

III

EL ENSAYO CLÍNICO CONTROLADO

El ensayo clínico es un experimento planificado. Su objetivo es evaluar la eficacia de intervenciones sanitarias, médicas o quirúrgicas. En un ensayo clínico el primer paso es la formulación de los objetivos o, dicho de otro modo, de la pregunta que se desea contestar. Esta pregunta será el principal determinante de sus aspectos metodológicos, como los criterios de inclusión y exclusión de los pacientes, el tipo de ensayo clínico, el número de pacientes incluidos en el ensayo, su duración y los parámetros o variables clínicas que se van a medir. En general, es preferible que el número de preguntas sea limitado, y que éstas sean muy concretas: "una respuesta aproximada a un problema bien formulado es mucho más valiosa que una respuesta exacta a un problema aproximado".¹

En la figura 4 se resumen las fases de un ensayo clínico. En primer lugar, se selecciona una muestra de la verdadera población de pacientes con la enfermedad de interés. A continuación se realiza la asignación aleatoria de los participantes a cada una de las modalidades de tratamiento. En tercer lugar, se aplican las intervenciones previstas a los grupos de tratamiento y en cuarto lugar se examina a los pacientes con el fin de conocer los resultados de las intervenciones comparadas. En cada una de estas fases, conceptualmente sencillas, hay posibilidades de introducir sesgos que distorsionen la impresión que el investigador puede hacerse de cada uno de los tratamientos probados. De ahí que la consideración de cada uno de estos procesos deba ser cuidadosamente valorada antes del comienzo del ensayo.

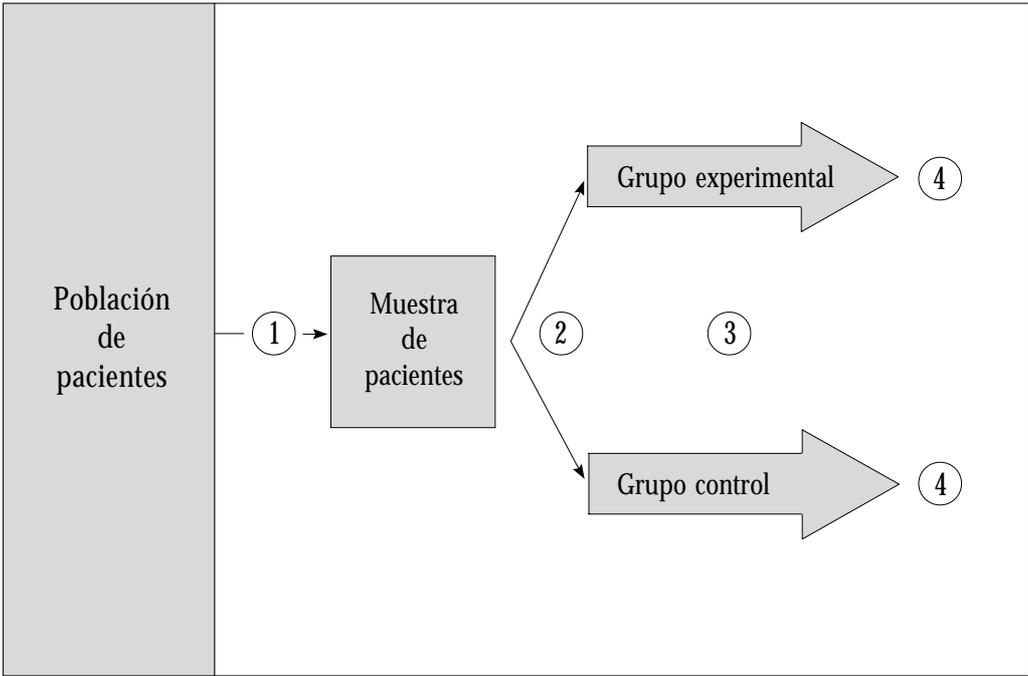


Figura 4. Las diferentes fases conceptuales de un ensayo clínico: (1) en primer lugar se selecciona la muestra de pacientes a partir de una población de referencia; (2) a continuación se procede a la distribución aleatoria; (3) seguidamente se aplican las intervenciones previstas y después (4) se miden los resultados. En cada una de estas fases se deben considerar cuidadosamente los posibles mecanismos por los que se puede incurrir en algún sesgo.

LOS PILARES DEL ENSAYO CLÍNICO

A continuación se comentarán los pilares básicos del ensayo clínico: la determinación del **origen y tamaño de la muestra**, la **asignación aleatoria**, la **exposición a los tratamientos** y el **enmascaramiento**.¹³

Selección de los participantes y tamaño de la muestra

Un protocolo de ensayo clínico debe especificar la **fuentes** de la que se van a reclutar los participantes, así como los **criterios de inclusión y exclusión** de los pacientes. Además, antes del comienzo del ensayo se debe estimar cuántos participantes son necesarios.

La **fuentes u origen de los pacientes** puede ser un determinante de la representatividad de la muestra, así como de su homogeneidad. Así por ejem-

plo, si se desea hacer un ensayo clínico sobre el tratamiento de la insuficiencia arterial periférica, se puede identificar a la población participante en diferentes centros o niveles del sistema de atención a la salud. Si los participantes se reclutan en centros de atención primaria, será costoso identificar al número necesario (porque la mayoría de los pacientes que acuden a un centro de atención primaria no sufren arteriopatía periférica). Una alternativa puede ser acudir a un servicio hospitalario de angiología. La muestra finalmente incluida en el ensayo será diferente según de dónde procedan los participantes: será más representativa de la población de referencia de pacientes con insuficiencia arterial periférica si los participantes son reclutados en un centro de atención primaria; los pacientes reclutados en un servicio hospitalario de angiología tenderán a presentar formas más evolucionadas y graves de la enfermedad. De ahí que la fuente de los participantes pueda determinar la validez externa del ensayo (véanse el Capítulo V y *Validez externa* en el Glosario).

Los **criterios de inclusión** pueden ser más o menos restrictivos, según los objetivos del ensayo. Así por ejemplo, se pueden comparar los efectos de dos antiinflamatorios no esteroideos en pacientes que reúnan los criterios diagnósticos de artritis reumatoide de la Asociación Reumatológica Norteamericana, o en pacientes con "dolores reumáticos" sin más especificación (es evidente que hay otras posibilidades intermedias entre estos dos extremos). Con el primer tipo de criterios, la muestra estudiada será clínicamente muy homogénea, con lo que será más fácil detectar diferencias entre ambos fármacos, en caso de que las haya (véase más adelante la figura 5). Con criterios más amplios la muestra estudiada será más heterogénea, pero las conclusiones del ensayo serán aplicables a una población de referencia más amplia (véanse *Ensayo clínico explicativo*, *Ensayo clínico pragmático* y *Validez externa* en el Glosario).

Los **criterios de exclusión** también afectan simultáneamente a la homogeneidad de la muestra estudiada y a la validez externa del ensayo. Algunos criterios de exclusión se aplican con el fin de proteger (por exclusión del ensayo) a posibles participantes que puedan ser grupos de riesgo elevado o que puedan presentar contraindicaciones específicas a alguno de los tratamientos probados. Aunque eso está justificado en las primeras fases del estudio de un nuevo fármaco, también es cierto que en ocasiones una motivación aparentemente ética impide el estudio científico de la terapéutica en ciertas poblaciones. Así por ejemplo, durante mucho tiempo y salvo excepciones, los niños han sido generalmente excluidos de los ensayos clínicos por motivos "éticos", pero al impedir que los niños sean sujetos de investigación clínica controlada, en ocasiones se favorece que posteriormente se conviertan en sujetos de una práctica clínica incontrolada.

El **tamaño de la muestra** de pacientes que debe ser incluida en un ensayo clínico depende de cuatro factores: la homogeneidad de las poblaciones participantes, la cuantía de las diferencias que se pretende poner de manifiesto y los errores de tipo I y II que se consideran aceptables.

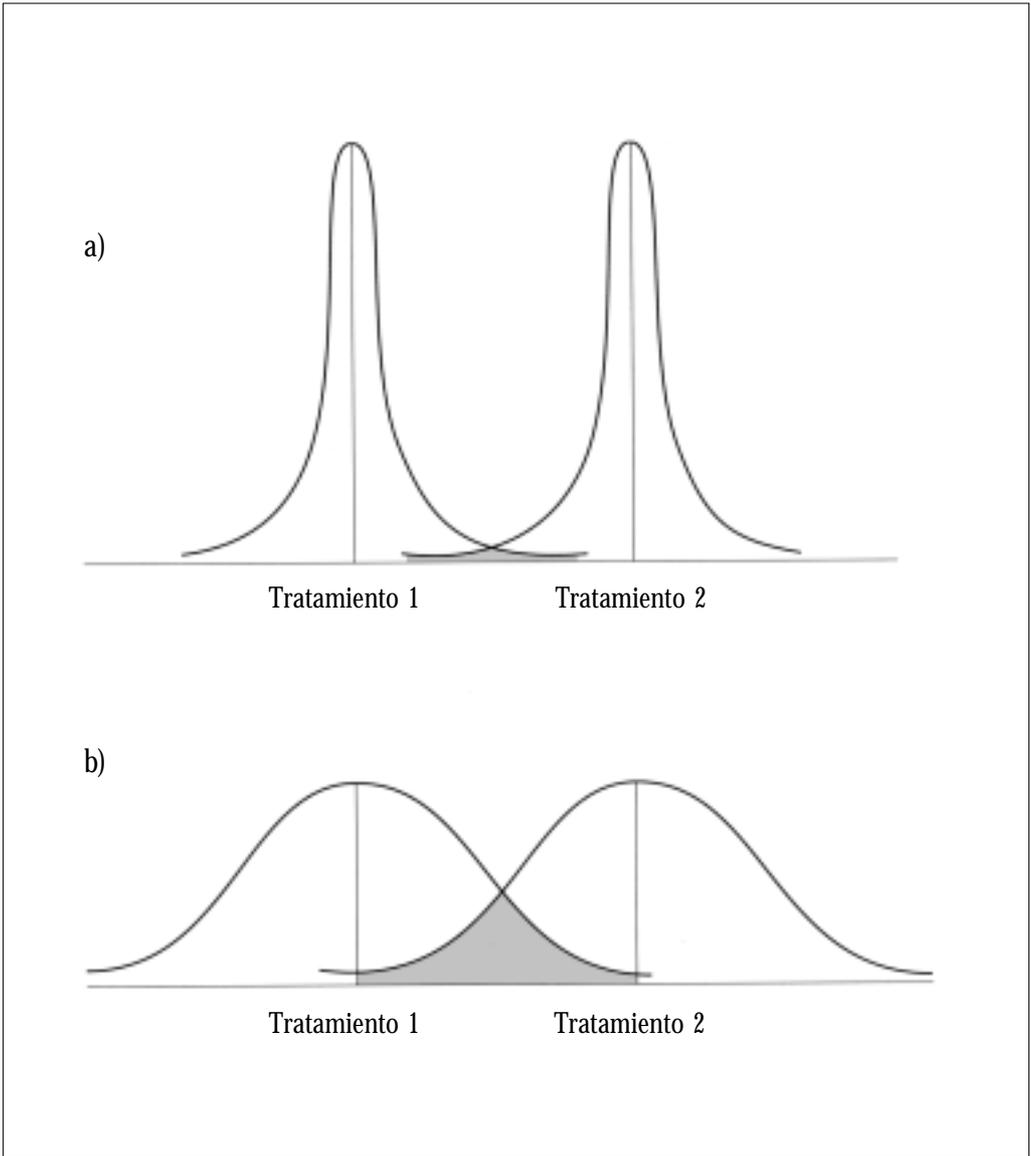


Figura 5. Distribuciones de frecuencias de valores de una variable continua. a) Los criterios de inclusión en el ensayo han sido restrictivos, con el fin de reunir una muestra de características homogéneas; la distribución de frecuencias de los valores de la variable continua (por ej., peso, presión arterial, etc.) se dispersa poco a partir de su valor medio y es relativamente fácil demostrar diferencias entre ambos grupos de tratamiento, porque, al ser homogéneas, las dos poblaciones se superponen en grado mínimo. b) Los criterios de inclusión en el ensayo han sido amplios, con el fin de reunir una muestra más representativa de la población de referencia; la diferencia entre los valores medios de ambos grupos es la misma que en el caso a, pero es más difícil demostrar que estas diferencias son estadísticamente significativas porque hay un grado notable de superposición en la respuesta de ambos grupos al tratamiento.

Las muestras de pacientes más homogéneas confieren mayor probabilidad de detectar pequeñas diferencias (véase la figura 5). Si la muestra de participantes es más heterogénea, para detectar diferencias entre ambos grupos será necesario incluir un mayor número de pacientes en el ensayo. [En la práctica esta consideración se traduce en que, con el fin de incluir una población homogénea, en muchos ensayos se tiende a aplicar criterios restrictivos de inclusión de pacientes; esto confiere mayor probabilidad de detectar pequeñas diferencias, pero a la vez invalida el carácter de "muestra" de los participantes, y en consecuencia plantea problemas de generalización de los resultados obtenidos; véase el Capítulo V.]

Análogamente, cuanto mayores sean las diferencias que se desee demostrar entre dos intervenciones, menor será el número necesario de pacientes. [Intuitivamente, es fácil comprender que es necesario un número más elevado de pacientes para descartar que en una enfermedad grave el tratamiento disminuye la mortalidad de 20% a 18%, que para descartar que la reduce de 20% a 10%; con menos pacientes se puede descartar ya una diferencia tan amplia como ésta.]

El **error de tipo I** consiste en rechazar la hipótesis nula cuando ésta es verdadera: en otras palabras, en afirmar que existen diferencias entre los dos grupos de tratamiento cuando en realidad éstas no existen. A la probabilidad de cometer un error de tipo I se la denomina α . Véase la figura 6: dado que en general no se sabe si $A > B$ o bien $B < A$, la posibilidad de cometer un error de tipo I es doble, por lo que se fija una probabilidad 2α . La mayoría de los ensayos clínicos se realizan previa fijación de una probabilidad de error $2\alpha = 0,05$.

El **error de tipo II** consiste en afirmar que no existen diferencias cuando éstas en realidad existen. La probabilidad de cometerlo se denomina β y el poder del estudio es la recíproca de esta probabilidad ($1-\beta$). Generalmente se fija el poder del estudio en 0,80 o en 0,90 (que corresponden, respectivamente, a valores de β de 0,20 o 0,10). (Véanse en el Glosario *Análisis de subgrupos, Dragado de datos, Hipótesis nula, P, Paradoja de Simpson, Poder estadístico, Prueba de significación y Significación estadística*).

La asignación aleatoria

Como ya se ha comentado, la asignación aleatoria define y diferencia el ensayo clínico controlado de los estudios de cohortes, porque es la única intervención metodológica que teóricamente da lugar a una distribución equilibrada de las características de los pacientes entre los diferentes grupos de tratamiento. En consecuencia, asegura que los grupos incluidos en el ensayo sean semejantes en todas las características relevantes menos en una: la intervención que cada uno recibe.

Con el fin de evitar sesgos en la formación de los diferentes grupos, la asignación se debe realizar después de que se haya comprobado que el paciente reúne los criterios de inclusión y exclusión preespecificados y de que, una vez

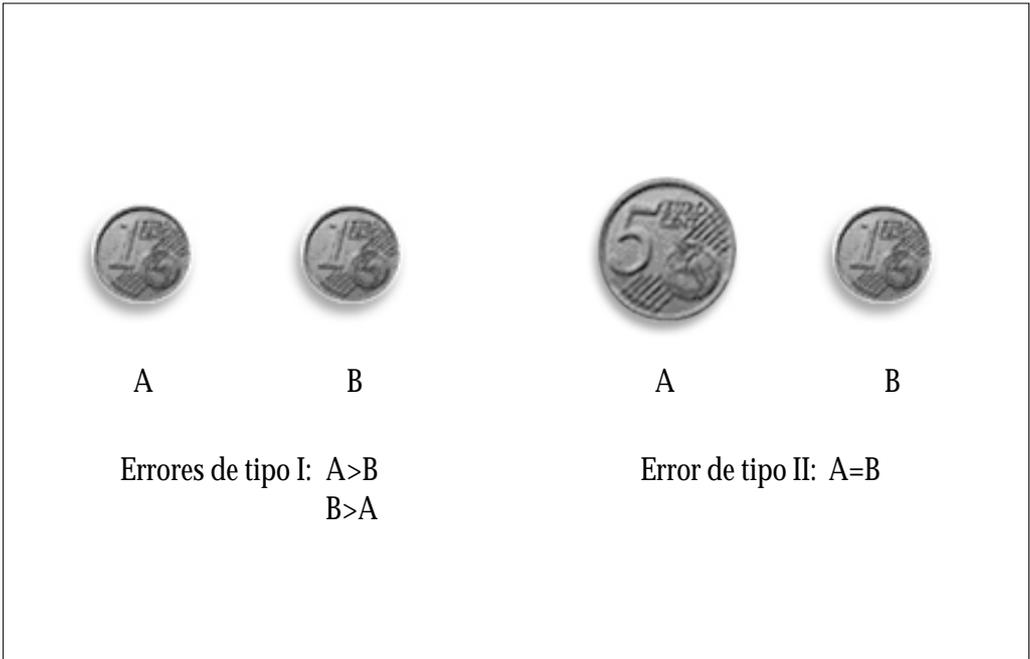


Figura 6. El error de tipo I consiste en rechazar una hipótesis nula verdadera, es decir, en afirmar que hay diferencias entre los grupos de tratamiento cuando en realidad no las hay. El error de tipo II consiste en aceptar la hipótesis nula siendo ésta falsa, es decir en afirmar que no hay diferencias cuando en realidad sí las hay. Nótese que hay dos posibilidades de cometer un error de tipo I, pero una sola de cometer un error de tipo II.

informado, haya dado su consentimiento para participar (véanse *Asignación oculta* y *Asignación no oculta* en el Glosario).

Es evidente que la mala suerte puede impedir que los grupos formados por asignación aleatoria simple sean comparables, sobre todo si el número de pacientes incluidos en el ensayo es reducido. Así por ejemplo, cuando el número total de pacientes es de 10, existe un 5% de probabilidades de que se produzca un desequilibrio extremo, del orden de 2 pacientes asignados a un grupo y 8 al otro; si el número de pacientes asignados es de 100, existe un 5% de probabilidades de desequilibrio del orden de 40:60, relación mucho menos desfavorable que la anterior.^{1,2}

En ocasiones conviene asegurar que las principales características pronósticas conocidas se distribuyen de manera equilibrada entre los grupos de tratamiento. Para ello se puede efectuar una asignación aleatoria por bloques, proceso que consiste en estratificar primero a los pacientes según su característica pronóstica. Así por ejemplo, en un ensayo en el que la edad constituya un

factor pronóstico relevante, los pacientes pueden ser inicialmente estratificados en tres grupos de edad; en un ensayo multicéntrico en el que se presuponga un importante efecto dependiente del centro, los pacientes pueden ser estratificados por centro (véase la figura 7). A continuación, dentro de cada estrato se procede a la asignación aleatoria, con lo que se asegura que la variable que ha sido motivo de estratificación (en los dos ejemplos anteriores, la edad y el centro) se distribuya de manera equilibrada entre ambos grupos de tratamiento.

En ocasiones puede ser conveniente asegurar que más de una variable se distribuya simultáneamente de manera equilibrada entre los grupos de tratamiento. En estos casos se puede aplicar el método de la minimización (véase *Asignación aleatoria por minimización* en el Glosario).

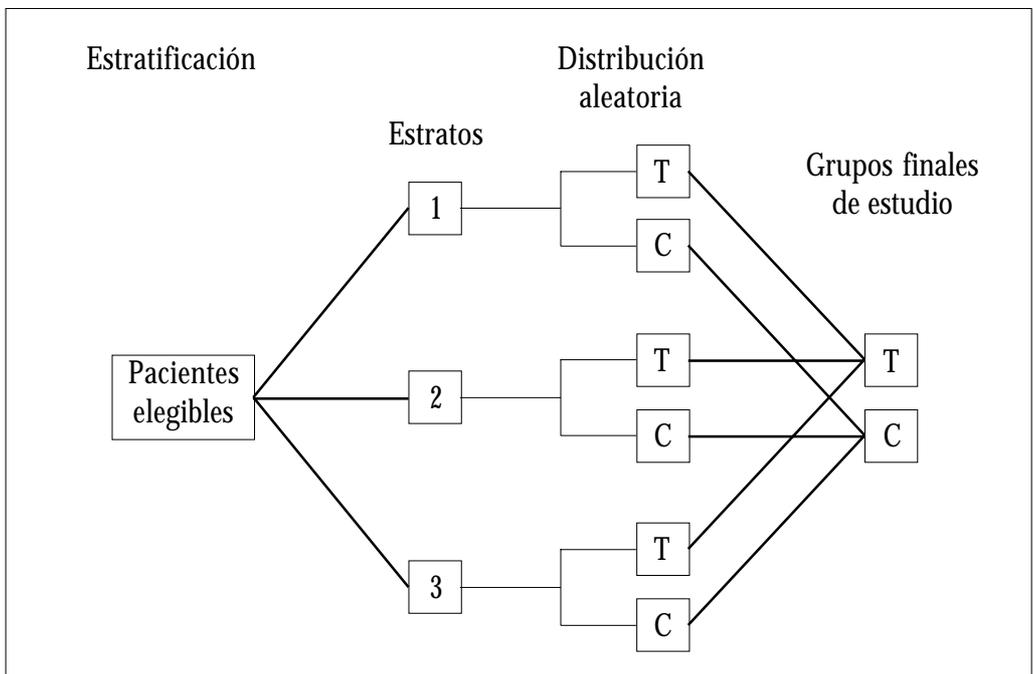


Figura 7. Asignación aleatoria estratificada: primero se forman estratos según una variable relevante. A continuación los pacientes de cada estrato son distribuidos al azar entre las diferentes posibles modalidades de tratamiento (T=tratamiento; C=control).

En cualquier caso, es fundamental que la asignación sea realizada una vez que se haya decidido la inclusión del sujeto en el ensayo clínico (después de que haya otorgado su consentimiento), pero sin conocer la modalidad de tratamiento que va a recibir. La importancia de que la asignación aleatoria se realice a ciegas se puso de manifiesto en un análisis de 145 artículos publicados que

describían ensayos clínicos.³ Los trabajos fueron divididos en tres grupos: 57 ensayos en los que la asignación aleatoria se había hecho a ciegas, 45 en los que podía no haberse realizado a ciegas y 43 en los que los controles habían sido seleccionados por un proceso no aleatorio. Por lo menos una variable pronóstica estaba distribuida de manera desequilibrada en un 14% de los del primer grupo, un 27% de los del segundo y un 58% de los del tercero. Se hallaron diferencias estadísticamente significativas ($p < 0,05$) entre las tasas de letalidad del grupo experimental y del grupo de referencia en un 9% de los estudios del primer grupo, un 24% de los del segundo y un 58% de los del tercero.³

La exposición a los diferentes tratamientos

En un ensayo clínico se produce un entrecruzado no intencionado de los tratamientos cuando un paciente que había sido asignado a un grupo recibe el tratamiento que corresponde a otro grupo. Así por ejemplo, en un ensayo sobre cirugía de cortocircuito coronario, un paciente asignado a recibir tratamiento médico podría decidir someterse al tratamiento quirúrgico durante el período de supervisión porque sus síntomas empeoran; un paciente asignado al tratamiento quirúrgico podría rechazar el procedimiento y optar por un tratamiento médico.

Los entrecruzados de este tipo son frecuentes en los ensayos sobre cortocircuito coronario. En el ensayo clínico de la *Veterans Administration* norteamericana sobre esta técnica,⁴ un 17% de los pacientes asignados al tratamiento médico se sometieron a tratamiento quirúrgico durante los primeros 21 meses; en el mismo período un 6% de los pacientes originalmente asignados al tratamiento quirúrgico lo rechazaron. La tasa acumulada de entrecruzado entre los asignados originalmente a tratamiento quirúrgico aumentó hasta un 30% a los 8 años y un 38% a los 11 años.⁵ En un estudio europeo similar (*European Coronary Surgery Study*), las tasas de entrecruzado a los 5 años fueron de un 24% para los pacientes asignados al tratamiento médico y 7% para los asignados al quirúrgico.⁶ En el *Coronary Artery Surgery Study* (CASS), las tasas de entrecruzado fueron de 24% para los asignados a tratamiento médico y 8% para los asignados al quirúrgico.⁷

Tasas de entrecruzado de estas magnitudes plantean dudas sobre posibles sesgos. En el estudio CASS, los pacientes que pasaron del tratamiento médico originalmente asignado al tratamiento quirúrgico presentaban más síntomas y mayor grado de afectación anatómica que los que se mantuvieron en el tratamiento originalmente asignado. Las tasas anuales de entrecruzado para los pacientes con afectación de un solo vaso, dos vasos o tres vasos fueron, respectivamente, de 2%, 4% y 8% ($p < 0,001$).⁸ Por el contrario, los pacientes que pasaron del tratamiento quirúrgico al médico presentaban menos síntomas y menor afectación anatómica: un 15% de los que tenían afectación de un solo vaso rechazaron el tratamiento quirúrgico y pasaron al médico, mientras que eso sólo lo hizo un 3% de los que presentaban afectación vascular triple

($p = 0,02$).⁷ Por lo tanto, en este ensayo la tasa de entrecruzado guardaba relación con la magnitud de la isquemia miocárdica y en consecuencia con el riesgo de ulteriores acontecimientos coronarios. Este sesgo tiende a favorecer el tratamiento médico, porque los pacientes más graves de entre los originalmente asignados al tratamiento médico tendían a adoptar el tratamiento quirúrgico, con lo que se determinaba una disminución de la proporción de acontecimientos entre los receptores de tratamiento médico.

Hay varias maneras de resolver este problema, todas ellas imperfectas. La primera consiste en analizar los datos según el tratamiento recibido. Para ello, se agrupa a los pacientes según el tratamiento que han recibido realmente, independientemente de la asignación original. La segunda es el análisis por intención de tratar; para realizarlo, los pacientes son agrupados según el tratamiento originalmente asignado. El análisis por intención de tratar se adecúa más al criterio de medición de la efectividad que al de la eficacia (véanse estos dos términos en el Glosario), y es más propio de los ensayos clínicos pragmáticos que de los explicativos (véanse *Análisis por intención de tratar*, *Ensayo clínico pragmático* y *Ensayo clínico explicativo* en el Glosario). Habitualmente el análisis principal de los resultados se hace por intención de tratar. [En cierto sentido, lo que importa desde un punto de vista práctico es conocer cómo evolucionará un paciente al que se le ofrece un determinado tratamiento, independientemente del modo cómo lo siga posteriormente, porque lo que importa es decidir qué se va a proponer al paciente.]

El enmascaramiento

Como se indica en la figura 4, uno de los principales pasos en la realización de un ensayo clínico es la evaluación de los resultados. Se ha comentado anteriormente que la evaluación del curso clínico de una enfermedad puede ser influida por diversos factores, como el efecto placebo, el efecto Hawthorne o la regresión a la media (véanse estos términos en el Glosario).

Quizá es menos conocida la razón por la que en muchas ocasiones conviene "cegar" no sólo al paciente, sino también al observador, a la persona que registra los resultados del ensayo. Para ello, nada más útil que reseñar la historia de Clever Hans.⁹

Clever Hans era un caballo. Parecía entender y responder preguntas sobre una increíble variedad de temas, con una precisión inaudita. Los directores de zoológicos y domadores de circo decían de él que era un caballo genial. Se hizo tan famoso, que su nombre se citaba en canciones populares y aparecía en etiquetas de licor. Tal popularidad condujo a la constitución de una comisión para investigarlo. Su dueño, el Sr von Osten, un viejo maestro de escuela, quería que el caballo fuera sometido a investigación científica y co-

operó con la comisión. La comisión comprobó que, sin ningún entrenamiento, la gente podía comunicar con Clever Hans; llegaron a hacerlo hasta 40 personas. Clever Hans demostró que, en ausencia de su dueño, era capaz de responder preguntas y que su inteligencia no era fraudulenta. Era capaz de resolver problemas aritméticos comunes, de reconocer el valor de diferentes monedas e incluso leía alemán bastante bien. Decía qué hora era y distinguía algunas canciones otrora populares, pero pasadas de moda. Seleccionaba a personas concretas entre una multitud e identificaba a personas por fotografías. Reconocía los colores. Cuando cometía un error, parecía reconocer inmediatamente su magnitud, ya que, cuando se le preguntaba en cuánto se había equivocado, solía dar la respuesta correcta. Pero, a diferencia del famoso Mr Ed, no podía hablar, por lo que se comunicaba golpeando su pata derecha en el suelo o moviendo la cabeza. Si sabía que el número de golpes que tenía que dar era elevado, comenzaba a golpear rápidamente; en caso contrario lo hacía más lentamente. Las preguntas de aritmética eran contestadas con un determinado número de golpes; respondía a las preguntas referentes a otros temas por medio de un código en el que los números significaban palabras o símbolos. Los resultados de la investigación se comenzaron a conocer al cabo de unos meses. Se descubrió que Hans no necesitaba que se le pronunciara la pregunta, bastaba pensar en ella. Cuando una operación matemática le era preguntada por dos personas que conocían todos sus términos (por ej., los sumandos de una suma), el caballo contestaba correctamente, pero si cada una de las dos personas aportaba un término de la operación y no conocía el planteado por la otra, Clever Hans se equivocaba irremisiblemente. Además, se pudo comprobar que cuando la persona que preguntaba se alejaba del caballo, éste comenzaba a cometer errores, tantos más cuanto más lejos estaba quien le preguntaba; que cuando se le tapaban los ojos se equivocaba más; que cuanto más oscuro era el ambiente más se equivocaba, y que en la oscuridad absoluta era un caballo estúpido. Estaba claro que el misterio era de tipo visible, y no una cuestión de poderes extrasensoriales. En 1904 Oskar Pfungst pudo comprobar que todos los comunicantes que habían obtenido respuestas del caballo, mientras le miraban como golpeaba la pata o movía la cabeza, le estaban dando pistas inconscientemente por medio de modificaciones mínimas de la posición de su cabeza debidas a cambios de tensión muscular. Estos movimientos sutiles eran extraordinariamente difíciles de evitar, incluso si la persona que preguntaba sabía que estaba dando pistas al caballo de este modo. Pfungst incluso registró en un quimógrafo los cambios de tensión muscular de estos observadores.

Esta historia muestra cuán sutil y en ocasiones irreconocible puede ser la comunicación inconsciente de las personas. Si un caballo podía detectar las esperanzas de quienes le hacían preguntas, ¿cuántas esperanzas y aspiraciones del investigador no detectará un paciente incluido en un ensayo clínico?⁹ Se dice que las personas tienden a hacer lo que creen que se espera de ellas, y no es menos cierto que los pacientes tienden a evolucionar como se espera de

ellos que evolucionen. Parece innecesario explicar por qué, en ciertas circunstancias, es conveniente que no sólo el paciente, sino también el investigador u observador, desconozcan cuál es el tratamiento administrado. Es cierto que en principio si se hace un ensayo clínico es porque no se conoce cuál de las opciones en él incluidas es la mejor. Sin embargo, de manera consciente o inconsciente, el investigador puede depositar más esperanzas en una de las intervenciones, en detrimento de la otra (u otras), y por este motivo es preferible que desconozca qué toma cada paciente.

Con el fin de asegurar que un ensayo se realiza realmente a doble ciego, es preciso que la forma farmacéutica, vía de administración y aspecto de los fármacos comparados no difieran (o bien, en ciertas circunstancias, que el observador que registra los datos clínicos ignore el tratamiento que recibe cada paciente). Los ensayos clínicos en los que se comparan fármacos con vías diferentes de administración o con pautas diferentes de administración plantean un problema especial para poder ser realizados a doble ciego. Con el fin de preservar la naturaleza doblemente ciega del ensayo, en estos casos se suele aplicar lo que se denomina un doble enmascaramiento (*double dummy* en inglés). Así por ejemplo, si el fármaco *A* debe darse cada 8 horas y el fármaco *B* cada 24 horas, los pacientes tratados con el fármaco *B* deberán recibir placebo de *A* a las mismas horas que los pacientes tratados con el fármaco *A* reciben su medicación. Análogamente, si la vía de administración de *A* es inyectable y la de *B* es oral, los pacientes tratados con *A* deben recibir también un placebo de *B* por vía oral y los tratados con *B* deben recibir un placebo de *A* por vía inyectable. No en todas las situaciones es necesario o preferible hacer el ensayo a doble ciego. El ensayo clínico a doble ciego no está indicado:

- si implica riesgos innecesarios para el paciente (por ej., en caso de administración de un placebo por vía parenteral de manera repetida y durante un período prolongado);
- si no es posible disponer de una formulación galénica adecuada;
- si los efectos farmacológicos permiten identificar fácilmente al menos uno de los fármacos estudiados (por ej., en un ensayo clínico en el que se comparen un bloqueador β -adrenérgico y un diurético, la bradicardia producida por el primero y la diuresis del segundo invalidarían el doble ciego), o bien
- si, por cualquier circunstancia, se considera que el diseño a doble ciego puede perjudicar la relación entre médico y paciente.

Por otra parte, algunas variables "duras" (por ej., mortalidad, paro cardíaco) no son sensibles al efecto placebo y no necesitan que el ensayo se realice a doble ciego.

TIPOS DE ENSAYOS CLÍNICOS

Se han propuesto numerosas clasificaciones de los ensayos clínicos. Según el grado de desarrollo clínico del fármaco, se clasifican en los de fase I, II, III o IV. Según el grado de enmascaramiento, pueden ser abiertos, a ciego simple, a doble ciego o a triple ciego. Según que cada participante reciba sólo una de las modalidades de tratamiento o todas ellas, pueden ser paralelos o cruzados. Según se predetermine un número de pacientes participantes o éstos se vayan comparando (generalmente por pares) a medida que avanza el ensayo, pueden ser no secuenciales o secuenciales (véanse un ejemplo en la figura 8 y los siguientes términos en el Glosario: *Ensayo clínico abierto*, *Ensayo clínico a ciego simple*, *Ensayo clínico a doble ciego*, *Ensayo clínico a triple ciego*, *Ensayo clínico cruzado*, *Ensayo clínico paralelo*, *Ensayo clínico secuencial*, *Fase I*, *Fase II*, *Fase III* y *Fase IV*).

ANÁLISIS DE DATOS

Este no es el lugar para hacer una descripción detallada de las técnicas utilizadas para el análisis de datos. Sin embargo, conviene comentar por lo menos cuatro aspectos: la comparabilidad de los grupos formados por la asignación aleatoria de los pacientes a los tratamientos, el análisis de subgrupos, el sentido de las diferencias estadísticamente significativas y la elección de la prueba estadística más adecuada.

Comparabilidad de los grupos

Una vez terminada la recogida de datos, el primer paso del análisis consiste en comprobar que los grupos que se han formado con la técnica de asignación aleatoria son realmente comparables en sus variables pronósticas conocidas (véase *Características basales* en el Glosario). Eso no siempre ocurre (debido a "mala suerte") y puede hacer fracasar un esfuerzo de muchas personas durante mucho tiempo. Lógicamente, cuanto mayor es el número de pacientes incluidos en un ensayo clínico, menor será la probabilidad de desequilibrios o diferencias en las características basales de los grupos comparados.

Para evitar este problema, se han propuesto métodos de asignación (por bloques, por minimización) con los que se pretende disminuir la probabilidad de desequilibrios entre los grupos (véanse *Asignación aleatoria por bloques* y *Asignación por minimización* en el Glosario). Así por ejemplo, en un estudio multicéntrico en el que se presupone que el factor centro puede tener influencia sobre el pronóstico, puede ser preferible una asignación por bloques, en la que el "bloque" es cada centro participante (véase la figura 7). (En estos casos es preferible que la asignación de pacientes en cada centro sea realizada central-

mente, con el fin de asegurar que se ha seguido el mismo método en todos los centros participantes).

Cuando los grupos difieren en alguna característica basal que pueda influir sobre el pronóstico (por ej., la edad), las diferencias pueden ser ajustadas *a posteriori* en el análisis de datos (esta no es una situación ideal, pero raramente ocurre algo ideal). Para ello se aplican básicamente dos tipos de técnicas: el análisis estratificado (o análisis de subgrupos) y la regresión múltiple.

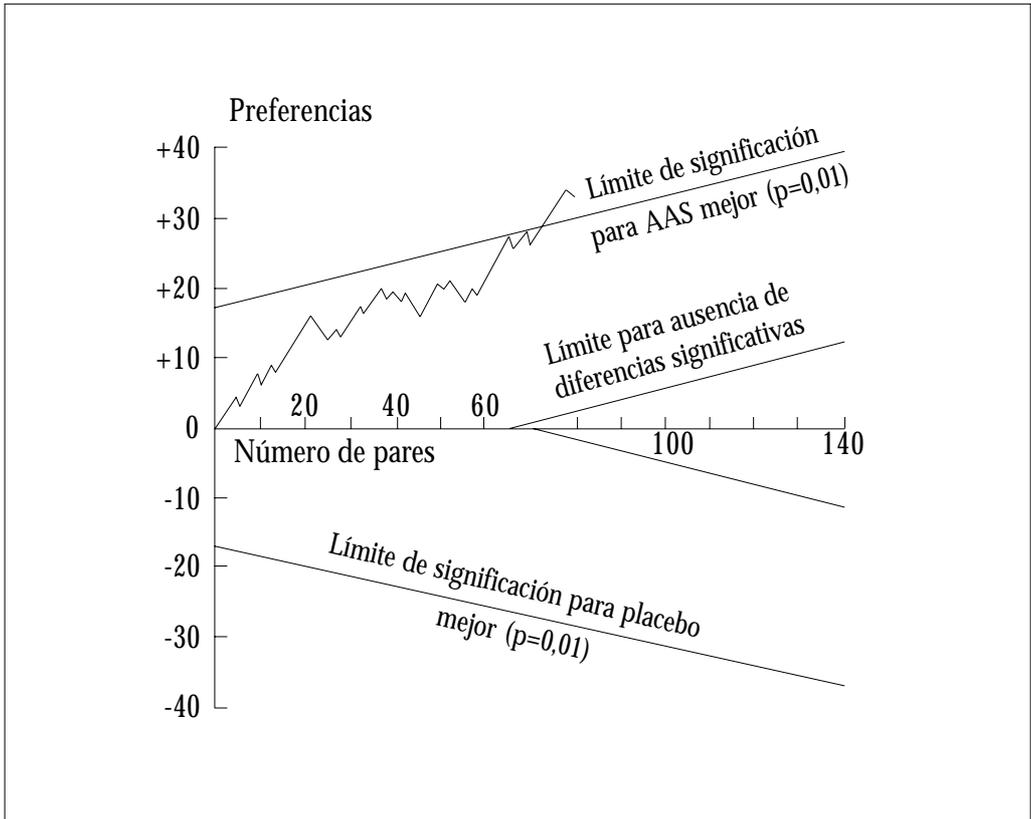


Figura 8. Valoración secuencial de los resultados de un ensayo clínico sobre la prevención de infarto agudo de miocardio y mortalidad con ácido acetilsalicílico (AAS) en pacientes con angina inestable. En un diseño secuencial de este tipo, se comparan pares de pacientes (en este caso particular, uno tratado con AAS y el otro con placebo, cuyos tratamientos han sido decididos por aleatorización). La línea quebrada representa los resultados que se fueron obteniendo con cada par de pacientes: la línea se traza hacia arriba si el acontecimiento estudiado (en este caso muerte o infarto agudo de miocardio no mortal) se produce en el paciente tratado con placebo ("AAS mejor") y hacia abajo si el acontecimiento se observa en el paciente tratado con AAS ("placebo mejor"). Los límites de significación estadística se determinan en función de los errores α y β esperados. (Tomada de Lewis y cols.¹⁰).

El **análisis de subgrupos** consiste en la repetición del análisis general en diferentes subgrupos o estratos de la población estudiada, seleccionados según una determinada característica (por ej., en hombres y mujeres, mayores y menores de 60 años, fumadores y no fumadores, etc.). Como se comenta en el apartado siguiente, el análisis estratificado o análisis de subgrupos es una técnica subóptima que puede inducir a error de tipo I. Véanse también *Análisis de subgrupos*, *Análisis estratificado*, *Dragado de datos*, *Error de tipo I*, *Estratificación* y *Paradoja de Simpson* en el Glosario.

Generalmente el **análisis de regresión múltiple** es más apropiado. En un análisis de regresión múltiple se considera simultáneamente un conjunto de variables seleccionadas previamente por el investigador, porque difieren de un grupo a otro y podrían influir sobre el pronóstico de la enfermedad estudiada. Véanse también *Análisis de regresión múltiple*, *Modelo lineal*, *Regresión múltiple*, *Variable dependiente* y *Variable independiente* en el Glosario.

Análisis de subgrupos y comparaciones múltiples

El análisis de subgrupos consiste en comparar sólo a los pacientes de cada grupo que tienen una característica determinada, en lugar de comparar a la totalidad de un grupo con el otro (por ej., en un ensayo clínico sobre insuficiencia cardíaca, comparar los resultados obtenidos sólo en los pacientes de cada grupo que presentaban la enfermedad como consecuencia de un infarto de miocardio, excluyendo a los demás).

El análisis de subgrupos debe ser considerado con mucha cautela. Puede ser útil para plantear nuevas hipótesis, que deberán ser comprobadas en nuevos ensayos diseñados a tal fin. Puede inducir a cometer un error intencionado de tipo I.¹¹ Véanse también *Análisis de subgrupos*, *Análisis estratificado*, *Dragado de datos*, *Error de tipo I*, *Estratificación* y *Paradoja de Simpson* en el Glosario.

En un reciente estudio se evaluaron diversos aspectos metodológicos en 50 publicaciones consecutivas de resultados de ensayos clínicos de cuatro prestigiosas revistas médicas (BMJ, JAMA, Lancet y NEJM) en 1997. Entre estos aspectos se incluían los métodos del análisis de datos. Dos terceras partes de los artículos presentaban resultados de subgrupos, pero en la mayoría de los casos sin que se hubieran aplicado pruebas estadísticas adecuadas para detectar que podían ser otras variables las causantes de las posibles diferencias. Los autores de este estudio concluyeron que una proporción elevada de los artículos daba una importancia excesiva al resultado de un análisis de subgrupos, que generalmente carecía de poder estadístico, y sugirieron que es necesario predefinir los análisis estadísticos de subgrupos que se van a realizar.¹²

Naturalmente, no es aceptable desde un punto de vista metodológico limitar la publicación de los resultados de un ensayo clínico sólo a los pacien-

tes incluidos que presentaban una característica determinada. Tampoco sería aceptable sacar conclusiones de superioridad o de equivalencia terapéutica de los tratamientos comparados a partir de un análisis de subgrupos que no se acompañe de los ajustes estadísticos necesarios y de una extremada prudencia en su interpretación (véanse también *Análisis de subgrupos*, *Corrección de Bonferroni*, *Dragado de datos* y *Paradoja de Simpson* en el Glosario).

Selección de la prueba estadística

Conviene elegir la prueba estadística adecuada según las variables incluidas en el ensayo. En la tabla 3 se da una orientación general sobre el tipo de prueba más conveniente según la naturaleza de las variables que se desee evaluar (véanse las definiciones de *Variables...* en el Glosario).

Para interpretar la validez estadística de una publicación es preciso conocer algunos principios básicos:¹³

- El **valor de p** indica la probabilidad de que un resultado determinado haya aparecido por casualidad. Cuando la probabilidad de una diferencia es 0,05, significa que si se repitiera 100 veces el mismo estudio en muestras diferentes de la misma población, en 5 ocasiones no se hallarían diferencias (y en 95 sí). Análogamente, si se repitiera 100 veces la medición de la misma variable, o bien si se midieran 100 combinaciones de variables y subgrupos, la probabilidad de cometer un error de tipo I aumentaría a medida que aumentara el número de comparaciones, de modo que con 20 comparaciones esta probabilidad alcanza un 25%; cuando se realizan comparaciones múltiples hay que aplicar una corrección estadística de los resultados, generalmente con el método de Bonferroni (véanse *Corrección de Bonferroni*, *P (probabilidad)* y las diferentes entradas *Prueba...* en el Glosario).
- En cualquier caso, hay que tener en cuenta que cuando se pregunta si el efecto terapéutico de una intervención es real o no, generalmente la respuesta no es dicotómica (sí o no), sino un continuo. Por este motivo, es preferible expresar los resultados con **intervalos de confianza**. Si se repitiera el mismo ensayo clínico centenares de veces, no siempre se obtendría el mismo resultado. Las muestras pequeñas y la presencia de números bajos en la tabla de 2×2 dan lugar a estimaciones poco precisas del intervalo de confianza (véase *Intervalo de confianza* en el Glosario).
- Las publicaciones que describen ensayos en los que se observa equivalencia terapéutica entre los grupos comparados (también llamados "negativos") son cada día más frecuentes. Para interpretar un ensayo "negativo" es importante preguntarse si un ensayo con una muestra de mayor tamaño hubiera demostrado una diferencia significativa. Para saberlo, es conveniente examinar el límite superior del IC95%. Sólo hay una probabilidad entre 40

(un 2,5%, porque el otro 2,5% de los resultados extremos está por debajo del límite inferior del IC95%) de que el resultado real sea éste o superior.

- A continuación, es conveniente preguntarse si este nivel de diferencia sería clínicamente importante. Si la respuesta es negativa, el ensayo puede clasificarse no sólo como negativo, sino también como definitivo. No obstante, si el límite superior del IC95% representara un nivel clínicamente importante de diferencia entre los dos grupos, el ensayo se podría considerar negativo pero no definitivo. Véanse también *Razón de ventajas*, *Reducción absoluta del riesgo*, *Reducción relativa del riesgo*, *Riesgo atribuible*, *Riesgo relativo* y *Número que es necesario tratar* en el Glosario).

Tabla 3. Técnicas estadísticas para el análisis de datos en ensayos clínicos en grupos paralelos.

Variable dependiente	Variable independiente			
	Binaria	Nominal	Ordinal	Intervalo
Binaria	χ^2	χ^2	Modelos log-lineales	Regresión logística
Nominal	χ^2	χ^2	Modelos log-lineales	-
Ordinal	U de Mann Whitney	Kruskall Wallis	Correlación de Kendall o de Spearman	Correlación de Kendall o de Spearman
Intervalo	Pruebas de t	Anova	Correlación de Kendall o de Spearman	Regresión

¿Estadísticamente significativo o clínicamente significativo?

Es preciso distinguir las diferencias clínicamente significativas de las que sólo son estadísticamente significativas.¹⁴ Una diferencia entre dos grupos puede ser estadísticamente significativa, pero tener escaso o ningún valor clínico. Véase la figura 5; cuando la muestra que participa en un ensayo clínico es homogénea en relación con la variable de interés o cuando el tamaño muestral es grande, es más probable encontrar diferencias que sean sólo significativas desde un punto de vista estadístico, pero con escasa repercusión clínica.

ASPECTOS ÉTICOS

La primera declaración sobre la ética de la investigación clínica fue consecuencia del conocimiento de los crímenes cometidos en los campos de concentración nazis durante la Segunda Guerra Mundial. Como dice Brennan,¹⁵ los principios heterogéneos de la bioética se basan en la repulsión, pero la repulsión no genera principios. La Declaración de la Asamblea Médica Mundial conocida como Declaración de Helsinki fue el primer documento que propuso criterios y medidas para proteger a los sujetos que participan en la investigación clínica. Véanse *Declaración de Edimburgo* y *Declaración de Helsinki* en el Glosario.

En la actualidad hay nuevos fenómenos que favorecen la aparición de posiciones favorables a revisar la Declaración de Helsinki.^{15,16} Por una parte, los elevados costes y la tremenda competencia industrial por desarrollar nuevos productos obligan a mejorar la eficiencia de todos los aspectos de la investigación; pero por otra parte el mercado no tolera bien el principio del compromiso moral que el investigador y a la vez médico debe tener con cada paciente o sujeto de investigación. En Estados Unidos es creciente la proporción de ensayos clínicos desarrollados en instituciones privadas por organizaciones de investigación clínica, en detrimento de la investigación realizada en centros universitarios (véanse *Organización de investigación clínica* y *Site management organization* en el Glosario). Por otra parte, se ha producido una rápida internacionalización de los ensayos clínicos de desarrollo de fármacos, sobre todo porque es menos caro realizarlos en países en desarrollo, en los que los investigadores potenciales suelen acoger más favorablemente las propuestas de colaboración. Al igual que en otros sectores industriales y de servicios, la globalización posibilita identificar el lugar en el que la producción de un determinado bien es más barata. La internacionalización acentúa las fisuras de los fundamentos de la bioética, porque no todos los países tienen las mismas tradiciones relativas al consentimiento informado y a la justicia distributiva.¹⁵ Además, la baja calidad de la atención médica en muchos lugares permite realizar ensayos más "eficientes" (es decir, menos costosos).

La primera condición que se debe cumplir para que un ensayo clínico se considere ético es que exista una duda razonable sobre la relación beneficio/riesgo que se puede obtener con cada una de las opciones comparadas; Bradford Hill llamó *equipoise* a este concepto (véase este término en el Glosario).¹⁷ Aunque este criterio pudiera parecer claro, conviene recordar que la interpretación de datos científicos está sujeta a subjetivismo y por lo tanto no todas las personas que examinen un protocolo de ensayo clínico deberían coincidir necesariamente sobre esta y otras cuestiones (véase *Beneficencia* en el Glosario).

La segunda condición debería ser que el ensayo esté bien diseñado. Si los resultados que va a producir la aplicación de un protocolo no van a poder ser interpretados de manera adecuada, por problemas de método que son ya reconocibles antes de comenzar el ensayo, no será ético exponer a los pacientes a los riesgos e incomodidades que supone su participación en él (véase *No maleficencia* en el Glosario).

El llamado Informe Belmont, elaborado por una Comisión del Congreso Norteamericano (véase este término en el Glosario) enunciaba los principios éticos básicos que deben animar la investigación clínica: el respeto a las personas, la beneficencia y la justicia (véanse estos términos en el Glosario).¹⁸ Gracia¹⁹ ha añadido otro principio a estos tres, el de no maleficencia, y clasifica los cuatro principios básicos en dos niveles: en el primero se situarían los principios de no maleficencia y justicia y en el segundo los de autonomía y beneficencia. En los casos en los que pudiera haber un conflicto entre principios situados en diferentes niveles, tendrían prioridad las consideraciones relacionadas con los principios del primer nivel, sobre los del segundo.^{18,19}

Consentimiento informado

En numerosas publicaciones se considera el consentimiento informado por escrito como el procedimiento óptimo o "de referencia" que garantizaría los estándares éticos de un ensayo clínico, y se califica el consentimiento verbal como una forma subóptima. Un documento escrito ofrece sin duda mejor garantía de que se han seguido procedimientos estandarizados. Sin embargo, es preciso reconocer que la insistencia en que el consentimiento sea escrito se debe sobre todo a sus implicaciones legales: el paciente firma un contrato con posibles consecuencias económicas, pues existe un seguro y se prevén unas indemnizaciones. En diversos trabajos se ha comprobado que en la práctica el grado de comprensión y la libertad de los pacientes para consentir estaban lejos de la perfección. Las barreras culturales y emocionales entre una persona enferma y un médico/institución son inevitables, e independientemente de lo ética o compasiva que sea la actitud que se adopte, las condiciones óptimas son más la excepción que la regla, porque la asimetría de conocimientos y de poder que caracteriza la relación habitual entre cuidadores y pacientes no se resuelve con la firma de un documento. En este sentido, se debería considerar que la experimentación es una prolongación sin

solución de continuidad de la práctica "apropiada" o "recomendable" (lo que invita a pensar, por ej., en las circunstancias y ocasiones en las que, en la práctica habitual, se somete a los pacientes a pruebas diagnósticas o a tratamientos agresivos sin que medie ya no consentimiento, sino ni tan sólo una mera explicación de las razones de tales molestias). En la actualidad, en muchos ensayos de escaso interés clínico, los pacientes son de hecho invitados a participar en una investigación cuyos objetivos son obtener un nuevo conocimiento trivial (o a veces ni eso), o satisfacer demandas reguladoras, pero no mejorar directamente el cuidado de los enfermos. El investigador debe estar seguro de que podrá garantizar al paciente que la investigación es bienintencionada (principio de beneficencia) y que no supondrá riesgos excesivos (principio de no maleficencia). Es difícil imaginar una hoja informativa que afirme de manera no ambigua que el protocolo del estudio es malo y que no es de esperar que produzca información útil.²⁰ Tanta ha sido en ocasiones la insistencia en el consentimiento informado, que se citan ensayos sobre el tratamiento trombolítico del infarto agudo de miocardio en los que leer la hoja de información al paciente anexa al documento de consentimiento necesitaba entre 30 y 60 minutos.²⁰ Una excesiva atención a los aspectos burocráticos de los reglamentos puede facilitar que se olvide que los derechos fundamentales de los pacientes deben estar por encima del cumplimiento de procedimientos y regulaciones.

Situar el consentimiento informado como el componente más importante de los aspectos "éticos" de la experimentación es una actitud éticamente equívoca. Las palabras y argumentos de apariencia ética pueden dar un falso aire de responsabilidad, atención y respeto. La legislación y las regulaciones locales deberían concentrarse en calidades substanciales, como la autonomía, el grado de incertidumbre y la igualdad de probabilidades (la *equipoise* de la que hablaba Bradford Hill¹⁷), así como el interés clínico y la adecuación metodológica del diseño del estudio.²⁰ Los ensayos clínicos deberían respetar y reforzar las regulaciones actuales sobre el consentimiento en la medida en que favorezcan la autonomía de pacientes y cuidadores. Como afirman Tognoni y Geraci,²⁰ al igual que con cualquier otro aspecto de la democracia, la investigación crítica es más importante que la obediencia repetitiva de las regulaciones establecidas.

La ética del ensayo controlado con placebo

La Declaración de Helsinki afirma que "en todo estudio médico se debe asegurar que cada paciente (incluidos los del grupo control si lo hubiere) recibe el mejor método diagnóstico y terapéutico". Más adelante, la Declaración afirma que un estudio que viole sus preceptos no debe ser aceptado para publicación. A pesar de ello, se aprueban, realizan y publican numerosos ensayos controlados con placebo. Muchos ensayos son realizados con la única finalidad de obtener la aprobación del registro de un medicamento, y no llegan

a ser publicados, por lo que es imposible conocer la proporción de ensayos controlados con placebo.²¹ La ética de incluir un grupo de control tratado con placebo en los ensayos clínicos ha sido objeto de amplio debate.²¹⁻³¹

En todas las áreas terapéuticas se han descrito ensayos controlados con placebo que plantean dudas sobre su ética. Ejemplos que se citan a menudo son un ensayo sobre el tratamiento de la oncocercosis con ivermectina en África, cuando la dietilcarbamacina había sido tratamiento aceptado de la enfermedad durante tres décadas (en este ensayo algunos participantes, pescadores analfabetos, firmaron su consentimiento informado con sus huellas dactilares),³² numerosos ensayos sobre artritis reumatoide,³³ ensayos sobre el tratamiento de la depresión,³⁴ el tratamiento de las náuseas y vómitos por quimioterápicos antineoplásicos,^{35,36} el tratamiento de la insuficiencia cardíaca congestiva y la disfunción ventricular,³⁷ y sobre el tratamiento de la hipertensión sistólica en geriatría.³⁸ En todos estos casos se conocían tratamientos de eficacia bien establecida en el momento de iniciar los ensayos.

El abuso del control con placebo no es casual. De hecho, los principales organismos reguladores pueden no aprobar ciertos fármacos porque sólo habían sido comparados con otros tratamientos establecidos, pero no con placebos.^{21,39} Es evidente que no todas las enfermedades objeto de estudio merecen la misma consideración en este sentido (no es lo mismo hacer un ensayo clínico controlado con placebo sobre el tratamiento de la endocarditis bacteriana, que sobre el del dolor postoperatorio), pero es igualmente evidente que los ensayos clínicos controlados con placebo vulneran el principio de que todos los pacientes, incluidos los del grupo control, deben recibir el mejor tratamiento disponible, o uno nuevo que se crea que es igual o mejor; por otra parte, aceptar los ensayos controlados con placebo atribuye a los investigadores y comités éticos la facultad de juzgar la cantidad de molestias y de incapacidad temporal que los pacientes pueden llegar a sufrir para asegurar el desarrollo de la investigación,²¹ no constituye un estímulo para las comparaciones directas entre fármacos, y dificulta la aplicación de criterios de selección de medicamentos para la prescripción.^{30,39}

Dadas estas premisas, se han propuesto algunas condiciones para el uso de controles tratados con placebos en ensayos clínicos:³¹

- que el ensayo no dure el tiempo suficiente para que la enfermedad tratada dé lugar a daño o pérdida irreversible (por ej., tres meses en los ensayos sobre artritis reumatoide o sobre hipertensión ligera);
- que los pacientes que no responden sean identificados y retirados, con el fin de protegerlos de las consecuencias de un tratamiento inadecuado (al igual que se aplican criterios estrictos sobre toxicidad para decidir la eventual retirada de pacientes, se deberían aplicar criterios estrictos de ineficacia: los participantes en un ensayo clínico que no cumplan criterios predefinidos al cabo del mismo período que en la condición anterior, deberían ser retirados);

- las hojas de información para el consentimiento del paciente deben incluir información sobre los posibles riesgos de lesión o efecto adverso irreversible resultante del uso del placebo o de un fármaco en investigación que resulte ser ineficaz, y
- redundaría en beneficio de todos (pacientes, industria, investigadores y reguladores) que la industria y las autoridades reguladoras (que de hecho son los principales decisores del diseño de los ensayos clínicos) definieran y justificaran conjuntamente una "posición ética" específica, con unas directrices claras, sobre esta cuestión.

PUBLICACIONES SOBRE ENSAYOS CLÍNICOS

Un informe que describa los resultados de un ensayo clínico debe contener la información suficiente para que el lector pueda hacerse un juicio informado sobre su validez interna y externa (véanse *Validez externa* y *Validez interna* en el Glosario). Diversos estudios han indicado que muchos artículos publicados que describen ensayos clínicos no dan los detalles suficientes para ello.

En 1996 un grupo multidisciplinario propuso la adopción de una serie de recomendaciones para la preparación y publicación de artículos sobre ensayos clínicos.⁴⁰ Se trata básicamente de una lista de comprobación acompañada de un diagrama de flujos. La lista de comprobación, resumida en la tabla 4, consta de 21 elementos básicos, referentes principalmente a los métodos, resultados y discusión. El diagrama de flujos da información sobre la selección, reclutamiento, seguimiento y evaluación final de los pacientes en ensayos clínicos paralelos.

En la figura 9 se da un ejemplo de tal tipo de diagrama. Está construida con los datos presentados en el informe publicado sobre un ensayo clínico cuyo objetivo era evaluar la eficacia del ácido acetilsalicílico (AAS) para prevenir la muerte y el infarto agudo de miocardio en pacientes con angina inestable.¹⁰ Como puede apreciarse, sólo una proporción reducida de los pacientes inicialmente identificados como posibles casos de angina inestable fueron incluidos en el ensayo. En el momento en el que éste fue realizado, todavía no se había confirmado que el efecto antiagregante plaquetario del AAS se traducía en un efecto terapéutico muy notable en pacientes con cardiopatía isquémica, entre ellos los que padecen angina inestable. Por este motivo la orientación de este ensayo fue más bien explicativa (véase *Ensayo clínico explicativo* en el Glosario), y con el fin de evitar la introducción de sesgos y factores de confusión (véanse estos términos en el Glosario), se aplicaron criterios muy estrictos de inclusión y exclusión. En cualquier caso, a la vista de la figura 9 cabe preguntarse si la naturaleza y la cuantía del efecto beneficioso del AAS en la angina inestable pueden ser extrapoladas a los pacientes que presentan características que ingenuamente podrían ser calificadas de "atípicas", pero que, como los datos

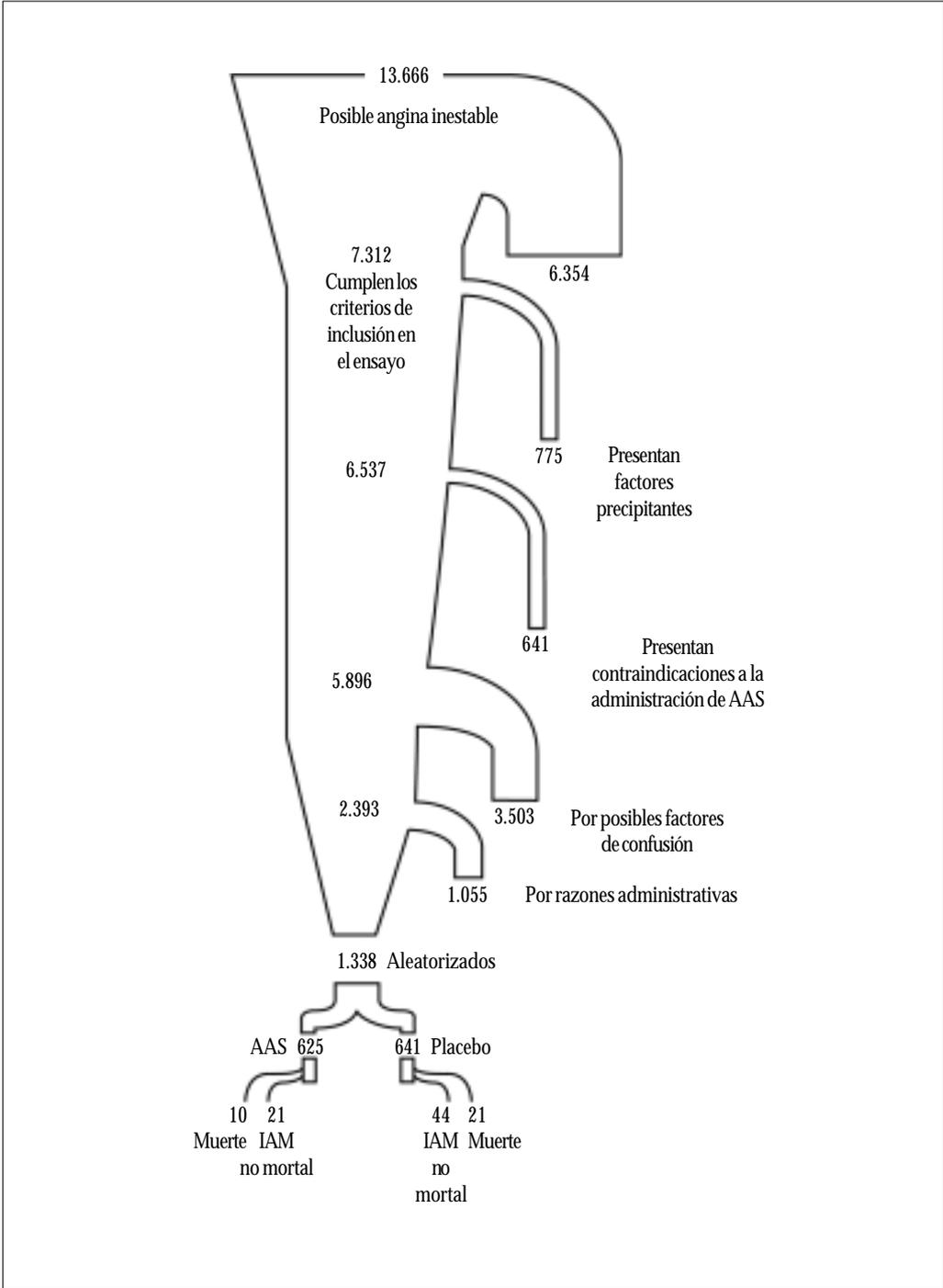


Figura 9. Diagrama del reclutamiento de pacientes en el ensayo clínico de la figura 8. A partir de 13.666 candidatos a participar, el proceso de selección, basado en la aplicación de criterios estrictos predeterminados de inclusión y exclusión, dejó sólo como participantes a menos de un 10%

de los candidatos originales. Se identificaron 13.666 pacientes con un cuadro clínico compatible con angina inestable, de los que 6.354 no cumplían los criterios de inclusión, con lo que quedaban 7.312 pacientes. Pero de éstos, 775 (10,6%) presentaban algún factor precipitante considerado criterio de exclusión: 377 tenían insuficiencia cardíaca congestiva, 138 taquiarritmias con frecuencia ventricular >150/minuto, 108 anemia con hemoglobina <10 g/100 ml y 152 hipoxia con una pO_2 <50 mm Hg o fiebre, hipotensión, hipertensión grave o tirotoxicosis. Quedaban aún 6.537 pacientes, de los que 641 no pudieron ser incluidos porque presentaban contraindicaciones a la toma de AAS: 302 ya estaban tomando anticoagulantes, 44 padecían alguna hemorragia, 140 habían presentado síntomas de úlcera péptica en los 3 meses anteriores y 155 tenían una historia de alergia o intolerancia al AAS. De los 5.896 pacientes restantes, 970 habían tomado AAS en 3 o más ocasiones en la semana precedente, 430 habían tomado algún otro antiagregante plaquetario en las 24 h anteriores, 380 habían padecido un infarto en las 6 semanas precedentes, 307 habían sido sometidos a una intervención quirúrgica o a cateterización cardíaca en las semanas anteriores, 518 presentaban otra enfermedad que hubiera podido afectar su supervivencia al año, 380 estaban participando en otro ensayo clínico, 110 eran en realidad de sexo femenino (y, por una u otra razón, eso había pasado inadvertido en el primer cribado) y no pudieron ser incluidos; finalmente, 408 pacientes presentaban alguna otra razón (por ej., diagnóstico más de 48 h después del ingreso) para no ser incluidos. Tras estas exclusiones quedaban 2.393 candidatos a participar en el ensayo, de los que 1.055 tuvieron que ser excluidos por razones administrativas, con lo que finalmente fueron incluidos en el ensayo 1.338 pacientes, es decir, aproximadamente un 10% de los inicialmente considerados. (Elaborada con datos de Lewis y cols.¹⁰)

muestran, son muy comunes. Los pacientes incluidos en ensayos clínicos son pacientes "limpios", con una sola enfermedad, cuyas manifestaciones clínicas son típicas; sin embargo, los pacientes reales presentan varias enfermedades a la vez, contraindicaciones al uso de ciertos fármacos, vienen siendo tratados con otros medicamentos, pueden tener una edad extrema, etc. Estas diferencias entre las poblaciones que participan en ensayos clínicos y las de la práctica real se comentan con más detalle en el Capítulo V.

ASPECTOS NORMATIVOS Y ECONÓMICOS

Se adjunta a esta guía el texto del Real Decreto 561/1993,³¹ por el que se establecen los requisitos para la realización de ensayos clínicos con medicamentos (véase también *Regulación legal* en el Glosario). En otras definiciones incluidas en el Glosario (*Ley del Medicamento, Investigador, Monitor, Promotor, Producto en fase de investigación clínica, Protocolo, Seguro*) se explica el sentido de las principales disposiciones contenidas en este Real Decreto, así como las responsabilidades de promotor, monitor e investigador y las atribuciones de los Comités Éticos de Investigación Clínica, las Comunidades Autónomas y la Dirección General de Farmacia y Productos Sanitarios.

Tabla 4. Estructura de los informes y artículos sobre ensayos clínicos controlados. Recomendaciones del grupo CONSORT (*Consolidation of Standards for Reporting Trials*).⁴⁰

Título	Identificar el estudio como un ensayo clínico con distribución aleatoria
Resumen	Presentar un resumen estructurado
Introducción	Indicar las hipótesis iniciales, los objetivos clínicos y los análisis de subgrupos y de covariables planificados
Métodos	
Protocolo	Describir: La población planificada y sus criterios de inclusión y exclusión Las intervenciones planeadas y el momento de su aplicación Las variables primaria y secundaria(s) y la(s) diferencia(s) mínima(s) relevante(s), e indicar cómo se proyectó el tamaño de muestra deseado Las razones y los métodos del análisis estadístico, con detalle de los principales análisis comparativos, y si se realizaron por intención de tratar Normas predefinidas de detención del tratamiento (si proceden)
Asignación	Describir: Unidad de aleatorización (por ej., individuo, grupo, geográfica) Método utilizado para generar el plan de aleatorización Método de ocultación de la asignación y momento en el que fue realizada Método aplicado para separar al generador del ejecutor de la asignación
Enmascaramiento (obligatorio)	Describir el modo (por ej., cápsulas, comprimidos); similitud de los tratamientos (por ej., aspecto, sabor); control de la pauta de aleatorización (localización del código durante el ensayo y cuándo fue abierto); y

pruebas de enmascaramiento efectivo entre participantes, personas que aplicaron la intervención, evaluadores del resultado y analistas de datos

Resultados

Flujo y seguimiento de los participantes

Presentar un perfil del ensayo (véase la figura 9) en el que se resuman el flujo de participantes, los números y el momento de la asignación aleatoria, las intervenciones y las medidas en cada grupo aleatorizado

Análisis

Indicar el efecto calculado de la intervención sobre las variables primaria y secundaria, con un estimador y medida de la precisión (intervalos de confianza)

Presentar los resultados en números absolutos cuando sea posible (por ej., 10/20, en lugar de 50%)

Presentar datos resumidos y resultados de pruebas estadísticas adecuadas descriptivas e inferenciales detallados, de manera que permitan la realización de análisis alternativos y la repetición de los análisis efectuados

Describir las variables pronósticas en cada grupo de tratamiento y las medidas tomadas para ajustarlas

Describir las desviaciones del protocolo y sus razones

Comentario

Indicar la interpretación específica de los resultados, con fuentes de sesgo e imprecisión (validez interna) y discutir la validez externa, con especificación de medidas cuantitativas cuando sea posible

Describir la interpretación general de los resultados a la luz de la totalidad de los datos disponibles

Bibliografía

1. De Abajo FJ. Fundamentos de los ensayos clínicos. En: Carvajal A (ed) Farmacoepidemiología. Valladolid: Universidad de Valladolid, 1993:83-105.
2. Pocock SJ. Clinical trials. A practical approach. Chichester: Wiley, 1991:84-5.
3. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358-61.

4. Murphy ML, Hultgren HN, Detre K, Thomsen J, Takaro J, Participants of the Veterans Administration Cooperative Study. Treatment of chronic stable angina. A preliminary report of survival data of the randomized Veterans Administration cooperative study. *N Engl J Med* 1977;297:621-7.
5. The Veterans Administration Coronary Artery Bypass Surgery Cooperative Study Group. Eleven-year survival in the Veterans Administration randomized trial of coronary bypass surgery for stable angina. *N Engl J Med* 1984;311:1333-9.
6. European Coronary Surgery Study Group. Long-term results of prospective randomized study of coronary artery bypass surgery in stable angina pectoris. *Lancet* 1982;2:1173-80.
7. Coronary Artery Surgery Study (CASS). A randomized trial of coronary artery bypass surgery. Survival data. *Circulation* 1983;68:939-50.
8. Lamas GA, Pfeffer MA, Hamm P, Wertheimer J, Rouleau JL, Braunwald E. Do the results of randomized clinical trials of cardiovascular drugs influence medical practice? *N Engl J Med* 1992;327:241-7.
9. Modell W. Clever Hans: the horse of Mr. von Osten. *Clin Pharmacol Ther* 1969;10:125-6.
10. Lewis HD, Davis JW, Archibald DG, *et al.* Protective effects of aspirin against acute myocardial infarction and death in men with unstable angina. Results of a Veterans Administration Cooperative Study. *N Engl J Med* 1983;309:396-403.
11. Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991;266:93-8.
12. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355:1064-9.
13. Wade OL, Waterhouse JAH. Significant or important? *Br J Clin Pharmacol* 1977;4:411-2.
14. Greenhalgh T. Statistics for the non-statistician. II: "Significant" relations and their pitfalls. *BMJ* 1997;315:422-5.
15. Brennan TA. Proposed revisions to the Declaration of Helsinki - Will they weaken the ethical principles underlying human research? *N Engl J Med* 1999; 341:527-30.
16. Levine RJ. The need to revise the Declaration of Helsinki. *N Engl J Med* 1999;341: 531-4.
17. Bradford Hill A. Suspended judgement. Memories of the British Streptomycin Trial in Tuberculosis. The first randomized clinical trial. *Contr Clin Trial* 1990;11:77-9.
18. Galende Domínguez I. Problemas éticos de la utilización de seres humanos en investigación biomédica. Consecuencias sobre la prescripción de medicamentos. *Med Clín (Barc)* 1993;101:20-3.

19. Gracia D. Fundamentos de bioética. Madrid: Eudema, 1989.
20. Tognoni G, Geraci E. Approaches to informed consent. *Contr Clin Trial* 1997;18:621-7.
21. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994;331:394-8.
22. Clark PI, Leaverton PE. Scientific and ethical issues in the use of placebo controls in clinical trials. *Annu Rev Public Health* 1994;15:19-38.
23. Okada F. New versus old placebos. *Ann Pharmacother* 1994;28:1201.
24. Henry D, Hill S. Comparing treatments. *BMJ* 1995;310:1279.
25. Rothman, KJ. Placebo mania. *BMJ* 1996;313:3-4.
26. Lurie P, Wolfe SM. Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. *N Engl J Med* 1997;337:853-6.
27. Angell M. The ethics of clinical research in the Third World. *N Engl J Med* 1997;337:847-9.
28. de Abajo FJ, Gracia DM. Etica del uso de placebo en investigación clínica. *Investigación y Ciencia* 1997;254:90-9.
29. García-Alonso F, Guallar E, Bakke OM, Carné X. El placebo en ensayos clínicos con medicamentos. *Med Clin (Barc)* 1997;109:797-801.
30. Laporte JR. El placebo en ensayos clínicos con medicamentos. *Med Clin (Barc)* 1998;111:558.
31. Stein CM, Pincus T. Placebo-controlled studies in rheumatoid arthritis: ethical issues. *Lancet* 1999;353:400-3.
32. Greene BM, Taylor HR, Cupp EW, *et al.* Comparison of ivermectin and diethylcarbamazine in the treatment of onchocerciasis. *N Engl J Med* 1985;313:133-8.
33. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis: results of two metaanalyses. *Arthritis Rheum* 1990;33:1449-61.
34. Williams Jr JW, Mulrow CD, Chiquette E, Noël PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Ann Intern Med* 2000;132:743-56.
35. Citron ML. Placebos and principles: a trial of ondansetron. *Ann Intern Med* 1993;118:470-1.
36. Aspinall RL, Goodman NW. Denial of effective treatment and poor quality of clinical information in placebo controlled trials of ondansetron for postoperative nausea and vomiting: a review of published trials. *BMJ* 1995;305:844-6.
37. Flather MD, Yusuf S, Kober L, Pfeffer M, Hall A, Murray G, Torp-Pedersen C, Ball S, Pogue J, Moyé L, Braunwald E, for the ACE-Inhibitor Myocardial Infarction

- Collaborative Group. Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: a systematic overview of data from individual patients. *Lancet* 2000;355:1575-81.
38. Staessen JA, Fagard R, Thijs L, Celis H, Arabidze GG, Birkenhäger WH, Bulpitt CJ, de Leeuw PW, Dollery CT, Fletcher AE, Forette F, Leonetti G, Nachev C, O'Brien ET, Rosenfeld J, Rodicio JL, Tuomilehto J, Zanchetti A, for the Systolic Hypertension in Europe (Syst-Eur) Trial Investigators. Randomised double-blind comparison of placebo and active treatment for older patients with isolated systolic hypertension. *Lancet* 1997;350:757-64.
39. Furberg CD, Herrington DM, Psaty BM. Are drugs within a class interchangeable? *Lancet* 1999;354:1202-4.
40. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Smel D, Stroup DF. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;637-9.