del riesgo y, si no se consideran necesarios, debe discutirse los motivos en el protocolo.

En muchas ocasiones, existen factores con posible influencia en la variable principal o, simplemente, capaces de reducir la variabilidad. El efecto centro en estudios multicéntricos y los valores basales de la variable principal son dos casos especialmente frecuentes. Si la importancia de estos u otros factores es tal que se considera deseable tenerlos en cuenta en el análisis principal, es necesario preverlo en el protocolo. En caso de duda se aconseja definir como principal el análisis no ajustado y, como secundario, el ajustado. No es aconsejable utilizar observaciones pos aleatorización como variables de ajuste del análisis principal porque éstas pueden estar afectadas por los tratamientos.

El efecto del tratamiento puede depender también de ciertos factores o variables y, si esta posibilidad se anticipa en fase de diseño, el análisis de este fenómeno puede formar parte del análisis principal. En general, es preferible iniciar el estudio de este fenómeno incluyendo términos de interacción en el modelo más que realizando análisis de subgrupos.

Por último, se menciona la importancia de los métodos y de las herramientas utilizadas para el manejo y el análisis de los datos. Los programas informáticos deben ser sometidos a prueba para garantizar que no introducen distorsiones indeseadas, y estas pruebas deben documentarse.

Evaluación de seguridad y tolerabilidad

Aunque los métodos y medidas óptimos para evaluar la seguridad y la tolerabilidad pueden depender de diversos factores, los datos de laboratorio, los signos vitales y los acontecimientos adversos suelen constituir el cuerpo principal de los datos de seguridad y tolerabilidad. Los acontecimientos adversos graves y los abandonos de tratamiento debidos a acontecimientos adversos son especialmente importantes. Otros criterios

como los acontecimientos adversos que emergen tras iniciar el tratamiento, pueden ser de utilidad, especialmente cuando se analizan poblaciones con abundante comorbilidad.

En cuanto a los datos de laboratorio, se debe prestar atención a las unidades y a los intervalos de referencia y, en caso de que no sean constantes dentro de un mismo EC, los valores observados deben estandarizarse mediante transformaciones para poder realizar una evaluación unificada.

La utilización de una metodología consistente a lo largo del programa de desarrollo clínico y, en particular, el uso de diccionarios estándar para la codificación de acontecimientos adversos y otros términos médicos, es un punto de capital importancia para permitir una descripción de resultados compacta e informativa, y también para posibilitar la combinación ulterior de datos procedentes de diversos EC.

Tanto por la abundancia y complejidad de los datos, como por su carácter exploratorio, los análisis de seguridad suelen involucrar técnicas descriptivas apoyadas por el cálculo de intervalos de confianza, cuando éstos faciliten la interpretación. Las representaciones gráficas son un instrumento muy valioso. En cuanto a los datos de laboratorio, se recomienda realizar análisis cuantitativos (medias, medianas, etc.) pero también cualitativos, evaluando las frecuencias con que aparecen valores por encima de determinados límites. Si deben analizarse cantidades considerables de variables, las pruebas de significación pueden ser útiles como mecanismo para atraer la atención hacia aquellas que requieran mayor profundización.

Los análisis de seguridad y tolerabilidad suelen realizarse sobre el conjunto de pacientes que hayan sido sometidos al menos a una dosis del tratamiento. La incidencia de acontecimientos adversos puede estimarse de distintas formas. Aunque es habitual utilizar el número de pacientes expuestos como denominador, según las situaciones, puede ser preferible tener en cuenta la duración de la exposición (personas-año) y utilizar técnicas de análisis de supervivencia.

Informes

Este capítulo trata muy brevemente ciertas cuestiones relativas al informe de los resultados del análisis de los datos de un EC, dirigiendo al lector hacia una guía que aborda con mayor detenimiento el informe de resultados («Structure and Content of Clinical Study

Reports»). Aunque el contenido de este capítulo es un tanto repetitivo, cabe destacar ciertos puntos.

Todas las decisiones tomadas durante la revisión ciega de los datos deben documentarse en el informe final, y es de capital importancia distinguir entre las que se tomaron durante la revisión ciega de los datos y las tomadas con posterioridad al desenmascaramiento de los tratamientos. La credibilidad de un estudio dependerá en gran medida de las discrepancias existentes entre lo previsto antes del desenmascaramiento y lo finalmente realizado.

Es frecuente que, durante la revisión ciega de los datos surjan aspectos imprevistos que susciten nuevas hipótesis, distintas de las que motivaron el EC. La exploración detallada de estas hipótesis es perfectamente admisible, pero este tipo de análisis debe distinguirse estrictamente de los análisis previstos antes de tener acceso a los datos.

Un detalle relativo a los resultados de las pruebas de significación es que debe documentarse el grado de significación («p-value») exacto (por ejemplo, $p=0.034$) en lugar de indicar niveles de significación discretos para los que la prueba produce rechazo de la hipótesis nula (por ejemplo, $p<0.05$).

Por último, de cara a que el juicio estadístico oriente no solo el análisis sino también la interpretación de los resultados y la forma de presentarlos, se reconoce que el estadístico del ensayo debería formar parte del equipo que redacta el informe final y, en particular, debería aprobar dicho informe.

Comentario final

Adicionalmente, la guía incluye como anexo un glosario en el que se definen algunos de los tecnicismos utilizados en el texto. El documento puede obtenerse en la dirección de Internet <http://www.ifpma.org/ich5e.html>.

Referencias

1. CPMP. Note for guidance. Biostatistical methodology in clinical trials in applications for marketing authorisations for medicinal products. Diciembre 1994.
2. FDA. Guideline for the format and content of the clinical and statistical sections of a new drug application. Julio 1988.
3. Japanese Ministry of Health and Welfare. Guidelines on statistical analysis of clinical studies. Marzo 1992.
4. Morgan D, Bacchieri A, Bay C, Cobo E, Van Ewijk P, Huitfeldt B, et al. Qualified statisticians in the European pharmaceutical industry: report of an EFSPi working group. *DIA Journal* (en prensa).

Las normas de buena práctica clínica y los CEIC

Las normas de buena práctica clínica (BPC) actualmente en vigor, consensuadas en el seno de las Conferencias Internacionales de Armonización (ICH) y aceptadas por las autoridades reguladoras de Estados Unidos, Europa y Japón,¹ persiguen el mutuo reconocimiento de los resultados de los ensayos clínicos que se realicen de acuerdo con sus recomendaciones. Sin embargo, en la actualidad no tienen el mismo rango legal en todo el territorio ICH. También hacen algunas recomendaciones de carácter ético que ya están incluidas en la legislación de la mayoría de los países occidentales.

En general, la investigación clínica con medicamentos en España ha evolucionado favorablemente hacia el cumplimiento de las normas de BPC, aunque todavía no puede compararse con la de otros países, como Estados Unidos. Es obvio que debemos considerar y aprender de la experiencia de otros, pero también debemos valorar las posibilidades de su aplicación en nuestro medio e ir adoptando, de forma paulatina, aquellos cambios deseables aunque difíciles de conseguir.

No son muchas las novedades que las normas de BPC-ICH incorporan respecto a lo legalmente establecido en nuestro país; sin embargo, algunas de estas novedades son importantes. Considerando la legislación nacional (principalmente el Real Decreto 561/93), las normas de BPC-ICH incorporan los siguientes puntos.

- 1º La obligación de contar en los CEIC con al menos un miembro *ajeno a la institución*, sin especificar si éste debe ser sanitario o no sanitario.
- 2º Mantener una *lista actualizada y pública de los miembros que componen el Comité* (no requiere especificar los miembros que estuvieron presentes en una reunión determinada, como algunos promotores solicitan).
- 3º Disponer por escrito de unos *procedimientos normalizados de trabajo* (PNT) que incluyan la composición de sus miembros (nombres y calificaciones), la autoridad que lo acredita, convocatorias y plazos de notificación para las reuniones, evaluación inicial y seguimiento de los ensayos, así como frecuencia de las revisiones de seguimiento, opinión o informe favorable de la evaluación inicial y de los cambios mayores y menores, justificada por escrito, y comunicaciones con el investigador.

4º La obligación de emitir *informes escritos de sus decisiones u opiniones –incluidas las de rechazo–* y de remitirlos al investigador o a la institución, tanto sobre los protocolos de ensayo clínico como sobre las modificaciones posteriores a su aprobación.

5º Permiten establecer en sus PNT un «*quorum*» *distinto para las decisiones a tomar*, por ejemplo para la revisión abreviada de cambios menores.

6º *Archivo de los documentos* relevantes (PNT, lista de miembros, actas y correspondencia, etc.) durante un periodo de al menos 3 años tras la finalización de un ensayo.

Para un correcto funcionamiento de los CEIC es preciso que se realice el seguimiento periódico de sus actividades y se compruebe el cumplimiento de los PNT mediante las correspondientes inspecciones por parte de las autoridades competentes. El programa de inspecciones de la Comunidad de Madrid se plantea desde esta óptica, ya que al garantizar el buen funcionamiento de los CEIC se persigue proteger a los sujetos experimentales y garantizar el respeto a los derechos humanos. Además, en los casos en que proceda, las inspecciones al equipo investigador y al promotor permitirán comprobar la veracidad de los datos y de los resultados de estos ensayos clínicos.

Sin embargo, ¿es necesaria tanta burocracia para asegurar que los resultados obtenidos son ciertos? ¿Se compromete siempre la validez interna y externa del estudio si no se cumplen las normas de BPC? ¿Cómo repercuten en el cuidado del paciente? Por otro lado, para los ensayos clínicos cuya finalidad no es la autorización de comercialización o justificar cambios en la misma, ¿qué grado de cumplimiento de estas normas es exigible?

Es evidente que el objetivo es la homogeneización de criterios con la finalidad de que las agencias reguladoras acepten los resultados de los ensayos clínicos realizados según las normas de BPC. No obstante, para facilitar esto hay que reflexionar sobre la posibilidad y conveniencia de la simplificación de los trámites. Por ejemplo, en el caso de trámites que no aportan nada a la seguridad y cuidado del paciente, a la validez interna (*que los resultados obtenidos sean ciertos*) o a la validez externa (*que los resultados sean generalizables*) del estudio.

Por otro lado, para los ensayos clínicos cuya finalidad no es la autorización de comercialización, habría que considerar qué aspectos de las BPC no son necesarios y, por tanto, susceptibles de ser eliminados como obligatorios.

Una vez que el péndulo ha llegado a los extremos (*nos creemos los resultados de los estudios sin normas ni comprobaciones/no creemos nada que no se haya realizado e inspeccionado para verificar el cumplimiento de las normas de BPC*), quizás deberíamos pensar un poco más en los protagonistas del ensayo clínico: en primer lugar, el **paciente**, que es quien asume el riesgo y, en segundo lugar, el **investigador**, que ante todo es médico y, por tanto, responsable del cuidado y mejor interés de sus pacientes.

Por ello, debemos ser conscientes de que una oportuna flexibilidad de las exigencias (siempre sin menoscabo

de la protección de los sujetos del ensayo, del respeto a los derechos humanos y de la validez de los resultados) tal vez sea la mejor manera de crear un ambiente propicio en el que los actores implicados en la investigación clínica atiendan a lo importante, en un orden de prioridad correcto. Como decía Aristóteles, «la virtud está en el término medio».²

Este texto fue presentado en parte como ponencia a la Mesa Redonda “Armonización en la inspección y control de medicamentos” celebrada en el XX Symposium de A.E.F.I. “Nuevo entorno para la industria Farmacéutica”, San Lorenzo del Escorial, 21 al 23 de octubre de 1999.

Referencias

1. Note for Guideline for Good Clinical Practice (CPMP/ICH/135/95).
2. Aristóteles. *Ética a Nicómaco*.

Inés Galende Domínguez. Servicio de Ordenación y Asistencia Farmacéutica. Consejería de Sanidad de la Comunidad de Madrid

Noticias

Comité de Redacción. Desde el primer número de *Investigación Clínica y Bioética*, la Dra. María Antonia Serrano ha sido miembro de su Comité de Redacción. Por motivos personales, a principios de año presentó su dimisión al Presidente de la Sociedad Española de Farmacología Clínica. *ICB* le agradece toda la dedicación prestada. Ha sido sustituida por la Dra. Mónica Saldaña, Farmacóloga clínica, Agencia Española del Medicamento.

Máster en Bioética y Derecho. 1ª edición a distancia, 4ª edición presencial. Dos cursos académicos: 2000-2001, 2001-2002. Organización: Les Heures, Universitat de Barcelona y Observatori de Bioética i Dret, Parc Científic de Barcelona. Información: Assumpta Cajal, teléfono 934 28 3710, acajal@fbg.ub.es

Consejo editorial: Diego Gracia (Director), Pau Salvà (Jefe de Redacción), Carles Vallvé (Secretario), Olav Bakke, José Nicolás Boada, Xavier Carné, Joan Costa, Jesús Frías, Inés Galende, Juan A. García Vicente, Cándido Hernández, Ángela Idoipe, Antonio Portolés, Carlos Rodríguez, Carlos Romeo Casabona.

Comité de redacción: Olav Bakke, Inés Galende, Mónica Saldaña, Pau Salvà, Carles Vallvé.

© Sociedad Española de Farmacología Clínica
ISSN 1131-8910
Depósito Legal B-9353-1993

Servicio de Farmacología Clínica
CSU Vall d'Hebron - 08035 Barcelona
e-mail: icb@icf.uab.es

Tel.: 934 28 3029
Fax: 934 89 4109
Suscripción: 2.500 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. Las autoridades sanitarias de Castilla-La Mancha, Murcia y de la Comunidad Valenciana también contribuyen a su publicación. Además *ICB* recibe apoyo económico de las siguientes instituciones: Almirall-Prodesfarma, Amgen, Asta, AstraZeneca, Bayer, Biomedical Systems, Boehringer Ingelheim, Byk-Elmu, Esteve, Faes, Farmaindustria, Ferrer, Glaxo Wellcome, Grifols, Hoechst Marion Roussel, Janssen-Cilag, Knoll, Lácer, Merck Farma y Química, Merck Sharp & Dohme, Novartis, Pharmacia-Upjohn, Phoenix International España, Roche, Rhône-Poulenc Rorer, Rovi, Sanofi-Synthelabo, Schering España, Searle, Serono, SmithKline Beecham, Solvay, Uriach, Vita-Invest y Zambon. Los artículos y notas publicados en *ICB* no pueden ser reproducidos sin permiso escrito.