

## Ensayos clínicos controlados

---

*Juan Carlos López Alvaranga  
Gregorio Tomás Obrador*

### INTRODUCCIÓN

Cuando se investiga una enfermedad, es frecuente que se observe discrepancia en los diagnósticos emitidos por diferentes observadores. Esta divergencia puede ser aún mayor al comparar los tratamientos recomendados, en parte porque se han aprendido de acuerdo a escuelas. Por ejemplo, existe mucha variación en los esquemas de tratamiento insulínico para diabéticos tipo 1: tratamiento convencional de dos aplicaciones diarias, tratamiento con seis aplicaciones diarias y aun mezclas con diferentes tipos de insulina.

Por otro lado, muchos de los tratamientos que se han prescrito carecen de fundamento, y otros, como lo muestra la historia de la medicina, han dañado a quienes los han utilizado. Como ejemplos relativamente recientes podemos citar la recomendación indiscriminada de vitaminas para mejorar el estado de ánimo; también el uso de la talidomida para la émesis de mujeres embarazadas, que tuvo como consecuencia una epidemia de niños con focomelia o amelia.

Lo anterior, unido a los logros de la investigación farmacológica actual, hacen necesario el estudio de forma sistemática de los nuevos tratamientos. Los ensayos clínicos controlados (ECC) son el tipo de diseño epidemiológico más adecuado

para evaluar la eficacia de los tratamientos, porque proporcionan el máximo grado de inferencia causal y permiten un mejor control de posibles factores confusores.

## DEFINICIÓN

Los ECC son un tipo especial de estudio longitudinal de cohortes, en los que el investigador asigna a un grupo seleccionado de sujetos a recibir (grupo experimental) una u otra (grupo control) intervención (generalmente un tratamiento) y los sigue en el tiempo para comparar su efecto. A los ECC también se les conoce como estudios experimentales, porque a diferencia de los observacionales, el investigador controla todos los aspectos del estudio. En otras palabras, el investigador realiza un auténtico experimento, similar al que se realiza en los laboratorios.

En el capítulo 5 se definen los elementos esenciales de los ECCs, como la selección de los pacientes, la aleatorización y el cegamiento. También se discuten los problemas asociados con su realización, como viabilidad, problemas éticos, validez externa y costo. En este capítulo se analizan los tipos de ECC y se examinan en mayor detalle los elementos esenciales de los mismos. También se incluye una breve explicación de las diferentes fases de la investigación farmacológica.

## TIPOS DE ENSAYOS CLÍNICOS

### ■ ENSAYOS CLÍNICOS HISTÓRICOS

En este tipo de ensayo clínico, los resultados del grupo experimental se comparan con los de un grupo histórico que no recibió el tratamiento. Son pocas las enfermedades que no requieran de un grupo control para evaluar la eficacia de un tratamiento. La razón es una de ellas, porque una vez iniciado el proceso clínico la muerte es inevitable en la mayoría de los casos. Si se demuestra que un medicamento evita la muerte por rabia, no es necesario tener un grupo control con placebo. Además, tal estudio sería inadmisiblemente desde el punto de vista ético.

Los ensayos clínicos históricos tienden a dar falsos positivos, porque el pronóstico del control histórico frecuentemente es peor que el del paciente del grupo experimental. Es por ello que la mayoría de los ECC comparan los resultados del tratamiento con los de un grupo control concurrente, que se someta a las mismas exigencias metodológicas que el grupo experimental.

### ■ ENSAYOS CLÍNICOS PARALELOS

Este tipo de diseño es el más frecuente y consiste en que los pacientes son asignados por el investigador a recibir de forma aleatoria uno de los tratamientos que

se estudian (figura 9-1). Generalmente se compara la *sustancia activa* contra *placebo*, lo que tiene por objeto aumentar el contraste de los efectos benéficos y adversos. Lo deseable es que exista una gran diferencia en los efectos benéficos y poca o ninguna en los adversos. Si existe un tratamiento aceptado como estándar para la enfermedad al momento de realizar el estudio, la sustancia activa se suele comparar contra dicho *tratamiento estándar o convencional*.

### ■ ENSAYOS CLÍNICOS DE DISEÑO FACTORIAL

En este tipo de diseño, los grupos son aleatorizados para recibir dos o más tratamientos y/o el placebo o tratamiento convencional (figura 9-1). En el estudio sobre Modificación de la Dieta en la Enfermedad Renal (*Modification of Diet in Renal Disease* o MDRD) se utilizó este tipo de diseño. Los pacientes fueron inicialmente asignados a recibir control normal o estricto de la presión arterial, y después a una dieta baja o muy baja en proteínas. La hipótesis del estudio fue que las dos intervenciones, el control de la presión arterial y la dieta hipoproteica, independientemente disminuyen la tasa de progresión de la insuficiencia renal crónica, la cual se evaluó por el grado de disminución de la tasa de filtración glomerular. Una ventaja del diseño factorial es que permite evaluar dos hipótesis en un sólo estudio, lo cual aumenta la relación costo-beneficio. Es importante remarcar que una condición del diseño factorial es que no exista interacción entre los tratamientos, para así evitar que se confunda el efecto de uno de los tratamientos con el otro.

### ■ ENSAYOS CLÍNICOS CRUZADOS

En los ECC cruzados, los dos grupos reciben ambos tratamientos secuencialmente pero de una forma aleatorizada. Es decir, un grupo recibe los tratamientos con secuencia A→B y el otro con secuencia B→A (figura 9-1). Este tipo de diseño se utilizó para determinar la eficacia del acarbosa en pacientes diabéticos tipo 2 con falla secundaria a hipoglucemiantes orales. 17 pacientes con *diabetes mellitus* tipo 2 que cumplieron criterios para falla secundaria al tratamiento con mezcla de hipoglucemiantes orales (clorpropamida más metformina) fueron aleatorizados para recibir una de dos secuencias: la secuencia 1, que consistía en acarbosa durante 12 semanas, un período de lavado de 3 semanas, y placebo durante otras 12 semanas; o la secuencia 2, que consistía en placebo durante 12 semanas, un período de lavado de 3 semanas, y acarbosa durante otras 12 semanas. El período de lavado es el tiempo que debe pasar para que los efectos residuales del medicamento que se utilizó en el período anterior desaparezcan. La duración de dicho período la determina el investigador basado en la farmacocinética y farmacodinamia del medicamento en estudio.

El diseño cruzado tiene las siguientes ventajas: 1) comparado con los estudios paralelos, sólo requiere la mitad del tamaño de muestra; 2) la variabilidad es menor debido a que los mismos sujetos reciben los dos tratamientos. No obstante, el diseño cruzado tiene la desventaja de que si un número pequeño de pacientes no concluyen el estudio, el poder estadístico para demostrar que las diferencias entre los dos tratamientos son válidas puede ser insuficiente, debido a la reducción del tamaño de la muestra.

### SELECCIÓN DE LOS PACIENTES

La validez interna de los ECC depende, en buena parte, de que la distribución de los factores que influyen los resultados que se están estudiando sea similar en el grupo experimental y en el grupo control; de este modo las diferencias encontradas serán debidas únicamente al tratamiento. Aunque la aplicación de criterios de selección estrictos aumenta la validez interna del estudio, es decir, la validez de las conclusiones, lo hace a expensas de disminuir la validez externa o generalización de los resultados a otros grupos. Por ejemplo, en el estudio MDRD antes mencionado, se incluyeron muy pocos diabéticos tipo 1 e individuos de raza negra, por lo que los resultados del estudio no se pueden generalizar a estos grupos de pacientes.

### ASIGNACIÓN DE LOS TRATAMIENTOS

La asignación de los tratamientos se suele hacer de forma aleatoria, es decir al azar, de modo que cada uno de los sujetos incluidos en el estudio tenga la misma probabilidad de recibir cualquiera de los tratamientos que se están comparando. El objeto de la *aleatorización* es disminuir el error producido por factores "no controlados" que pueden influenciar el resultado final del estudio. Si todos los pacientes cumplieron con los criterios de selección y la aleatorización fue correcta, las diferencias observadas entre los grupos experimental y control serán debidas únicamente al tratamiento y no a las características de los pacientes.

Es importante señalar que, cualquier método que no utilice la probabilidad para asignar el tratamiento no es aleatorio. Por ejemplo, asignar tratamientos alternos a pacientes que son seleccionados secuencialmente es un modelo determinístico, no probabilístico.

La *aleatorización en bloques* es uno de los métodos probabilísticos más empleados. Consiste en formar bloques de todas las combinaciones posibles de tra-

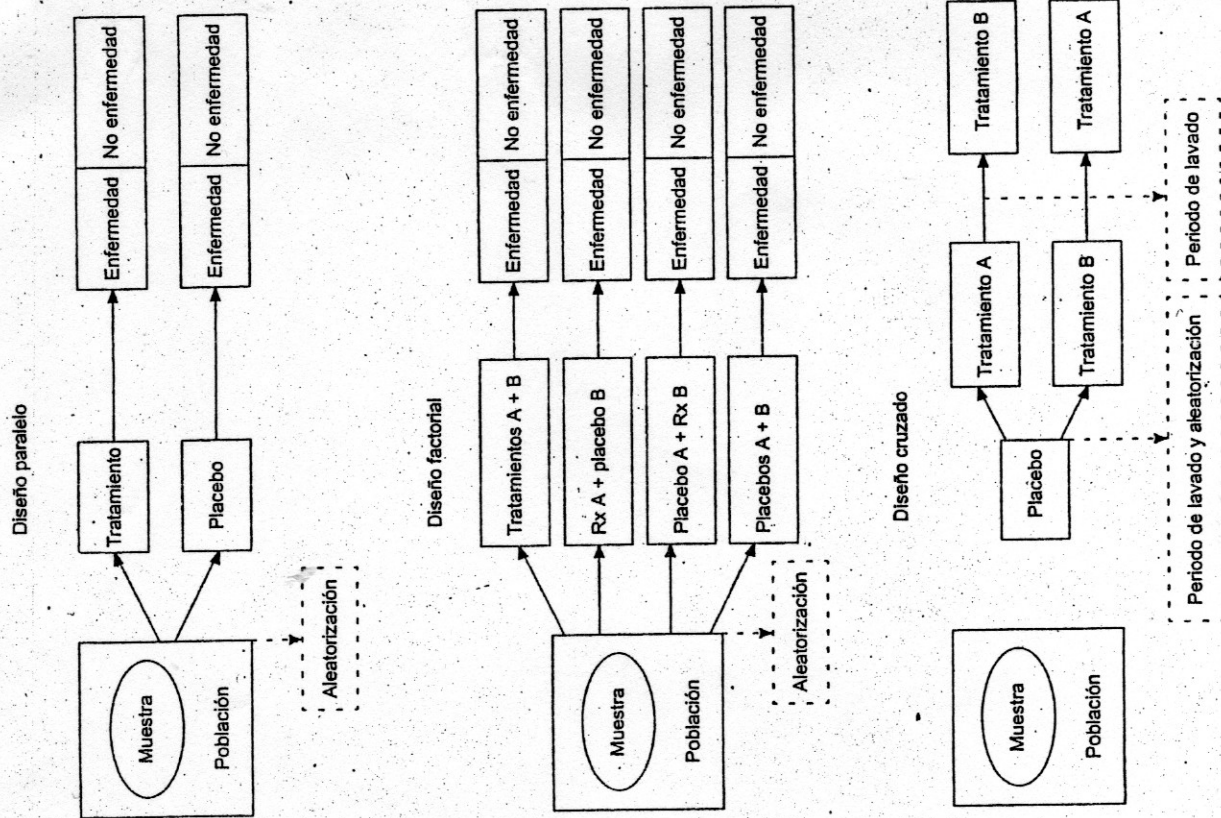


Figura 9-1. Tipos de ensayos clínicos. Rx = tratamiento.

tamiento. Por ejemplo, si se investigan dos tratamientos A y B se pueden formar bloques de cuatro con las siguientes combinaciones: AABB, ABAB, BABA, BBAA, ABBA y BAAB. Después de formar los bloques, la aplicación de los mismos se aleatoriza. De este modo se consigue que los grupos tengan el mismo número de pacientes y que la secuencia de los tratamientos A y B sea impredecible.

La *estratificación* es otro método para reducir el error experimental y aumentar el poder estadístico en el análisis. Los pacientes se estratifican de acuerdo al pronóstico, por ejemplo el sexo, edad o extensión de la enfermedad; después cada estrato se aleatoriza para recibir los tratamientos. De esta forma se reduce la probabilidad de que haya imbalances en las características de los pacientes asignados al grupo experimental y al control.

## CEGAMIENTO

El cegamiento consiste en que ni los participantes ni los investigadores del estudio conocen a que grupo de tratamiento (experimental o control) pertenece cada sujeto. El término *ciego* se utiliza para indicar que el paciente no conoce a que grupo pertenece, y el *doble ciego* para indicar que tanto el paciente como el investigador o administrador del tratamiento están "ciegos" a dicha información. La razón del cegamiento es evitar el sesgo de observación, que consiste en que tanto el sujeto como el investigador pueden comportarse de una forma sistemáticamente diferente si saben que tratamiento están recibiendo.

Para conseguir el cegamiento, las pastillas o cápsulas de la sustancia activa y del placebo deben tener el mismo color, forma y sabor. El investigador principal guarda el código que contiene información de lo que está tomando cada paciente; dicho código no se abre, excepto en caso de urgencia. En los estudios *triple ciego*, ni el paciente, ni el administrador del tratamiento ni los investigadores saben que está tomando cada paciente, hasta que se completa el análisis del estudio. Este tipo de diseño pretende evitar por completo el sesgo de observación.

Aunque es deseable que todos los ECC sean doble ciegos, en algunos casos no se puede hacer cegamiento porque se trata de un medicamento con efectos secundarios potencialmente graves o que produce cambios que pueden ser fácilmente reconocidos por el paciente o por el investigador. Por ejemplo, si el tratamiento experimental es un bloqueador beta, el paciente y/o el médico pueden reconocerlo fácilmente por la bradicardia que produce este tipo de medicamentos. Los ensayos clínicos en los que no se realiza cegamiento se denominan *abiertos* o de marca abierta (*open u open label*).

## ANÁLISIS DE LOS RESULTADOS

### RIESGOS RELATIVO Y ABSOLUTO

Las medidas más comúnmente utilizadas en la eficacia de un tratamiento son la reducción del *riesgo relativo* o *absoluto* de la enfermedad con el tratamiento experimental comparado con el control. Si en un ECC se observa que la tasa de enfermedad con el tratamiento control A es de 40% y con el tratamiento experimental B de 20%, el riesgo relativo de enfermedad es 50% menor si se usa el tratamiento B que si se usa el tratamiento A (riesgo relativo = riesgo en el grupo control / riesgo en el grupo experimental / riesgo en el grupo control; RR =  $0.40 / 0.20 = 0.50$  o 50%). La reducción del riesgo absoluto se calcula midiendo la diferencia de la tasa de enfermedad en el grupo control y en el grupo experimental (riesgo absoluto = riesgo en el grupo control - riesgo en el grupo experimental; RA =  $0.40 - 0.20 = 0.20$  o 20%). Es decir, el riesgo absoluto de la enfermedad es 20% menor si se toma el tratamiento B que si se toma el tratamiento A.

Es importante señalar que aun cuando la selección, aleatorización y cegamiento hayan sido adecuados, las diferencias en la tasa de enfermedad pueden no ser debidas al tratamiento recibido. Factores que pueden confundir los resultados incluyen el que, después de aleatorizar a un paciente a recibir la intervención, se descubra que no tenía la enfermedad. También puede ocurrir que el paciente no haya tomado el medicamento como se le prescribió, que haya dejado de tomarlo, o que haya decidido añadir otro medicamento que también es efectivo para la enfermedad que se está estudiando. Por último, también puede haber "contaminación", es decir, que un paciente inicialmente asignado al grupo control se pase al grupo experimental, y viceversa. La presencia frecuente de estos factores hace que la interpretación de los resultados sea difícil.

### ANÁLISIS DE INTENCIÓN DE TRATAR

Como en la práctica clínica, es frecuente que los pacientes descontinúen o cambien el tratamiento que se les ha prescrito, se ha propuesto que se realice el análisis de los ECC de acuerdo con la asignación inicial del tratamiento, independientemente que algunos pacientes no lo hayan tomado como se les indicó. Esta estrategia de *análisis con intención de tratar* (*intent to treat analysis*) tiene la ventaja de ofrecer resultados más cercanos a lo que ocurre en la práctica clínica diaria, pero tiene la desventaja de que se puede obtener un resultado falso negativo si un número grande de pacientes no tomaron el medicamento como se les prescribió. Obviamente, el análisis con intención de tratar sería inadecuado si lo que se quiere determinar es la eficacia de un medicamento; es decir, de si produce el efecto que supuestamente

debe producir o no. En tal caso, se debe asegurar que los pacientes que fueron asignados a recibir el tratamiento experimental y el control lo hicieron adecuadamente.

### ■ ANÁLISIS DE ENSAYOS CLÍNICOS CRUZADOS

La medición de los resultados en ECC de tipo cruzado es más compleja. En el estudio que se mencionó anteriormente sobre el efecto de la acarbose en pacientes que fallaron al tratamiento previo con mezclas de hipoglucémiantes orales, se realizó una curva de tolerancia a la glucosa oral y se hicieron mediciones de resistencia a la insulina al inicio y final de los periodos de 12 semanas. Antes de realizar el análisis para ver si la acarbose es más efectiva que el placebo, se tuvo que verificar que el periodo de lavado fue lo suficientemente largo para que no hubiera efecto residual al inicio del tratamiento en el segundo periodo. También se evaluó el efecto del tiempo transcurrido sobre variables que pudieran alterar los resultados finales. Por ejemplo, los pacientes podrían haber ganado o perdido peso durante las 27 semanas del estudio, lo cual podría haber influenciado la medición de resistencia a la insulina al final del segundo periodo. Por tanto, en un ECC cruzado no se deben analizar los resultados hasta que no se demuestre ausencia de efecto de secuencia o tiempo, pues de otro modo no se puede descartar que los resultados finales sean erróneos. Si se demuestra que hubo un problema de secuencia o de tiempo, el estudio se puede analizar como si fuera paralelo, aunque se corre el riesgo de perder poder estadístico.

### FASES DE LA INVESTIGACIÓN FARMACOLÓGICA

La investigación farmacológica típicamente empieza con la síntesis química de un medicamento y continúa con la realización de estudios, primero en animales y después en humanos. En la etapa preclínica, la cual puede durar hasta 3-6 años, se realizan estudios en animales para determinar la farmacocinética y la farmacodinamia del medicamento. Específicamente, se evalúa la seguridad del fármaco mediante la determinación de la dosis letal 50, la dosis tóxica y la dosis terapéutica. En la etapa clínica de investigación en humanos hay 4 fases. La investigación de fase I tiene como objeto establecer el rango de dosis, única o repetida, que pueda ser tolerada y que es segura; es decir, libre de efectos secundarios graves y frecuentes. Generalmente se realiza en un grupo pequeño de sujetos sanos, sin un grupo control. Los estudios fase II tienen por objeto hacer una evaluación inicial de la eficacia del medicamento y determinar la dosis terapéutica. Generalmente se efectúan en un grupo pequeño de pacientes que padecen la enfermedad para la que supuestamente es eficaz el tratamiento.

Los estudios de fase III tienen por objeto probar de forma definitiva que el medicamento es eficaz y descubrir la presencia de efectos secundarios comunes. Generalmente se realizan en un grupo grande de pacientes que tienen la enfermedad y cumplen con los requerimientos de un ECC riguroso. De los estudios de fase III se obtiene la mayor parte de la información disponible en la literatura sobre un medicamento; también suelen aportar la evidencia necesaria para que las agencias reguladoras aprueben la venta del medicamento en las farmacias. Una vez que se permite la venta del medicamento, se continúan realizando estudios de fase IV para detectar efectos secundarios inusuales o menos frecuentes. En esta fase se pueden sugerir nuevas indicaciones, que requerirán de nuevos estudios de fase II o III. Por ejemplo, aunque la indicación inicial del minoxidil fue el tratamiento de la hipertensión arterial, en estudios de fase IV se observó que producía crecimiento del vello, por lo que posteriormente se empezó a utilizar para tratar la calvicie.

### CONSIDERACIONES ÉTICAS EN LOS ENSAYOS CLÍNICOS CONTROLADOS

Es necesario que un Comité de Ética independiente valore el protocolo de investigación. El Comité debe determinar si es ético la aplicación en humanos de la intervención que se pretende estudiar y si la alternativa terapéutica en el grupo control es adecuada. La participación de los sujetos debe ser voluntaria y bajo consentimiento informado. Para ello, se le debe informar al paciente sobre el objetivo del estudio, los efectos adversos del medicamento y los beneficios que se esperan obtener; también se le debe aclarar que, si decide no participar, continuará recibiendo atención médica. Un grupo independiente de investigadores debe monitorizar los resultados durante la realización del estudio; si se observa que una de las alternativas terapéuticas es ineficaz o que está produciendo daño a los pacientes, se debe aconsejar la terminación temprana del estudio.

### CONCLUSIONES

Los ensayos clínicos controlados son un tipo especial de estudio longitudinal de cohortes en los que el investigador generalmente asigna y compara la eficacia de dos tratamientos. El grupo que recibe el tratamiento experimental se compara con un grupo control que puede ser histórico, concurrente (diseño paralelo), o el mismo grupo experimental que recibe ambas terapias con una secuencia predeter-

# Interpretación de ensayos clínicos negativos o sin diferencia

*Francisco López-Jiménez  
David Paniagua  
Gervasio A. Lama*

## INTRODUCCIÓN

Un ensayo clínico negativo es aquel en el que no se encuentra una diferencia significativa entre los grupos comparados. En los últimos años se ha incrementado el número de publicaciones de ensayos clínicos negativos.<sup>1</sup> Un ensayo clínico sin diferencia estadística puede brindar información útil para el clínico, ya sea para descartar un tratamiento ineficaz o para demostrar que una intervención es tan efectiva como aquélla a la que se le comparó. Desechar un tratamiento inútil es plausible. Sin embargo, si uno o varios estudios con errores metodológicos sugieren erróneamente que cierto tratamiento no brinda beneficio clínico, es posible que se prive a pacientes de nuevas intervenciones que en realidad son eficaces. Por ejemplo, para finales de los años 70 se habían realizado varios estudios clínicos pequeños para evaluar la efectividad de la estreptoquinasa en el infarto agudo del miocardio, pero ninguno demostró utilidad terapéutica.<sup>2</sup> Esos hallazgos desfervecieron por algún tiempo el interés por la estreptoquinasa. En 1986 se

**Tabla 10-1. Lista de factores a evaluar al analizar un ensayo clínico sin diferencia estadística entre los grupos estudiados**

- ¿Se tuvo un poder estadístico adecuado?
- ¿Se seleccionó un objetivo clínico adecuado?
- ¿Fue la intervención insuficiente o inapropiada?
- ¿Existen datos que sugieran contaminación del estudio?
- ¿Se sobreutilizó el análisis de "intención de tratar"?
- ¿Fue suficiente el grado de adherencia al tratamiento experimental?
- ¿Se hizo un seguimiento clínico insuficiente o demasiado prolongado?
- ¿Existían condiciones clínicas que competían en riesgo con la enfermedad en estudio?
- ¿Se descartó que hubiera clasificación errónea del resultado en estudio (no sistemática)?
- ¿Fue la aleatorización adecuada?
- ¿Existían intervenciones paralelas compensatorias?
- En estudios no cegados, ¿existía el riesgo de compensación de riesgo?
- ¿Se descartó un seguimiento sesgado?

publicó el primer ensayo clínico a gran escala evaluando estreptoquinasa vs placebo, el cual demostró una reducción de 20% del riesgo de muerte hospitalaria.<sup>3</sup> De no haberse realizado estudios apropiados, uno de los pilares en el manejo del infarto agudo del miocardio, para estas fechas, sería todavía considerado como ineficaz.

Durante la lectura crítica de publicaciones de ensayos clínicos controlados en los que no se encontró diferencia, se debe evaluar los posibles errores metodológicos que conllevan a un hallazgo falso negativo. A continuación se enumeran algunas de las más importantes (tabla 10-1).

## CAUSAS POSIBLES DE UN RESULTADO SIN DIFERENCIA

### ■ PODER ESTADÍSTICO INADECUADO

Para definir si los pacientes con infarto agudo del miocardio de cara inferior deberían ser hospitalizados, en 1978 se realizó un ensayo clínico controlado donde se comparó manejo hospitalario con manejo ambulatorio.<sup>4</sup> El estudio no mostró diferencia en mortalidad o complicaciones y los autores concluyeron que los pacientes con infarto inferior podrían ser tratados de manera ambulatoria. Desgraciadamente, las conclusiones ahí expresadas tuvieron poco fundamento, pues el estudio carecía de lo que se conoce como poder estadístico.

Poder estadístico es la probabilidad de que un estudio clínico encuentre una diferencia real entre los grupos comparados. El principal determinante del poder estadístico es el número de eventos clínicos a estudiar. Ya que el número de eventos clínicos está determinado tanto por las características de los pacientes estudiados, como por el número de pacientes incluidos en el estudio, a mayor el número de pacientes, mayor el poder estadístico. Otro determinante es el grado de variabilidad en la respuesta al tratamiento, en los casos en que el objetivo se expresa con variables continuas como presión arterial, peso corporal, etc. A menor variabilidad en la respuesta, mayor el poder estadístico. Un poder estadístico inadecuado es quizá la causa más común de ensayos clínicos negativos. En un análisis de estudios negativos publicados en revistas médicas, se demostró que 84% de los estudios con resultados negativos no tuvieron un poder estadístico adecuado para demostrar una diferencia de 25% en el objetivo estudiado.<sup>5</sup>

### ¿Cómo saber si un estudio negativo tuvo un poder estadístico adecuado?

Si en un estudio negativo los autores no hacen referencia a un tamaño inapropiado de la muestra, como una posibilidad que explique la falta de diferencia entre los tratamientos en estudios, queda a la acuciosidad del lector determinar si se incluyeron suficientes pacientes para demostrarla. En el Apéndice se proporciona una fórmula simple para calcular el número mínimo de pacientes que un estudio debe incluir para tener un poder estadístico adecuado. Si durante la lectura crítica de un estudio sin diferencia se demuestra que el ensayo tuvo un poder estadístico adecuado, se debe proceder a descartar otros factores que pueden enmascarar una diferencia.

### ■ SELECCIÓN DE UN OBJETIVO INADECUADO

Para demostrar la efectividad de la prevención secundaria de enfermedad coronaria con medicamentos hipolipemiantes, ensayos clínicos preliminares utilizaron como objetivo la regresión de la obstrucción coronaria, evaluada mediante coronariografía.<sup>6,7</sup> Los resultados fueron desalentadores, pues los hipolipemiantes no lograron una regresión angiográficamente significativa. Estudios clínicos posteriores donde se utilizaron objetivos clínicos, como muerte cardiovascular o infarto agudo del miocardio, demostraron una reducción importante de la incidencia de eventos clínicos.<sup>8,9</sup> La aparente discrepancia entre los resultados angiográficos y clínicos sugiere que el efecto benéfico no se limita, o no está directamente relacionado a cambios angiográficos y que puede ser resultado, entre otras cosas, de la mejoría en la función endotelial de las arterias coronarias.<sup>10</sup> De haberse limitado la investigación clínica a subrogados diagnósticos, se habría concluido erróneamente que los hipolipemiantes no tienen efecto clínico benéfico (tabla 10-2).

**Tabla 10-2. Lista de subgrupos inadecuados, que históricamente han fallado en predecir resultados clínicos de intervenciones médicas**

Subgrupo	Evento que pretendía predecir
Grado de estenosis coronaria	Eventos coronarios futuros
Mejoría en células CD4 en VIH+	Mayor supervivencia
Reducción de arritmias ventriculares	Menor muerte súbita y mayor supervivencia
Disminución de colesterol sérico con fibratos	Reducción de eventos coronarios

## ■ INTERVENCIÓN INSUFICIENTE O INAPROPIADA

En un estudio reciente donde se evaluó la efectividad de apoyo psicológico como medida de prevención secundaria en el infarto del miocardio, no se encontró diferencia en mortalidad o eventos clínicos entre el grupo con intervención y el grupo con manejo convencional.<sup>11</sup> A primera vista, los resultados sugieren que el mejoramiento en el bienestar emocional de los pacientes no beneficia el pronóstico de los sujetos que han tenido un infarto agudo del miocardio. Al revisar el estudio más a fondo, se puede encontrar que los pacientes con apoyo psicológico no sólo no mejoraron su pronóstico, sino que su estado emocional tampoco mejoró y que, paradójicamente, ¡su nivel de depresión aumentó! Este estudio sólo puede concluir que, la clase de apoyo psicológico utilizada no fue la adecuada para mejorar el estado emocional de los pacientes, y la hipótesis de que, mejorando el bienestar emocional se mejora el pronóstico de los pacientes después de un infarto, quedó por el momento sin respuesta.

En otras ocasiones, la falta de diferencia entre dos tratamientos puede ser consecuencia de dosis insuficiente del medicamento. En un estudio que evaluó la utilidad clínica de beta carotenos para prevenir infarto del miocardio, no se encontró beneficio alguno.<sup>12,13</sup> Una de las principales críticas de tal estudio fue la dosis insuficiente de carotenos que se utilizó, que fue muy inferior de la que en estudios experimentales sugería tener un efecto antioxidante aceptable.<sup>14</sup> Estudios con dosis mayores están actualmente en desarrollo, por lo que las conclusiones actuales se limitan a que los beta carotenos no son efectivos para prevenir el IAM a la dosis en que se investigaron, que probablemente fue insuficiente.

## ■ CONTAMINACIÓN DEL ESTUDIO

Idealmente, en un ensayo clínico controlado sólo los pacientes asignados a la intervención experimental son los que reciben ésta. En algunas ocasiones, sin embargo, sucede que un número considerable de pacientes que se asignaron al grupo control también reciben la intervención experimental, ya sea por que los

médicos a cargo están convencidos de las ventajas no demostradas de la terapia en estudio o porque los pacientes se las arreglan para no arriesgarse a estar en un grupo placebo. En un estudio de pacientes con SIDA (comunicación personal, M. D. Hugges), se encontró que algunos pacientes intercambiaban entre ellos la mitad de las dosis de los medicamentos que les fueron asignados, para así asegurarse de que, en caso de estar en el grupo placebo, al menos recibirían la mitad de la dosis experimental. Otro ejemplo es el estudio SOLVD,<sup>15</sup> en el que no se disminuyó la mortalidad en pacientes con insuficiencia cardíaca al tratarlos con enalapril. En este estudio, 19% de los pacientes asignados a placebo recibieron también un inhibidor de la enzima convertidora de angiotensina.

Al mezclarse la intervención entre los pacientes asignados y no asignados a ella, se disminuye el poder estadístico del estudio, y por ende la posibilidad de encontrarse una diferencia entre los grupos.<sup>16</sup> La contaminación es más probable en ensayos clínicos donde el tratamiento no es fácil de cegar y, donde la aceptación o popularidad del tratamiento experimental, hace que los médicos o los pacientes teman perderse el beneficio de la terapia en estudio. Este apartado está estrechamente ligado a la sobreutilización del análisis de intención a tratar que se menciona a continuación.

## ■ SOBREUTILIZACIÓN DEL ANÁLISIS DE INTENCIÓN DE TRATAR

El análisis de los datos de un ensayo clínico controlado regularmente comprende dos aspectos fundamentales:

- La efectividad global de la maniobra. Los pacientes se analizan de acuerdo a la intervención asignada, independientemente de si la recibieron o no, del grado de apego a ella, o si completaron el tiempo requerido o la suspendieron tempranamente. Este análisis se le conoce como análisis de intención de tratar.<sup>17</sup>
- El segundo aspecto a evaluar es la eficacia del tratamiento, donde se valora la respuesta a la maniobra únicamente en los pacientes que, después de la aleatorización, sí recibieron, toleraron y completaron el tratamiento. Un estudio clínico puede carecer de hallazgos significativos al hacer el análisis de acuerdo a la intención de tratar, pero tener diferencias significativas al analizar el efecto sólo en los que sí recibieron el tratamiento completo. Las razones de tal discrepancia no siempre pueden ser atribuibles a un tratamiento difícil de ser aceptado, tolerado o administrado, sino que pueden ser debidas a prejuicios o idiosincrasia de los pacientes estudiados, o a la logística y organización del estudio. Un estudio donde el análisis de intención de tratar sea negativo pero que muestre clara ventaja para el grupo de pacientes que sí recibieron y completaron el tratamiento puede proporcio-



nar información útil para el clínico al mostrar cuál puede ser la respuesta al tratamiento en algunos pacientes en particular.

### ■ FALTA DE ADHERENCIA AL TRATAMIENTO EXPERIMENTAL

En algunas ocasiones, los pacientes no toman el tratamiento experimental o lo toman de manera irregular, particularmente cuando el régimen a seguir es complejo, cuando el tratamiento experimental produce efectos colaterales perceptibles por el enfermo y cuando existe poca motivación del paciente para participar en el estudio. La falta de apego al tratamiento experimental disminuye la capacidad de un estudio para encontrar una diferencia entre dos tratamientos.<sup>18</sup> En un estudio que no demuestra diferencia entre los grupos tratados, se espera encontrar en el artículo una descripción de cómo se evaluó la adherencia al tratamiento. Ordinariamente se hace contando el medicamento sobrante o verificando la droga o sus metabolitos en el suero o la orina de los pacientes.<sup>19</sup>

### ■ SEGUIMIENTO INSUFICIENTE O EXAGERADO

El tiempo necesario para que una intervención experimental tenga el efecto esperado puede ser impredecible. Si el análisis de los resultados se realiza prematuramente, es posible que el efecto de la intervención no se detecte. Por ejemplo, en algunos estudios, que comparan intervenciones médicas con quirúrgicas, un seguimiento breve puede desfavorecer la cirugía, pues en muchos casos ésta se asocia a efectos adversos y complicaciones a corto plazo, mientras que su beneficio real puede ser perceptible tiempo después.<sup>20</sup>

De manera contraria, cuando el seguimiento es muy largo, se puede disminuir la probabilidad de mantener una diferencia entre el efecto de los tratamientos. Por ejemplo, en un estudio que evalúe el efecto de un medicamento antihipertensivo en la mortalidad global a largo plazo, nos encontramos que eventualmente todos los pacientes morirán (figura 10-1).

### ■ COMPETENCIA EN EL RIESGO

Entre más comorbilidades tengan los pacientes que se incluyan en un estudio, mayor el riesgo de que mueran por causas ajenas a la enfermedad en estudio, limitando así las posibilidades que se demuestre una diferencia clínicamente significativa entre los grupos tratados. Por ejemplo, entre más ancianos sean los pacientes incluidos en un estudio de hipertensión arterial y mortalidad, mayor el número de enfermedades y de factores de riesgo cardiovascular que compiten con la

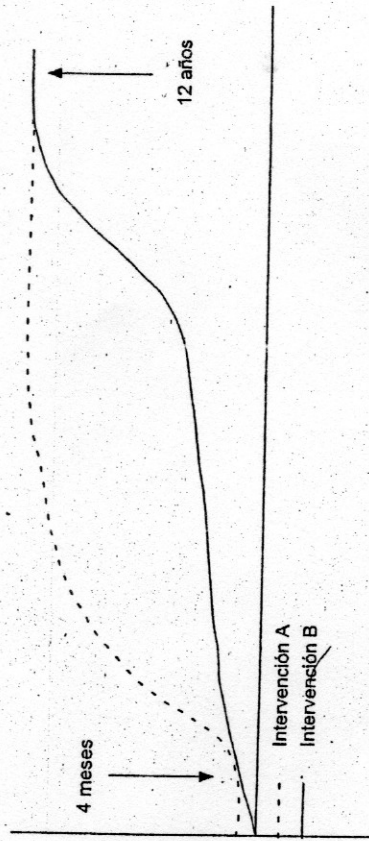


Figura 10-1. Representación gráfica de un seguimiento insuficiente (4 meses) y otro demasiado prolongado (12 años) que pueden enmascarar diferencias clínicas importantes entre dos intervenciones.

hipertensión arterial elevada como causa de muerte. Aun, un tratamiento eficaz en reducir la tensión arterial y en disminuir el número de eventos clínicos relacionados a ésta, tendría dificultades en mostrar una diferencia en mortalidad.

### ■ CLASIFICACIÓN ERRÓNEA DEL RESULTADO EN ESTUDIO

En cualquier estudio clínico, es posible que los pacientes se clasifiquen erróneamente de acuerdo a la ocurrencia o no del resultado a evaluar. Este problema requiere especial atención cuando el resultado es susceptible a gran variabilidad intra o interobservador. Si la clasificación errónea sucede en un número significativo de pacientes y no es sistemática, se corre el riesgo de no detectar una diferencia real entre los grupos en estudio.<sup>21</sup> En un estudio negativo, en el que el objetivo a evaluar no sea mortalidad total, se debe buscar la manera en que se definieron y evaluaron los eventos a comparar. También se debe buscar la variabilidad inter o intraobservador al clasificar los eventos, la cual debe ser mínima. De esta forma, se puede considerar que una clasificación errónea no fue la causa de la ausencia de diferencia entre los grupos en estudio.

### ■ ALEATORIZACIÓN INADECUADA

Al asignar la intervención mediante el azar se pretende lograr dos o más grupos con características similares. Sin embargo, cuando el número de pacientes es limitado, la distribución de ciertas características puede quedar desbalanceada y

provocar resultados erróneos. Si el grupo de pacientes a los que se asignó el tratamiento en estudio tiene más características desfavorables, es posible que esto enmascare un efecto benéfico de la maniobra. En un estudio sin diferencia se debe determinar si las características que pueden estar asociadas a un resultado adverso, se distribuyeron equitativamente entre los grupos.

### ■ INTERVENCIONES PARALELAS COMPENSATORIAS

En un estudio clínico a largo plazo, en el que la intervención produce una mejora tanto en síntomas como en estudios de laboratorio, es posible que los pacientes en el grupo control reciban otras intervenciones paralelas de manera compensatoria. Por ejemplo, en un estudio que evalúe dolor articular en pacientes con artritis reumatoide, es posible que los pacientes que no reciben el beneficio de un tratamiento experimental útil, de cualquier manera intenten disminuir su grado de dolor ingiriendo otros analgésicos. Esto ocasiona que desaparezca la diferencia en el grado de dolor entre los grupos estudiados. Cuando existe la posibilidad de tener intervenciones paralelas compensatorias que puedan hacer imperceptible una diferencia, se debe buscar en el artículo información detallada del uso de tales intervenciones en cada grupo o incluirlas como parte del resultado a evaluar.

### ■ COMPENSACIÓN DE RIESGO

La adherencia a intervenciones preventivas es motivada en gran parte por el temor a sufrir la consecuencia a prevenir. Al mismo tiempo, la confianza que genera estar bajo cierto tratamiento o medida preventiva, teóricamente puede inducir al individuo a exponerse a riesgos que de otra manera no tomaría. Por ejemplo, se demostró que después de hacerse obligatoria la utilización de cascos en motociclistas, la velocidad promedio de circulación entre ellos aumentó. Similarmente, después de reglamentarse el uso de cascos entre beisbolistas, los lanzadores comenzaron a enviar la pelota más cerca de los bateadores.<sup>22</sup> En un estudio donde se evaluó el efecto de consejo médico después de efectuarse la prueba de VIH, la incidencia de gonorrea se triplicó en los 6 meses posteriores a la prueba en la gente cuyo resultado de VIH fue negativo.<sup>23</sup> La compensación de riesgo puede eliminar los beneficios de las intervenciones preventivas. Es posible que en ensayos clínicos donde el tratamiento no sea cegado, se presente la compensación de riesgo y que se atenúe o desaparezca la diferencia entre los grupos en estudio.

### ■ SEGUIMIENTO SEGGADO

En un ensayo clínico donde se compara mortalidad a largo plazo, el análisis de supervivencia mediante las curvas de Kaplan-Meier y el análisis Cox de riesgos

proporcionales asumen una pérdida no sesgada de los sujetos durante el seguimiento.<sup>24</sup> Esto significa que las razones por las que se excluyeron los pacientes del seguimiento fueron similares entre los grupos a comparar. Sin embargo, suponemos que en un ensayo clínico controlado los pacientes tratados con la cirugía X tienen un pronóstico real muy superior que los pacientes tratados con el medicamento Y. Subsecuentemente, una buena parte de los pacientes tratados quirúrgicamente tienen mejoría substancial y dejan de ir a sus consultas de seguimiento, mientras que los pacientes con resultados menos favorables asistieron a sus consultas médicas durante más años. Este fenómeno puede hacer que desaparezca una diferencia clínicamente significativa al analizar la supervivencia de los pacientes en ambos grupos. En un estudio negativo no es sencillo determinar si la pérdida de pacientes en el seguimiento fue sistemática. Los autores deben dejar claro que el seguimiento tuvo la misma intensidad en ambos grupos y que fue cegado al tratamiento.

## ESTUDIOS DE EQUIVALENCIA

En ocasiones es importante demostrar que un nuevo tratamiento es tan eficaz como el estándar, particularmente cuando el nuevo tratamiento es más barato, más fácil de administrar o tiene menos efectos colaterales que el tratamiento de referencia. Sin embargo, concluir que dos intervenciones tienen la misma eficacia o que cierta intervención no es mejor que placebo, es un proceso relativamente complejo. En investigación clínica no es lo mismo ausencia de evidencia que evidencia de ausencia o, en otras palabras, si un estudio controlado no demuestra diferencia estadísticamente significativa, no significa que la efectividad de los tratamientos comparados es similar.<sup>25</sup> En un estudio negativo que fue diseñado explícitamente para encontrar diferencia entre los grupos tratados, únicamente se puede afirmar que tal estudio no pudo demostrar una diferencia estadísticamente significativa. Un estudio diseñado para probar equivalencia, tiene que demostrar que el poder estadístico fue adecuado y que no se incurrió en ninguno de los errores en la interpretación y análisis mencionados anteriormente, antes de aceptar sus conclusiones como válidas. El principal inconveniente de un estudio de equivalencia es que requieren un mayor número de pacientes que los estudios diseñados para encontrar diferencias, lo que los hace más costosos y más difíciles de realizar que los estudios para mostrar diferencia.

## CONCLUSIONES

Este capítulo no pretende convencer de que, todos o la mayoría de los estudios negativos no pudieron demostrar una diferencia real sólo por problemas metodoló-

gicos. Muchos estudios prometedores no pasan nunca de ser promesas. De hecho, de manera inversa, algunas intervenciones que nunca han demostrado científicamente ser de beneficio clínico, pero que se han usado empíricamente por siglos, necesitan múltiples estudios de gran escala con resultados negativos para convenecer de su ineficacia.<sup>26</sup> Las conclusiones de un estudio negativo se deben tomar con cautela. La utilidad de la información que proporcionan depende en parte de la calidad metodológica con la que se realizan, de la plausibilidad biológica de los resultados y de la evidencia científica existente a la que sean confrontados.

## APÉNDICE

Fórmula para determinar el número mínimo de pacientes requeridos en un ensayo clínico para tener un poder estadístico de al menos 80%:

$$n = \frac{p_1 * (100 - p_1) + p_2 * (100 - p_2)}{(p_2 - p_1)^2} \times 7.9$$

n = Número de pacientes requeridos en cada grupo a comparar, p1= Porcentaje estimado de pacientes con respuesta favorable en el grupo control; p2= Porcentaje estimado de pacientes con respuesta favorable en el grupo experimental. ■

## BIBLIOGRAFÍA

8. Shepherd J, Cobbe SM, Ford I *et al.* Prevention of Coronary Heart Disease with pravastatin in men with hypercholesterolemia. *N Engl J Med* 1995; 333: 1301-7.
9. Sacks FM, Pfeffer MA, Moye LA *et al.* The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. *N Engl J Med* 1996; 335: 1001-9.
10. Ludmer PL, Selwyn AP, Shook TL. Paradoxical vasoconstriction induced by acetylcholine in atherosclerotic coronary arteries. *N Engl J Med* 1986; 315: 1046-51.
11. Frasure-Smith N, Lespérance F, Prince RH *et al.* Randomized trial of home-based psychosocial nursing intervention for patients recovering from myocardial infarction. *Lancet* 1997; 350: 473-9.
12. Hennekens GH, Buring JE, Manson JE *et al.* Lack of effect of long term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N Engl J Med* 1996; 334: 1145-9.
13. Omenn GS, Goodman GE, Thornquist MD *et al.* Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med* 1996; 334: 1150-5.
14. Doering WVE. Antioxidant vitamins, cancer and cardiovascular disease. *N Engl J Med* 1996; 335: 1065-6.
15. The SOLVD Investigators. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. *New Engl J Med* 1992; 327: 685-91.
16. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997; 16: 1017-29.
17. Pocock SJ. *Clinical Trials. A practical approach.* New York, John Wiley & Sons, Chichester, 1993.
18. MacFadyen RJ, Struthers AD. The practical assessment of compliance with ACE-inhibitor therapy: a novel approach. *J Cardiovasc Pharm* 1997; 29: 119-24.
19. Costa FV. Compliance with antihypertensive treatment. *Clin & Exper Treat* 1996; 18: 463-72.
20. Howard G, Chambliss LE, Kronmal RA. Assessing differences in clinical trials comparing surgical vs nonsurgical therapy. *JAMA* 1997; 278: 1432-36.
21. Hennekens CH, Buring JE. Intervention studies. In: Mayrent SL (ed.). *Epidemiology in Medicine.* Boston, Little Brown, 1987: 178-212.
22. Hemenway D. Risk Compensation. In: Hemenway D (ed.). *Prices and choices. Microeconomic Vignettes.* New York, University Press of America, 1993: 213-20.
23. Zenilman JM, Erickson B, Fox R, Reichart CA, Hook EW. Effect of HIV post-test counseling on STD incidence. *JAMA* 1992; 267: 843-5.
24. Feinstein AR. Proportional hazards analysis. In: Feinstein AR (ed.). *Multivariable Analysis. An introduction.* New Haven, Yale University Press, 1996: 370-97.
25. Ware JH, Antman EM. Equivalence trials. *N Engl J Med* 1997; 337: 1159-61.
26. Bonfill X, Gabriel R, Cabello J. La medicina basada en la evidencia. *Rev Esp Cardiol* 1997; 50: 819-25.

1. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994; 272: 122-4.
2. Lau J, Antman E, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of the therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327: 248-54.
3. GISSI. Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1986; 1: 397-401.
4. Hill JD, Hampton JR, Mitchell JRA. A randomized trial of home versus hospital management for patients with suspected myocardial infarction. *Lancet* 1978; 1: 837-41.
5. Freiman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error, and sample size in the design and interpretation of the randomized controlled trial: survey of 71 'negative' trials. *N Engl J Med* 1978; 299: 426-32.
6. Leaf A. Management of hypercholesterolemia: are preventive intervention advisable? *N Engl J Med* 1989; 321: 680-4.
7. Brett AS. Treating hypercholesterolemia: how should practicing physicians interpret the published data for patients? *N Engl J Med* 1989; 321: 676-80.