

Estudios clínicos experimentales

Juan José Calva-Mercado, M.C., M.Sc.⁽¹⁾

Los estudios clínicos experimentales, también conocidos como ensayos clínicos controlados, son estrategias diseñadas para evaluar la eficacia de un tratamiento en el ser humano mediante la comparación de la frecuencia de un determinado evento de interés clínico (o desenlace) en un grupo de enfermos tratados con la terapia en prueba con la de otro grupo de enfermos que reciben un tratamiento control. Ambos grupos de enfermos son reclutados y seguidos de la misma manera y observados durante un mismo periodo de tiempo. La esencia de los ensayos clínicos es que el propio investigador decide qué individuos serán sometidos al tratamiento en prueba (grupo experimental) y quiénes estarán en el grupo comparativo (grupo control o de contraste). Es justamente esta propiedad la que distingue a los ensayos clínicos experimentales de los estudios observacionales (de cohortes) ya que en éstos el propio paciente o su médico tratante decide quién se somete (y quién no) a la maniobra en evaluación; esta decisión obedece a múltiples razones, algunas estrechamente ligadas al pronóstico de la enfermedad. Debido a que los ensayos clínicos controlados son estudios diseñados con antelación, la asignación de la maniobra experimental por el investigador puede seguir diversos procedimientos; cuando es mediante un sorteo el estudio se conoce como un ensayo clínico aleatorizado, o controlado por sorteo.

Cabe aclarar la connotación de diversos términos empleados en la anterior definición y a lo largo del presente capítulo. El uso del término tratamiento (o tera-

pia) es con un sentido amplio; se refiere no sólo a un medicamento sino que también incluye otro tipo de intervenciones (o maniobras) tales como un procedimiento quirúrgico, una medida preventiva (o profiláctica), un programa educativo, un régimen dietético, etcétera. De igual manera, el término evento (o desenlace) se puede referir a toda una diversidad de resultados, tales como: mediciones bioquímicas, fisiológicas o microbiológicas, eventos clínicos (intensidad del dolor, aparición de infecciones oportunistas, desarrollo de un infarto al miocardio, recaída de una leucemia aguda, etc.), escalas de actividad de una enfermedad (como la del lupus eritematoso generalizado), mediciones de bienestar o funcionalidad (calificación de Karnofsky, escala de calidad de vida) o el tiempo de supervivencia. Por último, el grupo control se refiere al grupo de individuos que reciben una intervención que sirve de contraste para evaluar la utilidad relativa de la terapia experimental y que no necesariamente tiene que ser un placebo pues en ocasiones lo más adecuado (y ético) es que sea el tratamiento estándar, es decir, la mejor alternativa terapéutica vigente en el momento del diseño del experimento clínico.

¿Por qué y cuándo son necesarios?

La pregunta nos remite a reflexionar sobre dónde surge la idea de que un cierto tratamiento pueda modificar la historia natural de una enfermedad y acerca de la necesidad de contar con suficientes observaciones

(1) Subdirección de Investigación Clínica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán. México, D.F., México.

Solicitud de sobretiros: Juan José Calva Mercado. Subdirección de Investigación Clínica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán. Vasco de Quiroga 15, Delegación Tlalpan, 14 000 México, D.F., México.
Correo electrónico: jcalva@quetzal.innsz.mx

sistematizadas para conocer el verdadero efecto de la terapia en cuestión en los seres humanos, antes de prescribirlos de manera rutinaria a los enfermos.

En ocasiones, surge la idea de que un fármaco pudiera ser clínicamente útil al comprender su íntimo mecanismo de acción, así como la patogenia a nivel celular y molecular de una determinada enfermedad. De hecho, en ocasiones resulta muy tentador intentar predecir el efecto de un cierto medicamento con base únicamente en la información generada en el laboratorio, tanto *in vitro* como en modelos animales de experimentación. Sin embargo, por muy profundo que sea este conocimiento siempre será incompleto si no se tiene la experiencia en seres humanos intactos; de no ser así, se corre el riesgo de sorpresas desagradables. Por ejemplo, se sabe que el antimetabolito citarabina interfiere con la síntesis de la pirimidina y que es capaz de inhibir *in vitro* a varios virus con ADN, incluyendo al virus herpes varicela-zóster. A algún ingenuo clínico se le ocurriría que pudiera beneficiar con este medicamento a sus enfermos con herpes zóster generalizado; afortunadamente a un grupo de médicos escépticos se les ocurrió comparar la evolución de un grupo de estos enfermos, a quienes se les administró la citarabina, con la de otro grupo de enfermos semejantes, quienes sólo recibieron placebo, y demostraron no sólo lo ineficaz del antiviral sino una peor evolución en los enfermos que la habían recibido, explicable por sus efectos inmunosupresores.

Otras dos fuentes de ideas sobre el posible valor de una cierta terapia suelen ser: a) las observaciones empíricas de clínicos perspicaces; así, por serendipia surgió la idea de que la amantadina pudiera ser de utilidad en los enfermos de Parkinson puesto que a quienes se les prescribía para prevenir la influenza mostraban mejoría de sus manifestaciones neurológicas o como el caso de la reducción de crisis de fiebre familiar del Mediterráneo con el uso de colchicina administrada con la intención de prevenir ataques de gota. En estos ejemplos, el valor de estos tratamientos no provino de un entendimiento de la patogenia de estas enfermedades, la que de hecho, aún no se conoce bien, y b) las observaciones de estudios poblacionales donde se llega a establecer una asociación directa (o inversamente) proporcional entre la frecuencia de una enfermedad y alguna condición ambiental; tal como ha sido el caso de una relativa menor frecuencia de enfermedad coronaria en poblaciones con mayor ingesta de alimentos ricos en antioxidantes, o de cáncer de colon y una dieta con alto contenido en fibra.

El punto esencial es que cualquiera que sea el origen de las hipótesis del posible beneficio de una terapia, éstas deben probarse, y demostrarse como ciertas,

mediante estudios clínicos; es decir, mediante la observación sistematizada y objetiva de su efecto en seres humanos que la reciben, y su comparación con lo que habitualmente sucede en un grupo de enfermos sin la terapia.

Hay circunstancias en las que, por una parte, se conoce muy bien la historia natural de una enfermedad y que es tan consistente que es posible, con razonable certeza, predecir su curso clínico (generalmente muy desfavorable) y, por la otra, que el beneficio de una cierta terapia en este tipo de enfermedades es tan dramático e incuestionable que resulta innecesario realizar un ensayo clínico para aceptar su uso generalizado. Es decir, es suficiente con comparar la evolución clínica de un grupo de casos tratados con la nueva terapia con lo que habitualmente pasa en los enfermos antes del acceso a ésta (controles históricos), tal y como ha sucedido con el beneficio de los antibióticos en el tratamiento de la neumonía bacteriana, de los antifímicos en la meningitis tuberculosa o de la cirugía abdominal en la apendicitis: no hay alguien en el mundo que haya exigido el haber realizado un ensayo clínico en el que un grupo de estos enfermos recibieran un placebo. Sin embargo, esta es una circunstancia excepcional; lo común es que diferentes individuos con la "misma" enfermedad tengan cursos disímiles (en ocasiones, en sentidos totalmente opuestos) e impredecibles, de tal suerte que esta incertidumbre pronóstica impide deslindar qué tanto la mejoría o las complicaciones observadas en un grupo de enfermos que reciben una nueva terapia es atribuible a ésta, o bien, es parte de su historia natural o consecuencia de otros determinantes ajenos a la intervención en evaluación. Por ejemplo, un grupo de enfermos puede tener una mejor evolución clínica a pesar de recibir un tratamiento totalmente inútil, porque son pacientes en etapas tempranas de su enfermedad, reciben otros medicamentos concomitantes o mejor atención médica, están bajo el efecto placebo de la terapia novedosa o porque comparten el entusiasmo de los investigadores. Es en estos casos, que constituyen más la regla que la excepción, cuando es indispensable (ética y científicamente) realizar (y exigir) el diseño, la conducción y la publicación formal de un estudio clínico experimental; el ensayo clínico controlado por sorteo es el diseño metodológico más confiable para distinguir si el beneficio atribuible a un tratamiento es real o sólo un espejismo.

En la historia de la medicina abundan los bochornosos ejemplos de tratamientos (novedosos en su época) a los que se le atribuyeron espectaculares beneficios (por lo que multitudes las recibieron) y que no fue hasta que un ensayo clínico aleatorizado los

puso en su justo lugar: en el archivo de terapias inútiles e, incluso, dañinas. Tal es el caso de la congelación gástrica en el tratamiento de la úlcera péptica, de la ligadura de la arteria mamaria interna en la prevención de la angina de pecho recurrente, de los esteroides en los pacientes con sepsis grave, de la plasmaféresis en los pacientes con polimiositis, entre muchos otros ejemplos.

También hay ejemplos de intervenciones terapéuticas o profilácticas ampliamente arraigadas en la práctica médica cotidiana actual pero cuyo beneficio real ha sido, en mayor o menor grado, cuestionado; tal como sucede con la vacuna antituberculosa con el bacilo de Calmette-Guerin, el uso de antibióticos en la otitis media aguda no complicada en los niños o el internamiento de enfermos en una unidad coronaria. Sin embargo, estos son ejemplos en los que resulta casi imposible (por razones éticas o logísticas) llevar a cabo en la actualidad un estudio clínico experimental, por lo que parece que estamos condenados a no resolver el dilema de su verdadero costo-beneficio. De ahí que hay quien recomienda que se realicen los ensayos clínicos controlados de manera temprana en la fase de evaluación de nuevas terapias, antes de que sean adoptadas de manera irreflexiva por la comunidad médica. Sin embargo, también debe tomarse en cuenta que lo común es que los ensayos clínicos sean estudios costosos y logísticamente complejos por lo que deben plantearse teniendo información completa de su farmacología, de su toxicidad y preliminar de su posible eficacia.

¿Qué determina la validez de sus resultados?

Esta pregunta tiene que ver con la confianza que lleguemos a tener de que los resultados del ensayo clínico revelen con exactitud la dirección y magnitud de lo que les pasa a los sujetos bajo estudio y que esto sea realmente atribuible al tratamiento en prueba. Es necesario que tengamos los elementos para evaluar si los resultados representan una estimativa no sesgada del efecto de la intervención o, si por el contrario, son conclusiones falsas determinadas por algún error sistemático. A continuación se describen los principales componentes, porque no son los únicos, que determinan la validez interna de un ensayo clínico.

La asignación del tratamiento en prueba por medio de un sorteo

La decisión de a cuál grupo pertenecerá cada uno de los individuos participantes en el estudio debe ser aje-

na a cualquier predisposición, o prejuicio, del investigador y la mejor manera de lograrlo es mediante el empleo de un proceso aleatorio: es únicamente el azar quien se encarga de asignar a los sujetos (figura 1).

En la década de los setenta y principios de los ochenta era común que a los pacientes con insuficiencia arterial cerebral se les practicara una cirugía de revascularización extra-intracranéa mediante la colocación de un puente en el que se unía una rama de la arteria carótida externa (la temporal superficial) con una rama cortical de la arteria carótida interna (la cerebral media). Los cirujanos que la practicaban entonces lo hacían con mucho entusiasmo convencidos de que su procedimiento realmente prevenía la oclusión de las arterias cerebrales en sus nobles pacientes. Esta creencia se basaba en la observación repetida de que grupos de pacientes así operados con menor frecuencia sufrían eventos isquémicos cerebrales en comparación con los enfermos que no eran sometidos a la cirugía. El detalle importante es que estas eran observaciones provenientes de la práctica clínica cotidiana, es decir, no se habían obtenido de un estudio planeado en donde un grupo de investigadores clínicos hubieran decidido, mediante un procedimiento semejante

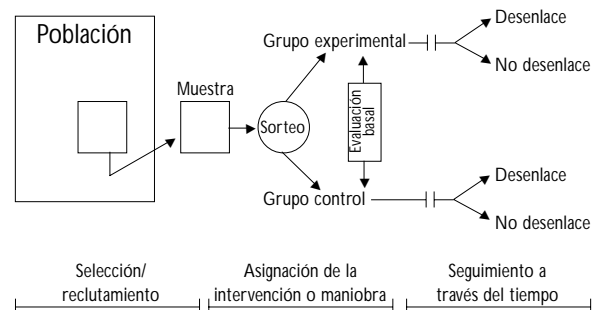


FIGURA 1. DIAGRAMA DE LA ESTRUCTURA DE UN ENSAYO CLÍNICO CONTROLADO POR SORTEO (ALEATORIZADO).

LA POBLACIÓN DE INTERÉS SE DEFINE CON BASE EN LOS CRITERIOS DE INCLUSIÓN/EXCLUSIÓN Y LA MUESTRA ESTUDIADA GENERALMENTE SE SELECCIONA Y RECLUTA POR SU ACCESIBILIDAD. LA ASIGNACIÓN DE LOS PARTICIPANTES AL GRUPO EXPERIMENTAL (QUIEN RECIBE LA INTERVENCIÓN BAJO EVALUACIÓN) O AL GRUPO CONTROL (QUIEN RECIBE LA INTERVENCIÓN DE CONTRASTE: EL PLACEBO O LA MANIOBRA HABITUAL) LA REALIZA EL PROPIO INVESTIGADOR MEDIANTE UN PROCEDIMIENTO ALEATORIO (SORTEO). EL EFECTO DE LA INTERVENCIÓN EN ESTUDIO SE MIDE AL COMPARAR LA INCIDENCIA DEL DESENLACE DE INTERÉS EN EL GRUPO EXPERIMENTAL CON LA DEL GRUPO CONTROL

al de un “volado”, quiénes era operados y quiénes no. En el quehacer clínico cotidiano no es nada raro que el cirujano ofrezca sus servicios (o que el paciente se deje operar) preferentemente a pacientes que están relativamente en mejores condiciones que otros; es decir, que desde antes de la cirugía gozan de un inherente mejor pronóstico, en cuanto al riesgo de oclusión arterial, de complicaciones de la cirugía e, incluso, de supervivencia; en cambio (y por razones totalmente entendibles), no suelen hacerlo con pacientes más deteriorados, en quienes hay menos posibilidades de que su intervención quirúrgica resulte exitosa. Esta práctica puede hacer entonces que, como los pacientes con menos posibilidades de una buena evolución son incluidos preferentemente en el grupo (cohorte) control (o de contraste), un procedimiento ineficaz parezca (de manera distorsionada) como todo lo contrario. Es justamente mediante el sorteo (también conocido como asignación aleatoria de la maniobra, que sólo puede hacerse en un estudio experimental) que se pretende que todos los pacientes tengan la misma oportunidad de recibir el tratamiento en prueba; de tal suerte que las diferencias en el desenlace clínico entre los sometidos (o no) al tratamiento sean explicables únicamente por este hecho y no por otros determinantes ajenos a la maniobra en estudio. Resulta interesante que fue mediante un enorme estudio experimental en humanos, aleatorizado, multicéntrico y publicado en 1987, que se logró demostrar tanto que a los pacientes así operados en realidad les va peor que a los tratados únicamente con medicamentos en el periodo posoperatorio inmediato, como que la evolución clínica a largo plazo era idéntica entre los sometidos o no a la revascularización quirúrgica.

La evolución clínica de los enfermos obedece a múltiples causas, el tratamiento sólo es una de ellas; de tal forma que es común que la gravedad de la enfermedad, la presencia de comorbilidad y toda una gama de determinantes pronósticos (unos que podemos identificar y otros muchos que desconocemos) hagan que tengamos una falsa impresión de las bondades (perjuicios) de los tratamientos que ofrecemos. Es por ello que las decisiones de los médicos, basadas en información proveniente de estudios observacionales (o de experimentales no aleatorizados) estén propensos a caer en este tipo de errores. De hecho, se ha documentado que los estudios en los que el tratamiento se ha asignado por cualquier otro método que no haya sido un sorteo tienden a mostrar efectos de mayor beneficio (y frecuentemente falsos) que lo observado en los ensayos clínicos aleatorizados. La ventaja de la aleatorización es que logra, si el tamaño de la muestra es suficientemente grande, que los deter-

minantes de la evolución de los enfermos (tanto los conocidos como los desconocidos) estén distribuidos de manera equilibrada entre el grupo de enfermos que reciben la maniobra bajo estudio (grupo experimental) y los que no la reciben (grupo control).

Seguimiento completo de los individuos en estudio

Este punto tiene que ver tanto con que se logre un seguimiento completo de todos los sujetos participantes en el estudio durante el tiempo programado como con que se incluyan en el análisis a todos los individuos respetando su pertenencia al grupo (experimental o control) al que fueron originalmente asignados.

Todo paciente que ingresa al ensayo debe ser tomado en cuenta en el análisis y conclusiones; de lo contrario, si un número importante de ellos se reporta como “sin seguimiento suficiente” la validez del estudio se verá seriamente cuestionada. Mientras más individuos se pierden mayor será la posibilidad de sesgo en el estudio debido a que los pacientes que no completan su seguimiento pueden tener un pronóstico diferente al de quienes sí permanecen hasta el final del estudio. Las pérdidas en el seguimiento pueden obedecer a la aparición de eventos adversos al medicamento en estudio, a una mala respuesta al mismo, a muerte, etcétera o, al contrario, porque tienen una evolución particularmente benigna y su bienestar hace que no regresen a sus evaluaciones. Así, al ocurrir esto se puede tener una impresión distorsionada de las bondades (o de su ausencia) del tratamiento bajo prueba.

Tal como sucede en la práctica médica cotidiana, los pacientes en los estudios experimentales dejan de tomar sus medicamentos, ya sea por olvido o por efectos indeseables de los mismos. Excluir a estos individuos –que no se adhirieron de manera completa a los lineamientos del protocolo– del análisis puede igualmente generar sesgos. Igual que en el punto anterior, es frecuente que esta falta de apego esté estrechamente vinculada con un pronóstico diferente al de los enfermos cumplidores, tal como se ha llegado a constatar en algunos ensayos clínicos donde incluso los pacientes a quienes se les asignó recibir el placebo tuvieron una mejor evolución que los que mostraron un buen apego a su ingesta en comparación con los incumplidos. Al excluir a éstos del análisis uno corre el riesgo de sólo evaluar a pacientes con un mejor pronóstico, alterando la condición de equidad entre grupos dada por la asignación por sorteo de la maniobra.

Es justamente este principio de incluir en el análisis a todos los pacientes, tal como fueron asignados de

acuerdo con el sorteo, el que define el análisis de "intención a tratar". Esta estrategia metodológica conserva el valor de la aleatorización, es decir, el de lograr que los pacientes con diferentes pronósticos de la enfermedad queden igualmente distribuidos en los grupos (brazos) del ensayo.

Evitar que las expectativas de los pacientes y de sus evaluadores influyan en la medición de los desenlaces

Igual que los investigadores, los enfermos que saben que están recibiendo un nuevo tratamiento en experimentación generalmente tienen una idea prejuiciada sobre su eficacia. Estas expectativas, ya sea optimistas o desfavorables, pueden llegar a distorsionar la medición de los resultados particularmente cuando su evaluación es mediante indicadores propicios a ser influidos por la subjetividad del informante o del propio evaluador ("datos blandos"). Esto puede ocurrir cuando se miden síntomas (como el dolor), sentimientos de bienestar (como la calidad de vida), o signos clínicos sin definiciones suficientemente objetivas y poco reproducibles. En estas condiciones, y cuando los investigadores conocen quién de sus pacientes del estudio recibe el tratamiento experimental y quién no, se pueden sesgar los resultados al hacer una búsqueda diferencial de los desenlaces o al dar interpretaciones diferentes a los hallazgos.

La mejor manera de evitar este tipo de errores es mediante el enmascaramiento, que no es otra cosa que tratar de que ni el enfermo ni el investigador que lo evalúa (doble cegamiento) sepan si aquél se encuentra recibiendo la terapia en prueba o la intervención de contraste. Esto habitualmente se logra administrando un placebo (o la terapia habitual) con apariencia, sabor y textura indistinguibles de la terapia experimental. De no ser posible esto, y si el tipo de desenlaces a medir lo amerita, habrá que diseñar una estrategia en la que el investigador que los evalúa desconozca a qué grupo de estudio pertenecen sus pacientes.

Brindar igual atención médica, fuera de la terapia en estudio, a los grupos del ensayo

Si los cuidados médicos, incluyendo otras terapias diferentes a la que se estudia, se dan de manera diferente en el grupo experimental que en el control se corre el riesgo de comprometer seriamente los resultados del ensayo al cometer el sesgo denominado de "cointervención"; es decir, será difícil distinguir qué tanto de la diferencia (o no diferencia) observada entre los grupos es efecto de la terapia en estudio vs. consecuencia de una atención médica diferencial. Una

manera de evitar este tipo de error sistemático puede ser mediante el diseño de un ensayo doble ciego.

¿Qué determina que sus resultados sean aplicables a otros grupos de enfermos?

Lo común es que una de las principales motivaciones que lleva a un investigador a realizar un ensayo clínico es el deseo de que la terapia experimental eventualmente sea utilizada por otros médicos, en otros sitios, como una medida útil y que se brinde para el beneficio real del mayor número de enfermos. Esto dependerá fundamentalmente de tres aspectos.

Primero, de qué tanto se parezcan los enfermos que estudia a los enfermos en quienes posteriormente se quiera extrapolar la experiencia obtenida en el ensayo clínico. Esto depende del rigor en la selección de los sujetos a incluirse en el experimento; mientras más criterios de inclusión sea necesario cumplir menos posibilidades de que los resultados sean aplicables a otras poblaciones de enfermos. De hecho, el investigador suele enfrentarse al dilema de que por querer estudiar a una población relativamente homogénea de enfermos que le permita incrementar la eficiencia de su estudio (es decir, demostrar con claridad un efecto benéfico de la terapia experimental con el menor número de enfermos y costos posibles) establece criterios de inclusión al estudio demasiado estrictos que lo alejan de las características del quehacer médico cotidiano: de la "vida real".

Segundo, del desenlace que se haya elegido para valorar la eficacia de la terapia experimental. Habitualmente y por razones de eficiencia, el investigador suele elegir mediciones fisiológicas, bioquímicas, microbiológicas (o de otra estirpe) más fáciles de medir, en corto tiempo y con menos pacientes como sustitutos de desenlaces clínica o socialmente más importantes (como puede ser la calidad de vida o la supervivencia) bajo el supuesto de que aquéllas pueden predecir éstos. Así, no es raro que en los estudios experimentales se evoque el beneficio de una nueva terapia porque ésta mejora las pruebas de función respiratoria, o porque disminuye la carga viral en el plasma, o los niveles séricos de colesterol (por mencionar sólo algunos ejemplos), asumiendo que al lograrlo los enfermos vivirán más tiempo y mejor. Es necesario ser cauteloso en la selección de los resultados a evaluar y que sean realmente importantes para los individuos y para la sociedad; de lo contrario, se corre el riesgo de promulgar la bondad de una terapia nueva porque mejora una medición intermedia, pero que a la larga se demuestre su efecto nocivo en desenlaces clínicamente más importantes o incluso en la supervivencia de los

sujetos. Ejemplos de ello es lo sucedido con el uso de hipolipemiantes (como el clofibrato y ciertas estatinas) en la reducción del colesterol sérico o de ciertos antiarrítmicos (como la encainida) posterior a un infarto al miocardio.

Tercero, de que la intervención que se evalúa sea única y precisa, como lo es un medicamento. En este caso su reproducción por otros investigadores (u otros médicos) es fácil. Por el contrario, hay tratamientos que implican varios elementos cambiantes o que demandan un grado de habilidad muy particular de quien realiza la maniobra (como sería el caso del manejo médico de un cierta entidad clínica en un determinado ambiente, una intervención quirúrgica, un programa educativo, etc.) cuyos efectos no necesariamente serán los mismos cuando son realizados por otros médicos o investigadores.

¿Cómo se analizan sus resultados?

Lo más común es que los ensayos clínicos controlados midan la incidencia de algún evento en los grupos de individuos seguidos en un determinado lapso y que este evento se exprese de manera dicotómica (es decir, la presencia o no del desenlace: infarto al miocardio, recurrencia de una neoplasia, muerte, etc.) como la proporción de sujetos que llegan a presentarlo. Pongamos como ejemplo un estudio en el que 20% (0.20) de los enfermos en el grupo control fallecieron en contraste con sólo 15% (0.15) de los que recibieron el tratamiento en evaluación. En el cuadro I se resume la forma como se puede presentar el efecto de éste, a saber:

1. La diferencia absoluta (o la reducción del riesgo absoluto -RRA-), que se obtiene al sustraer la proporción de individuos que fallecieron en el grupo experimental (Y) de la proporción de individuos que lo hicieron en el grupo control (X): $X - Y = 0.20 - 0.15 = 0.05$ (5%).
2. El riesgo relativo (RR), es decir, el riesgo de morir en los pacientes sometidos a la terapia experi-

mental con relación al de los pacientes en el grupo control: $Y/X = 0.15/0.20 = 0.75$.

3. El complemento del riesgo relativo (o la reducción del riesgo relativo -RRR-) que se expresa como un porcentaje: $[1 - (Y/X)] \times 100 = [1 - 0.75] \times 100 = 25\%$. Esta cifra significa que el nuevo tratamiento reduce el riesgo de morir en 25% con relación a lo que ocurre en los pacientes del grupo control; mientras mayor sea la RRR mayor es la eficacia del tratamiento.
4. El número necesario de pacientes a tratar (NNT) indica si el beneficio ofrecido por la nueva terapia retribuye el esfuerzo y costo en su adquisición o implantación. Por ejemplo, una reducción de 25% en el riesgo de morir puede parecer impresionante, pero su impacto en el paciente o en la práctica clínica puede, sin embargo, ser mínimo. La utilidad de un tratamiento está no sólo en función de la reducción relativa del riesgo sino también del riesgo del desenlace adverso que se quiere prevenir (en nuestro ejemplo, la muerte); de tal forma que mientras menor sea este riesgo mayor será el número necesario de enfermos a tratar con la nueva terapia para prevenir una muerte (es decir, menor su impacto). En el cuadro II se ilustran dos circunstancias: una en la que el riesgo de muerte en una población de enfermos, en un determinado periodo de tiempo, es sólo de 1%; en contraste con otra población con un riesgo mayor (de 10%). En el primer caso, la nueva terapia reduciría el riesgo de fallecer en 25%, es decir, una reducción del riesgo absoluto de 0.0025 (o 25 muertes en 10 000 pacientes tratados). El número necesario de pacientes a tratar se obtiene al calcular la inversa de esta reducción del riesgo absoluto ($1/0.0025 = 400$); así, sería necesario tratar 400 pacientes durante un tiempo determinado para salvar una sola vida. En cambio, en el segundo caso, una reducción relativa de 25% de muerte en una población en mayor riesgo de morir (de 10%) lleva a una reducción del riesgo absoluto de 0.025 (o 25 muertes en 1 000

Cuadro I
MEDICIONES DEL EFECTO DE UN TRATAMIENTO EN EVALUACIÓN*

Reducción del riesgo absoluto (diferencia de riesgos) (RRA)	$X - Y$	$0.20 - 0.15 = 0.05$ o 5%
Riesgo relativo (RR)	Y/X	$0.15/0.20 = 0.75$
Reducción del riesgo relativo (RRR)	$1 - (Y/X)$ X 100 o $(X - Y)/X$ X 100	$1 - 0.75$ X 100 = 25% $0.05/0.20$ X 100 = 25%
Número de pacientes necesario a tratar para prevenir un evento (NNT)	$1/(X - Y)$	$1/(0.20 - 0.15) = 20$

* X= riesgo del evento en los pacientes sin el tratamiento (grupo control) 20/100= 0.20 o 20%
Y= riesgo del evento en los pacientes con el tratamiento (grupo experimental) 15/100= 0.15 o 15%

Cuadro II
EJEMPLO DEL EFECTO DEL PRONÓSTICO DE LA ENFERMEDAD EN EL NÚMERO
NECESARIO DE PACIENTES A TRATAR

Riesgo de muerte en los pacientes en el grupo control (riesgo basal): X	1% o 0.01	10% o 0.10
Riesgo relativo en los pacientes en el grupo experimental: Y/X	75% o 0.75	75% o 0.75
Reducción del riesgo relativo $1 - (Y/X) \times 100$ o $(X - Y)/X \times 100$	25%	25%
Riesgo de muerte en los pacientes en el grupo experimental: Y	$0.01 \times 0.75 = 0.0075$	$0.10 \times 0.75 = 0.075$
Reducción del riesgo absoluto $X - Y$	$0.01 - 0.0075 = 0.0025$	$0.10 - 0.075 = 0.025$
Número de pacientes a tratar para evitar una muerte $1 / (X - Y)$	$1 / 0.0025 = 400$	$1 / 0.025 = 40$

pacientes tratados) de tal suerte que se tendría que tratar a sólo 40 individuos para salvar una vida ($1/0.025 = 40$). Este ejemplo señala un elemento clave en la decisión de implantar una nueva terapia: considerar la magnitud del riesgo del desenlace adverso en los pacientes no tratados con ella. Para una misma reducción del riesgo relativo mientras mayor sea la probabilidad de padecer un evento indeseable si no se trata, mayor será el beneficio con la nueva terapia y menor el número de pacientes que tendremos que tratar para prevenir un evento. Cabe mencionar que en esta decisión debe considerarse también el costo, la factibilidad y grado de seguridad de la nueva terapia en cuestión. Si, por ejemplo, ésta conlleva un riesgo de 10% de un cierto efecto adverso por el tipo de medicamento, en la población de enfermos con un bajo riesgo de muerte, se tendrían 40 individuos sufriendo el efecto indeseable de la droga por cada vida salvada contra sólo cuatro si se le da a la población de enfermos con un mayor riesgo de muerte. Será, finalmente, el costo de la terapia, así como la gravedad y tipo de consecuencias del efecto adverso del medicamento, lo que nos haga decidir si 4, 40, 400, 4 000 o 40 000 pacientes necesarios a tratar es una cantidad importante o no.

Otro aspecto importante a evaluar en la medición de los resultados de un ensayo clínico es qué tan precisa fue la estimativa del efecto del tratamiento. La verdadera reducción del riesgo es algo que nunca llegaremos a conocer; lo más que podemos alcanzar es llegar a estimarla y el mejor estimado es el valor observado en el estudio (el llamado "estimado puntual"). Mediante el cálculo estadístico del intervalo de confianza (IC) uno puede establecer una zona de valores, alrededor de este estimado puntual, donde pudiera encontrarse el verdadero valor poblacional. Lo habitual, aunque arbitrario, es que se use el IC al 95%; es decir, se establece el intervalo que incluye al verdade-

ro valor de la reducción del riesgo relativo en 95% de las veces. Será raro (con una probabilidad de sólo 5%) que éste se encuentre más allá de los límites de este intervalo; lo que va de acuerdo con lo que convencionalmente se establece como el nivel de significancia estadística (o valor de p).

Examinemos un ejemplo. Si en un ensayo clínico se aleatorizaron 100 pacientes al grupo experimental y 100 pacientes al grupo control y se llegan a observar 15 muertes en el primer grupo y 20 en el segundo, el cálculo del estimado puntual de la reducción del riesgo relativo sería de 25%: $X = 20/100$ o 0.20, $Y = 15/100$ o 0.15, y $[1 - (Y/X)] \times 100 = [1 - 0.75] \times 100 = 25\%$. Por lo anteriormente dicho, pudiera ser que el verdadero valor de la RRR fuera significativamente menor o mayor que esta cifra de 25%, la que se obtuvo de una diferencia de tan sólo cinco muertes entre los dos grupos, y hasta pensarse que el tratamiento no fuera eficaz (una RRR de 0%) e incluso perjudicial (una RRR con un valor negativo). De hecho así es: estos resultados son consistentes tanto con una RRR de menos 38% (es decir, que a los pacientes que reciben el nuevo tratamiento tuvieran un riesgo de morir 38% mayor que los pacientes en el grupo control) y una RRR de casi 59% (es decir, que los pacientes en el grupo experimental tuvieran un riesgo menor de morir de casi 60%). En otras palabras, el intervalo de confianza al 95% (IC 95%) de este estimado de la RRR es de menos 38% a 59%, y el estudio en realidad no nos ayuda a decidir si el nuevo tratamiento es útil. Ahora, supongamos que en lugar de 100, se hubieran sorteado 1 000 pacientes por grupo y que se hubiera observado la misma proporción de desenlaces, es decir, 150 muertes en el grupo experimental ($Y = 150/1\ 000 = 0.15$) y 200 muertes en el grupo control ($X = 200/1\ 000 = 0.20$). Nuevamente, el estimado puntual de la RRR sería 25%: $[1 - (Y/X)] \times 100 = [1 - (0.15/0.20)] \times 100 = 25\%$. Se puede predecir que, en este ensayo con un número bastante mayor de individuos estudiados, el verdadero valor de la RRR esté más cerca del valor de 25%;

efectivamente, el IC 95% de la RRR en estos datos va de 9 a 41%.

Lo que estos ejemplos nos muestran es que mientras mayor sea el número de participantes en el ensayo clínico mayor será el número de eventos observados y mayor la certidumbre de que el valor verdadero de la RRR (o de cualquier otra medida de eficacia) está cercano al valor obtenido en el estudio. En el segundo ejemplo anterior, el valor posible más bajo de la RRR fue de 9% y el más alto, de 41%. El estimado puntual (en este caso 25%) es probablemente el que más se acerca al valor real (poblacional) de la RRR. Mientras más lejanos los valores del estimado puntual menos probable es que sean consistentes con el valor observado y aquéllos más allá de los límites del IC son valores con muy pocas posibilidades de representar la verdadera RRR, dado el estimado puntual (la RRR observada).

Siendo que mientras más grande es el tamaño de la muestra estudiada más estrecho es el IC, ¿cuál es un número suficiente de sujetos a estudiar? Cuando en un ensayo clínico se concluye que los resultados fueron positivos (es decir, que muestra que el tratamiento fue eficaz) es porque el valor inferior del IC del estimado de la RRR (el más bajo y aún consistente con los resultados del estudio) es “clínicamente importante”, es decir, lo suficientemente grande como para que el tratamiento sea prescrito a los enfermos. En esta circunstancia puede decirse que el tamaño de la muestra estudiada fue suficiente. Si, en cambio, se considera que el límite inferior de este IC no es “clínicamente importante”, entonces el estudio no puede ser considerado como definitivo, a pesar de que la diferencia entre tratamientos sea estadísticamente significativa (es decir, que excluya que la RRR sea de cero).

El IC también ayuda a interpretar los estudios con resultados “negativos”, en los que los autores concluyen que el tratamiento experimental no mostró ser mejor que el del grupo control: si el límite superior del IC muestra que la RRR pudo haber sido “clínicamente significativa” quiere decir entonces que el estudio no logró excluir un efecto importante de la nueva terapia. En el primer ejemplo anterior, el límite superior del IC fue una RRR de 59%; de tal forma que el beneficio del tratamiento pudo haber sido sustancial y se concluiría que aunque los investigadores no lograron demostrar que la terapia experimental fuera mejor que el placebo tampoco pudieron descartar un efecto importante de ella. Esto es lo que se conoce como un estudio con poca precisión (poder o sensibilidad) por haber estudiado un tamaño de muestra insuficiente.

Cabe señalar la asociación entre el IC y el valor de significancia estadística, o valor de p . Si éste es

igual o mayor a 0.05 (5%) quiere decir que el límite inferior del IC 95% de la RRR es el valor nulo (cero), o un valor negativo, y el RR, de uno, o menor; es decir, que no se logra descartar la hipótesis nula de no diferencia entre el tratamiento experimental y el placebo (o terapia en el grupo control). Conforme el valor de p disminuye (menor a 5%) el límite inferior del IC 95% de la RRR es mayor a cero y se dice que la diferencia entre tratamientos mostró ser estadísticamente significativa; es decir que, si la hipótesis nula es cierta, es muy pequeña la posibilidad de ver una diferencia de esta magnitud o mayor. La prueba estadística para el cálculo del valor de p dependerá del tipo de variable, de cómo esté expresado el evento desenlace (dicotómica, ordinal, continua) y su descripción no es motivo de este capítulo.

¿Es ético hacerlos?

Realizar un ensayo clínico controlado puede generar inquietudes de orden ético. Algunos pacientes se pueden llegar a incomodar por el hecho de ser interrogados o examinados para propósitos diferentes a los de su estricta atención médica, o bien, al saber que investigadores que no son sus médicos llegan a tener acceso a la privacidad de su expediente clínico. El diseño experimental significa que los tratamientos son asignados por un investigador, no por su médico tratante ni elegidos por el propio paciente. Cuando la asignación es mediante un sorteo o cuando hay un doble enmascaramiento (es decir, que ni el paciente ni sus médicos que lo atienden saben a qué medicamento se está sometiendo) los aspectos éticos se vuelven más complejos e incluso más importantes que los científicos.

Para que un estudio clínico experimental sea éticamente justificable debe cumplir la premisa de que, al momento del inicio del estudio, no haya evidencia de que alguno de los tratamientos ofrecidos en cada brazo (o grupo del ensayo, incluyendo el control) sea superior al(os) otro(s). No es raro que al plantearse la conducción de un estudio clínico experimental surja la inquietud de que a los participantes no se les esté ofreciendo la mejor opción terapéutica conocida; es decir, que se les está privando de una terapia experimental novedosa y superior a lo ya conocido o, por el contrario, que se les esté exponiendo a un nuevo tratamiento de dudosa utilidad y seguridad. En efecto, el estudio no sería ético si se tiene la suficiente información de que uno de los tratamientos es más eficaz (o dañino); pero si en verdad se desconoce si el tratamiento en evaluación es el óptimo, no habría justificación para dicha preocupación; de hecho, podría

argumentarse la falta de ética al estar prescribiendo abierta y rutinariamente un tratamiento cuya utilidad (y seguridad) relativa se desconocen porque se adoptó de manera prematura, antes de su evaluación formal. En este punto puede existir la dificultad de definir a partir de cuándo la evidencia es "suficiente" o "convinciente" para justificar o no la realización de un estudio experimental en humanos; desde luego que la perspectiva de un clínico (que procura el máximo beneficio de un individuo en particular) será diferente que la del investigador clínico (quien busca la "verdad" con rigor científico). Por otra parte, es habitual que en el proceso de desarrollo de un nuevo medicamento se le exija a quien lo manufactura que presente información de su eficacia obtenida en pequeños estudios no controlados (la llamada fase II) antes de encaminarse a un ensayo clínico controlado (fase III).

Aunque estos dilemas éticos pueden llegar a ser de difícil solución, en muchos centros de investigación del mundo el consentimiento informado y los comités institucionales de investigación en humanos han venido a constituir importantes salvaguardas de la ética. El consentimiento informado es requerido por la mayoría de las agencias financiadoras de proyectos de investigación experimental en seres humanos así como por las revistas científicas y habitualmente es obligatorio que sea firmado por el participante. El consentimiento informado implica que al individuo se le da la oportunidad de preguntar y enterarse de la naturaleza y detalles del estudio y se hace de su conocimiento que goza de entera libertad para decidir si participa en el estudio clínico así como de que no habrá ningún tipo de represalia (o cambio en su atención médica) si decide no hacerlo, o si abandona el estudio antes de lo programado. Sin embargo, a pesar de estos requisitos, y dada la creciente complejidad de los protocolos de los estudios experimentales, no es rara la circunstancia en la que el sujeto participante no llegue a estar completa y claramente informado. Los comités institucionales de investigación en humanos son cuerpos colegiados que suelen estar constituidos por investigadores, médicos clínicos, enfermeras, administradores, abogados y representantes de la comunidad y su tarea consiste en revisar escrupulosamente el protocolo de investigación, constatando que el planteamiento de la pregunta de investigación y el diseño sean adecuados (después de todo no es ético realizar un estudio con importantes errores metodológicos) y asegurando la preservación de la integridad y seguridad de los participantes en el estudio y que la carta de consentimiento informado esté correctamente elaborada. El comité puede también solicitar que los datos

del estudio sean analizados periódicamente por un comité externo y ante la eventualidad de encontrar una clara diferencia en la eficacia entre los tratamientos, que se interrumpa anticipadamente el estudio y se publiquen los hallazgos.

Finalmente, cabe comentar sobre la justificación de cuándo dar un placebo a los enfermos en el grupo control. Esto ocurre en dos circunstancias: una, cuando no existe tratamiento alguno que convincentemente haya mostrado ser eficaz en modificar favorablemente el curso de la enfermedad; es decir, que en una buena práctica médica en condiciones habituales un médico no prescribiría medicamento específico alguno (tal es el caso de la evaluación del zanamivir, un inhibidor de la neuraminidasa, en el tratamiento de la influenza o del uso de la warfarina en la prevención de fenómenos embólicos en la fibrilación auricular sin valvulopatía). La segunda circunstancia es cuando sí existe un tratamiento estándar efectivo pero se desea evaluar si la adición de otro medicamento ofrece una ventaja adicional. En este caso a ambos grupos de enfermos se les administra la terapia estándar, pero al experimental se le adiciona el nuevo medicamento en estudio y, en cambio, al grupo control un placebo indistinguible del nuevo medicamento (tal es el caso de la evaluación de si los ácidos grasos 3-omega, como tratamiento complementario, mejoran la funcionalidad de los enfermos con artritis reumatoide activa que además reciben medicación antiinflamatoria).

En cambio hay circunstancias en las que sería una grave falta ética dar únicamente un placebo y privar de un tratamiento activo ya ampliamente establecido y aceptado como útil a los individuos en el grupo control. Difícilmente alguien que desea probar que una cierta terapia es eficaz para mejorar el pronóstico de los pacientes sintomáticos por infección por el VIH la compararía con un grupo de enfermos en quienes se omitiera el tratamiento antirretroviral estándar en la actualidad: dos análogos nucleósidos inhibidores de la transcriptasa reversa más un inhibidor de proteinasas.

¿Cuáles son sus ventajas y desventajas?

Respecto a los estudios observacionales de cohortes, la principal ventaja que ofrecen los ensayos clínicos controlados (especialmente los que son mediante un sorteo) es que asignan la maniobra experimental (el tratamiento en prueba) independientemente de determinantes pronósticos o de la selección de la muestra; de tal suerte que disminuye el riesgo de sesgos de susceptibilidad diferencial (confusores) o de un muestreo

distorsionado. Además, el diseño experimental facilita el enmascaramiento de los individuos participantes y de sus evaluadores (doble cegamiento) ya que en un estudio observacional los tratamientos generalmente son decididos y prescritos abiertamente por los médicos tratantes.

Por otra parte, la posibilidad real de realizar los estudios experimentales puede verse seriamente comprometida por razones de índole práctica (logística) de costos y de ética. Además, en un ensayo clínico se puede llegar a una situación tal de control de variables (homogeneidad de la muestra estudiada, regulación del apego al tratamiento, atención y seguimiento del enfermo, etc.) que haga que sus resultados se alejen de la realidad del quehacer clínico cotidiano, restringiendo seriamente su extrapolación al resto de los enfermos; es decir, que en la compulsión por diseñar estudios metodológicamente impecables no se logre responder a las preguntas originales.

A pesar de esto último, los ensayos clínicos aleatorizados permanecen como el estándar de oro (o pa-

radigma) en la evaluación de la utilidad de nuevos tratamientos. A pesar de no ser infalibles y de, en ocasiones, dejar resquicios de incertidumbre, los estudios experimentales han contribuido enormemente a la evolución hacia terapias y medidas profilácticas más eficaces y seguras.

Además, el entendimiento de los principios científicos que rigen a los ensayos clínicos controlados le ha ayudado al médico a adquirir la destreza para discriminar con rapidez los estudios clínicos cuyas conclusiones son poco sustentables, por ser estudios de mal diseño y realización, de aquellos realmente válidos y aplicables a su propia práctica. En la medida en que los médicos sólo demos credibilidad a los trabajos publicados con solidez científica (es decir, que nos transformemos de lectores pasivos y puramente receptivos en evaluadores críticos) elevaremos la calidad de nuestras decisiones y el grado de beneficio que reciben nuestros enfermos.