

Guía para la lectura crítica de un artículo sobre efectos nocivos

Fernando J. García López¹, Vicente Barrio Lucia²

¹ Unidad de Epidemiología Clínica. Hospital Universitario Puerta de Hierro. Majadahonda, Madrid

² Servicio de Nefrología. Hospital Infanta Sofía. San Sebastián de los Reyes, Madrid

Nefrología Sup Ext 2010;1(4):7-15

doi:10.3265/NefrologiaSuplementoExtraordinario.pre2010.Nov.10701

En muchas ocasiones se le plantea al clínico la cuestión de si su paciente presenta una reacción adversa de algún tratamiento recibido o si la puede presentar si empieza a recibirlo. Se trata de una preocupación justificada, pues se calcula que la incidencia de efectos adversos en España relacionados con la asistencia hospitalaria afecta al 8,4% de los pacientes, 1,2 efectos adversos por cada 100 días-paciente. Pero la determinación de si un paciente puede sufrir un acontecimiento adverso como consecuencia de un tratamiento no es nada fácil, pues la aparición de la reacción adversa durante el tratamiento no es prueba suficiente de que esa reacción sea consecuencia directa de dicho tratamiento.

Ante este problema, el clínico debe enfrentarse a dos preguntas:

1. ¿Causa ese tratamiento (medicamento, intervención quirúrgica u otro tratamiento) ese efecto adverso en algunos pacientes?
2. ¿Puede causar ese tratamiento (medicamento, intervención quirúrgica u otro tratamiento) el efecto adverso en su paciente?

Para evaluar un artículo que aborde un problema de acontecimientos adversos, se recomendarán una serie de pasos para ayudar a responder a esas dos preguntas. No se entrará en cómo se ha seleccionado el artículo, que es materia que requiere una atención especial y que no es objeto de esta guía. Estas recomendaciones se basan en las guías para la lectura crítica de la bibliografía médica del grupo de Medicina Basada en la Evidencia de la Universidad de McMaster y en sus desarrollos posteriores.

Correspondencia: Fernando J. García López

Unidad de Epidemiología Clínica.

Hospital Universitario Puerta de Hierro. Majadahonda. Madrid
fjgarcia@medynet.com

¿SON VÁLIDOS LOS RESULTADOS DEL ESTUDIO?

Criterios primarios

Antes que nada, se trata de formular las preguntas fundamentales que permitirán eliminar aquellos artículos sin validez suficiente, y por tanto, incapaces de ayudar a la contestación de las preguntas.

¿Estuvo la pregunta de investigación claramente definida?

Al igual que en otras guías previas^{1,2}, una pregunta bien definida significa que, al menos, están claramente señalados tres aspectos importantes: la justificación, los componentes básicos del estudio (pacientes, intervención, comparación y resultado) y los objetivos e hipótesis.

1. En primer lugar, la justificación del estudio. Se deben explicar los antecedentes y la necesidad del estudio, qué motivos justifican su diseño y ejecución.
2. En segundo lugar, los cuatro componentes básicos del estudio, según las siglas PICO, de Pacientes, Intervención, Comparación y *Outcome* (desenlace o resultado):
 - Pacientes: criterios de elegibilidad, por ejemplo, de inclusión (que se relacionan con la edad, el sexo, el diagnóstico clínico y las comorbilidades asociadas) y de exclusión (que se suelen establecer para preservar la seguridad de los participantes), y el ámbito y lugar en donde se recogen los datos (país, ciudad, ámbito hospitalario, de atención primaria o de la población general) y si fue un estudio monocéntrico o multicéntrico.
 - Intervención: la intervención de la que se quiere conocer sus efectos nocivos hipotéticos, que puede ser una intervención farmacológica o de otro tipo; se debería definir el momento y modo de administración, su duración, así como factores asociados pertinentes, como cointervenciones.

- Comparación: igual que en intervención, pero referido al grupo control. En particular, tipo de comparación (placebo, tratamiento activo).
- Desenlace o resultado: el suceso o acontecimiento adverso que se pretende estudiar.

Se deben indicar las variables de resultado primarias y secundarias. La variable de resultado primaria es la variable prefijada de mayor importancia para el estudio y suele ser la que determina el tamaño de la muestra. Todos los estudios con relevancia clínica deben tener una variable de resultado primaria preestablecida.

Se debe explicar con claridad la definición de las variables de resultado, sus instrumentos de medida y sus criterios de validación, el momento y periodicidad de las mediciones, y, en caso de ser importante (por ejemplo, en aspectos que dependen del juicio del observador, como la interpretación de una prueba de imagen o de una biopsia), quiénes y cuántos practican la evaluación, cómo se compara la información aportada por los evaluadores y si están enmascarados con respecto a la asignación de los participantes en el ensayo.

3. En tercer lugar, los objetivos específicos y la hipótesis del ensayo:
Los objetivos son las preguntas que el ensayo pretende contestar, que, por lo general, se centran en la incidencia de efectos adversos de una intervención en comparación con otras intervenciones distintas. Las hipótesis son preguntas establecidas de antemano, más específicas y sujetas a evaluación estadística, que se contrastan para poder cumplir los objetivos.

¿Se utilizaron grupos de comparación claramente definidos que fueron similares en todos los aspectos importantes salvo en la exposición al tratamiento?

Ésta es la pregunta más importante, pues de la semejanza entre el grupo de intervención y el grupo de comparación en todas las variables distintas a la exposición en estudio depende la validez del estudio. La respuesta a esta pregunta depende fundamentalmente del tipo de diseño del estudio. Cuatro tipos de diseño pueden abordar el estudio de los acontecimientos adversos.

1. El primer diseño es el más válido, el ensayo controlado con distribución aleatoria. Por definición, si la distribución de los pacientes entre los grupos se efectúa al azar, y esa distribución se ha efectuado verdaderamente al azar, se igualan todas las posibles variables pronósticas que pueden influir en la aparición de las reacciones adversas. Cualquier aumento no explicado por el azar en la incidencia de los

efectos nocivos en el grupo de tratamiento con respecto al grupo de control, a igualdad del resto de variables, podrá achacarse razonablemente al efecto del tratamiento. Por ejemplo, el ensayo TREAT, en pacientes con diabetes de tipo 2 con enfermedad renal crónica sin diálisis y con anemia moderada, estudió el efecto del tratamiento con darbepoetina para obtener una hemoglobina de 13 g/dl en comparación con un grupo control que sólo recibían darbepoetina de rescate cuando la hemoglobina bajaba de 9,0 g/dl; la incidencia de ictus letal y no letal en el grupo con dianas altas de hemoglobina fue del 16% frente al 9% en el grupo con diana baja, casi el doble (razón de riesgos instantáneos, *hazard ratio*, 1,92, con un intervalo de confianza del 95% de 1,38 a 2,68)³. Es poco probable atribuir al azar estos resultados, que sugieren un efecto nocivo importante.

El problema con los ensayos controlados y aleatorizados es que sólo permiten detectar efectos adversos relativamente frecuentes; o los infrecuentes si el tamaño de la muestra es muy alto. Si un medicamento ocasiona una reacción adversa en un 2% de los casos, y en los pacientes del grupo control esa reacción adversa se observa en un 1% de casos, se necesitaría un ensayo con 5.030 sujetos, 2.515 asignados al medicamento y otros tantos al grupo control para poder detectar esa diferencia. La mayor parte de los ensayos clínicos se diseñan con hipótesis de eficacia, que suelen formularse con diferencias mayores en los porcentajes de éxito entre los grupos de tratamiento y control, por lo que no poseen potencia estadística suficiente para detectar reacciones adversas infrecuentes. Y si el efecto adverso es más frecuente, pero de aparición tardía, tras muchos años después del inicio de la administración del tratamiento, un ensayo clínico tampoco lo detectaría. Por otra parte, el diseño de un ensayo con objeto de detectar acontecimientos adversos como objetivo principal suele ser inaceptable desde el punto de vista ético.

Una alternativa puede provenir de las revisiones sistemáticas de ensayos clínicos, que pueden reunir un número alto de sujetos participantes en los ensayos individuales. Pero la suma de muchos ensayos pequeños no es igual que un ensayo grande y puede seguir ocultando la aparición de reacciones adversas. Además, la vigilancia y notificación de los efectos adversos en los ensayos clínicos no siempre se hace con el mismo rigor que la detección de los sucesos que componen las variables de resultado de eficacia. En conclusión, los ensayos clínicos son el diseño óptimo para la detección de acontecimientos adversos pero en muchas ocasiones son impracticables, de modo que los efectos nocivos deben estudiarse con otros diseños.

2. El segundo diseño, el estudio de cohortes, se asemeja al ensayo controlado y aleatorizado, pero la asignación de los pacientes a los grupos no se efectúa al azar sino que se hace por decisión de los clínicos o de los sujetos participan-

tes. Esto plantea dudas acerca de que el grupo expuesto a la intervención y el grupo de control sean semejantes en todas las variables ajenas a la exposición; que difieran en sus factores pronósticos iniciales de desarrollar con el tiempo las reacciones adversas que se quiere estudiar. Si se encuentra una asociación entre la exposición y las reacciones adversas, esa asociación puede no ser causal sino producida por terceras variables, «factores de confusión», asociados con la exposición y con las reacciones adversas, que sean las verdaderas variables causales.

Un caso particular lo plantea el fenómeno de la «confusión por indicación». Cuando un médico prescribe un tratamiento a un paciente y no lo hace a otro lo suele hacer porque el primero está peor y lo necesita más que el que está más sano. Si luego el paciente que recibe el tratamiento presenta un acontecimiento adverso y el que no lo recibe no, es posible que la causa del acontecimiento adverso no sea el medicamento sino la mayor gravedad del paciente que lo recibe. En ese caso, establecer un vínculo causal entre el medicamento y la reacción adversa sería una equivocación, pues la reacción adversa se habría debido a otra característica del paciente, no al medicamento⁴. Este problema de la «confusión por indicación» es muy difícil de abordar, pues exige la comparación de pacientes con la misma indicación pero que reciben tratamientos diferentes, en los que se pueden aplicar métodos estadísticos para ajustar las pequeñas disimilitudes que presenten. Este ajuste supone que se conocen todos los factores determinantes de la indicación, que la información sobre las variables potenciales de confusión es de buena calidad y que el número de pacientes es suficientemente alto para poder aplicar los modelos estadísticos multivariantes⁵. Un estudio reciente puso de manifiesto la existencia de «confusión por indicación» en un estudio de cohortes previo que encontraba una asociación entre dosis altas de epoetina-alfa y mayor mortalidad en pacientes de hemodiálisis prevalentes⁶, pero que esa asociación se fue atenuando hasta desaparecer a medida que se iban considerando sucesivas variables de confusión en el análisis estadístico.

Por último, en ocasiones los estudios de cohortes tampoco pueden analizar complicaciones raras o tardías de los tratamientos, pues exigen mucho tiempo y mucho dinero.

3. Los estudios de casos y controles permiten estudiar complicaciones raras y tardías y no requieren una muestra muy amplia. Se definen porque se toman como casos aquellos pacientes que desarrollan la complicación o la enfermedad que se quiere estudiar y como controles sujetos de un ámbito semejante que no la presentan. A partir de ahí se investiga si los casos y los controles han estado expuestos a diversos factores, entre ellos los tratamientos o las intervenciones médicas que se quieran estudiar. Aunque son estudios más rápidos y baratos que los ensayos controlados

y aleatorizados y que los estudios de cohortes, están más sujetos a sesgos y confusión. En particular, dos problemas amenazan su validez: la selección de los controles y los sesgos de información acerca de la exposición y de las variables de confusión. Los controles no sólo deben estar libres de enfermedad sino también deben ser similares a los casos con respecto a sus posibilidades de haber estado expuestos a las variables de exposición estudiadas; tanto los controles como los casos tendrían que haber tenido el mismo riesgo de exposición si no hubiera ninguna asociación entre la exposición y la enfermedad. Por otra parte, la investigación de la exposición se debe hacer minimizando sesgos de recuerdo de exposiciones pasadas que afecten de modo distinto a los casos que a los controles^{7,8}.

Perneger et al. publicaron un estudio de casos y controles para evaluar la influencia de la toma de paracetamol, aspirina y antiinflamatorios no esteroideos (AINE) en la etiología de la enfermedad renal crónica avanzada que requiere diálisis⁹. Tomaron como casos pacientes recién llegados a diálisis y como controles a sujetos escogidos al azar por teléfono en los mismos barrios de residencia de los casos que no estuvieran en diálisis. Al tratarse de medicamentos sin receta, no había ningún registro de prescripción, por lo que en las entrevistas a los casos y a los controles tuvieron que preguntarles acerca del consumo de medicinas que recordaban a lo largo de su vida. Esto pudo introducir cierto sesgo entre los dos grupos, pero los autores preguntaron acerca de tres tipos de medicamentos para intentar medir el sesgo. Encontraron una diferencia importante en la exposición al paracetamol, de modo que en comparación con las personas que no habían tomado 1.000 pastillas de paracetamol en su vida, las que habían tomado entre 1.000 y 4.999 en total tenían el doble de riesgo de padecer enfermedad renal crónica (razón de posibilidades [*odds ratio*] 2,0), mientras que las que habían tomado 5.000 o más tenían un riesgo mayor (*odds ratio* 2,4). Sin embargo, no encontraron asociación entre la aspirina y la enfermedad renal, lo que servía para minimizar el posible sesgo. Tampoco encontraron asociación entre los AINE y la enfermedad renal, salvo en aquellos casos con consumo masivo (más de 5.000 pastillas), que tenían una *odds ratio* de 8,8 de desarrollar enfermedad crónica avanzada.

4. Por último, las series de casos, sin grupo control, en ocasiones constituyen la única fuente de información válida acerca de los efectos adversos de las intervenciones terapéuticas o diagnósticas. Cuando los efectos son muy frecuentes, las series de casos pueden llegar a ser pruebas suficientemente sólidas para establecer relaciones presuntamente causales. Éste puede ser el caso de la fibrosis sistémica nefrogénica, como complicación de la administración de contraste con gadolinio en pacientes con enfermedad renal crónica^{10,11}. Sin embargo, cuando los efectos son infrecuentes, las series de casos,

como las descripciones de casos, sólo pueden generar hipótesis que habrán de comprobarse en ulteriores estudios.

¿Se midieron de igual modo las exposiciones al tratamiento y los resultados clínicos en ambos grupos?

La información acerca de los distintos tipos de diseño también orientará acerca de este criterio. Mientras que un ensayo controlado y aleatorizado que sea doble ciego la información acerca de la exposición y los resultados será igual en el grupo de tratamiento y en el grupo de control, en el resto de casos puede haber diferencias en la búsqueda de información entre los dos grupos que conduzcan a sesgos.

Un estudio basado en un registro hospitalario que incluyó a pacientes con los diagnósticos de hipertensión, edemas o insuficiencia cardíaca, como marcadores del uso de diuréticos, pretendió evaluar la asociación de la toma de diuréticos con el riesgo de cáncer. El riesgo de cáncer aumentó un 22% y el de hipernefoma se duplicó. Los autores reconocen que hubo un sesgo de verificación diagnóstica, es decir, que en los pacientes que tomaban diuréticos hubo una información más exhaustiva y completa que en el grupo de control. El estudio también adolece de otros problemas, como el riesgo de confusión por variables como el tabaco, que no fue abordado en el análisis¹².

¿Fue suficientemente exhaustivo y prolongado el seguimiento de los pacientes del estudio?

Ésta es una cuestión clave que depende de la naturaleza de la exposición y de la enfermedad. Si se pretende estudiar el efecto de la radiación electromagnética o del uso del teléfono móvil sobre el riesgo de cáncer, los estudios deberán ser suficientemente prolongados para descartar efectos tardíos.

Fayad et al. estudiaron la función renal de 100 niños sometidos a litotricia extracorpórea y emplearon un tiempo de seguimiento de 6 meses. Concluyeron que la litotricia extracorpórea es un procedimiento seguro¹³. Otro estudio, de Ibrahim et al., estudió las consecuencias a largo plazo de la donación de vivo de un riñón para trasplante; el tiempo medio de seguimiento fue de 12,2 años, con una desviación tipo de 9,2 años¹⁴. Corresponde al lector juzgar si se trata de períodos de seguimiento suficientemente prolongados en el tiempo, y si el seguimiento fue suficientemente exhaustivo, pues la prolongación del tiempo de seguimiento aumenta el riesgo de los perdidos en el seguimiento y, por tanto, amenaza la validez del estudio.

Criterios secundarios

¿Satisficieron los resultados algunas «pruebas diagnósticas de causalidad» razonables?

Sir Austin Bradford Hill enumeró una serie de criterios para distinguir asociaciones causales de las no causales en los estudios epidemiológicos observacionales¹⁵. Estos criterios se basan en otros previos que se remontan a los criterios de inferencia causal de Hume publicados en 1739¹⁶. Estas pautas no constituyen una lista de comprobación que permita asegurar la causalidad de una asociación, pues eso depende de cada artículo y de sus características particulares. En el caso de estudios que afronten los posibles efectos nocivos, los criterios anteriormente expuestos son más importantes. No obstante, se enumeran algunos de los criterios de causalidad de Hill más relevantes, por orden de importancia:

¿Se aseguró que la exposición precedió al efecto nocivo?

La causa siempre debe preceder al efecto, sin excepciones. En todos los artículos se debe asegurar que la exposición precede en el tiempo al efecto.

¿Hubo un «gradiente dosis-respuesta», es decir, un aumento de la dosis o de la duración de la exposición se asocia con un aumento de la magnitud del efecto nocivo?

Este criterio es bastante razonable, pero no siempre es cierto. En algunas ocasiones la asociación entre la exposición y la respuesta tiene forma de «J», como se ha descrito en la relación entre el consumo de sal y la hipertensión arterial¹⁷ o entre la presión arterial y la incidencia de ictus en la enfermedad renal crónica¹⁸. También puede ocurrir que cuando hay un «gradiente dosis-respuesta» entren en juego factores de confusión paralelos a la exposición que actúen acentuando la asociación cuando es el factor de confusión el que se relaciona con el desenlace. Por ejemplo, si se estudiara la relación entre la edad del padre y la frecuencia de aparición de síndrome de Down, esa asociación estaría confundida por la edad de la madre. Y en otras ocasiones el efecto de la exposición sobre la respuesta es estable y no depende de la cantidad de dosis, con lo que la ausencia del gradiente no es prueba contraria a la causalidad. Esto sucede, por ejemplo, en la fenilcetonuria, que se manifiesta en lactantes con déficit de la enzima hidrolasa de la fenilalanina, independientemente de la cantidad de fenilalanina ingerida en la dieta cuando ésta sobrepasa una cierta cantidad. Y también puede ocurrir, raramente, que dosis más bajas sean más nocivas que dosis mayores, como sucede con la irradiación tiroidea, que en pequeñas cantidades puede producir bocio nodular y cáncer, pero no a dosis mayores¹⁹.

El resto de criterios, excepto el de la fortaleza de la asociación, que se verá más adelante, tienen menor importancia.

¿Hubo pruebas de que el efecto adverso se redujo o desapareció al retirar el tratamiento y se recrudeció o reapareció cuando se reinstauró el tratamiento?

Ésta es una circunstancia que sólo se prueba en contadas ocasiones, por motivos éticos. Incluso cuando se produce no es prueba inequívoca de causalidad, pues puede acompañarse de factores de confusión asociados no suficientemente controlados.

¿Hubo «consistencia» (concordancia) entre los distintos estudios realizados sobre la asociación entre la exposición y el efecto?

Ésta es una prueba positiva cuando los distintos estudios son de naturaleza diversa. Pero su ausencia tampoco descarta causalidad, pues los efectos de una exposición dependen de muchas circunstancias.

¿Hubo especificidad en la asociación entre la asociación y el efecto?

Ésta es una prueba bastante endeble, pues la mayoría de las causas producen múltiples efectos patológicos diversos. Tómese el caso del tabaco, que produce cáncer en múltiples órganos, además de patología vascular diversa, una causa nada específica.

¿Hubo coherencia en la asociación, es decir, el grado en que una asociación causal que se postula es compatible con la teoría y el conocimiento preexistente²⁰?

¿Hubo verosimilitud biológica en la hipótesis de la asociación causal?

¿Hubo analogía de la asociación causal postulada con otras asociación semejantes ya probadas entre una exposición de la misma clase y el mismo efecto²⁰?

Por ejemplo, como se conocen los efectos de la talidomida y de la rubeola en el embarazo, se podría aceptar una asociación causal si otros medicamentos u otra infección viral produjeran los mismos efectos¹⁵.

Estas tres preguntas son también relativamente endebles, pues dependen de los conocimientos previos, que pueden ser equi-

vocados. Por ejemplo, durante mucho tiempo se consideraba que había suficientes razones biológicas para pensar que la administración de albúmina como solución intravenosa en el shock hipovolémico era superior a la administración de suero salino; sin embargo, los ensayos clínicos que han comparado esos dos procedimientos de rehidratación intravenosa no han podido observar un beneficio de la albúmina en lo que respecta a la mortalidad, mientras que la administración de albúmina aumenta la mortalidad en el caso de grandes quemados²¹. Estos resultados obligaron a modificar la verosimilitud biológica previa. A partir de cualquier hecho es posible construir en seguida una nueva hipótesis biológica verosímil¹⁹.

¿CUÁLES FUERON LOS RESULTADOS?

¿Cuál fue la fuerza de asociación entre la exposición y el efecto?

1. Esta pregunta también tiene que ver con la validez del estudio, pues la fuerza de la asociación es otro criterio de causalidad incluido en los formulados por Hill¹⁵. Una asociación más poderosa, es decir, con una medida de efecto mayor, puede constituir una prueba mayor de la existencia de una asociación causal²². Sin embargo, esa asociación puede estar confundida por factores de confusión suficientemente potentes como para hacer aparecer una asociación muy patente. Por otra parte, una asociación relativamente débil no es prueba de la ausencia de causalidad. El ejemplo más claro es la asociación del tabaco con la cardiopatía isquémica, que tiene una medida de efecto relativamente pequeña. Mientras el riesgo relativo de cáncer de pulmón asociado al tabaco es de casi 8, un riesgo 700% mayor en fumadores que en no fumadores, el riesgo de cardiopatía isquémica es de alrededor de 1,25, un 25% mayor en fumadores que en no fumadores. Pero esa asociación es causal y el tabaco ocasiona una mortalidad y morbilidad bastante mayor por cardiopatía isquémica que por cáncer de pulmón.
2. En los ensayos controlados y aleatorizados y en los estudios de cohortes, la medida de riesgo más habitual es el riesgo relativo o razón de riesgos, que es el cociente entre dos incidencias acumuladas, a su vez cocientes entre el número de casos con las reacciones adversas buscadas y el número de sujetos expuestos o no expuestos, según se hable del grupo de intervención o del grupo control; equivalente a la razón de riesgos es la razón de tasas, que es el cociente entre dos tasas de incidencia, a su vez cocientes entre el número de casos con las reacciones adversas buscadas y el tiempo en riesgo de los sujetos expuestos o no expuestos, expresado en años (o meses, semanas o días) en riesgo por paciente (años-paciente, meses-paciente, semanas-paciente o días-paciente). En ocasiones, sobre todo en los estudios de cohortes que aplican análisis multivarian-

tes para ajustar las variables de confusión, se emplean las razones de riesgos instantáneos (*hazard ratios*) que tienen una interpretación equivalente a las razones de riesgos o las razones de tasas. Una razón de riesgos de 1 indica que no hay una incidencia de reacciones adversas con la exposición estudiada con respecto al grupo no expuesto; mientras que una razón de riesgos superior a 1 indica que la incidencia de acontecimientos adversos en el grupo de exposición es mayor que en el grupo control y cuanto mayor sea la magnitud de la razón de riesgos, mayor es la incidencia relativa de reacciones adversas asociadas a la exposición en comparación con la no exposición. Por el contrario, una razón de riesgos inferior a 1 indica que la incidencia de efectos nocivos de la exposición es inferior a la incidencia sin la exposición y cuando menor de 1 sea la razón de riesgos menor será la incidencia relativa en el grupo expuesto en comparación con el grupo no expuesto.

3. En estudios de casos y controles no se puede conocer el riesgo en los grupos de exposición y no exposición, pues se desconoce el total de la población de donde provienen los casos y los controles. Por tanto, no se puede calcular directamente la razón de riesgos o la razón de tasas. En su lugar, se calcula la razón de posibilidades (*odds ratio*), que es una medida indirecta del riesgo relativo, al que se aproxima tanto más cuando menor es la incidencia de los efectos adversos estudiados. La razón de posibilidades es el cociente entre dos cocientes, dos posibilidades: el cociente entre los casos expuestos y los casos no expuestos, en el grupo de casos; y el cociente entre los controles expuestos y los controles no expuestos en el grupo de controles. Su interpretación es semejante a la del riesgo relativo.
4. Una medida muy útil para evaluar la magnitud del daño que puede producir una intervención es el número necesario para hacer daño (NND, NNH según las siglas inglesas de *number needed to harm*). En muchas ocasiones no se describe en los artículos, pero se puede calcular. Para ello es necesario conocer el riesgo absoluto de reacciones adversas con la exposición, $R_e = \text{casos}_e / N_e$, en donde casos_e es el número de casos con reacciones adversas en el grupo expuesto y N_e es el número total de sujetos expuestos; y sin la exposición, $R_{ne} = \text{casos}_{ne} / N_{ne}$, en donde casos_{ne} es el número de casos con reacciones adversas en el grupo no expuesto y N_{ne} es el número total de sujetos no expuestos; a partir de ello se calcula la diferencia entre esos dos riesgos, la diferencia de riesgos, $\Delta_r = R_e - R_{ne}$; por último, se calcula su inverso, de modo que $\text{NND} = 1/\Delta_r$. El NND indica el número de pacientes que reciben una intervención para producir una reacción adversa determinada; cuanto menor es el NND, más perjudicial es la intervención. El NND tiene una interpretación contraria al NNT, el número necesario para tratar.

De hecho, la mayor parte de las intervenciones producen beneficios y perjuicios, de modo que el lector deberá juzgar cuál es el balance entre ellos antes de tomar una decisión acerca de si se administra la intervención a un paciente. En un ensayo que comparaba la profilaxis con valganciclovir 200 días frente a 100 días después del trasplante en pacientes con alto riesgo, en NNT para evitar un episodio de enfermedad por citomegalovirus fue de 5, mientras que el NND para evitar un episodio de leucopenia después del día 100 fue de 7²³.

¿Qué precisión tuvo la estimación del riesgo?

Como en artículos anteriores de esta serie^{1,2,24,25}, la precisión se evalúa con los intervalos de confianza de las medidas de efecto. Un intervalo de confianza estrecho indica una medida precisa, mientras que un intervalo ancho indica una medida imprecisa, en donde queda bastante incertidumbre acerca de la magnitud de la medida.

¿SERÁN ÚTILES LOS RESULTADOS PARA LA ASISTENCIA DE LOS PACIENTES?

El clínico se debe plantear la siguiente pregunta:

¿Pueden extrapolarse los resultados del estudio a este paciente?

Más que preguntarse si el paciente hubiera reunido los criterios de inclusión del estudio, la pregunta debe formularse buscando si las diferencias con el estudio son tan importantes, en lo que respecta a los tratamientos (dosis, pautas) o las exposiciones, así como los factores asociados que pueden haber influido en sus efectos, como para concluir que sus resultados no tienen nada que ver con él.

¿Cuál es el riesgo de que este paciente sufra el efecto nocivo?

Ésta es la segunda pregunta formulada al comienzo, si el paciente puede sufrir el acontecimiento adverso por determinada intervención. Ya antes se habló del contraste entre el NNT y el NND, que expresa de una forma cuantitativa ese balance, cuando se pueda calcular. Algunos autores postulan el uso de una fracción decimal, *F*, que relaciona el riesgo de un paciente concreto con el riesgo de los pacientes estudiados en el artículo según su comorbilidad (por ejemplo, si el riesgo de su paciente es del doble que el de los del artículo *F* sería 2; si es igual, 1 y si es la mitad, 0,5). El NND se dividiría por *F* para calcular el riesgo de que el paciente sufra ese efecto adverso. En cual-

quier caso, siempre habrá que plantearse la pregunta de cuáles serán las consecuencias adversas de reducir o eliminar la exposición al agente nocivo.

¿Cuáles son las preferencias, preocupaciones y expectativas del paciente con respecto a este tratamiento?

La aversión al riesgo o la disposición a asumir efectos secundarios con objeto de conseguir resultados clínicos favorables son factores que el paciente tiene que juzgar. Se pueden introducir modificaciones en la fracción F para ajustar el balance entre NNT y NND para tener una orientación cuantitativa que le ayude al paciente a participar en la decisión.

¿Cuáles son los tratamientos alternativos disponibles?

Incluso cuando el NND ajustado sea desfavorable con respecto al NNT, se deben identificar los tratamientos alternativos disponibles, o la alternativa de no aplicar ningún tratamiento, que se debe proponer al paciente.

En fin, con esta información evaluada de modo sistematizado el clínico podrá disponer de instrumentos para decidir si el acontecimiento adverso que presenta su paciente se debe al tratamiento recibido, o si lo puede presentar con cierta posibilidad, y decidir si mantener o instaurar el tratamiento potencialmente causante de la reacción adversa. En la tabla 1 se expone un resumen de estos criterios.

Tabla 1. Guía para la lectura crítica de un artículo sobre efectos nocivos

¿SON VÁLIDOS LOS RESULTADOS DEL ESTUDIO?

Criterios primarios:

¿Estuvo la pregunta de investigación claramente definida?

- Justificación del estudio
- Componentes básicos del ensayo:
 - + Pacientes
 - Criterios de selección
 - Ámbito
 - + Intervención
 - + Comparación
 - + Desenlace o resultado
 - Variables de resultado primarias y secundarias
 - Definición de las variables de resultado
- Objetivos específicos e hipótesis del estudio

¿Se utilizaron grupos de comparación claramente definidos que fueron similares en todos los aspectos importantes salvo en la exposición al tratamiento?

- El caso de ensayos controlados con distribución aleatoria
- El caso de los estudios de cohortes
- El caso de los estudios de casos y controles
- El caso de las series de casos y las descripciones de un caso

¿Se midieron de igual modo las exposiciones al tratamiento y los resultados clínicos en ambos grupos?

- Sesgo de vigilancia diagnóstica

¿Fue suficientemente exhaustivo y prolongado el seguimiento de los pacientes del estudio?

Criterios secundarios:

¿Satisficeron los resultados algunas «pruebas diagnósticas de causalidad» razonables?

- ¿Se aseguró que la exposición precedió al efecto nocivo?
- ¿Hubo un «gradiente dosis-respuesta», es decir, un aumento de la dosis o de la duración de la exposición se asocia con un aumento de la magnitud del efecto nocivo?

Continúa en página siguiente>>

Tabla 1 (Continuación). Guía para la lectura crítica de un artículo sobre efectos nocivos

- ¿Hubo pruebas de que el efecto adverso se redujo o desapareció al retirar el tratamiento y se recrudesció o reapareció cuando se reinstauró el tratamiento?
- ¿Hubo «consistencia» (concordancia) entre los distintos estudios realizados sobre la asociación entre la exposición y el efecto?
- ¿Hubo especificidad en la asociación entre la asociación y el efecto?
- ¿Hubo coherencia en la asociación, es decir, el grado en que una asociación causal que se postula es compatible con la teoría y el conocimiento preexistente)?
- ¿Hubo verosimilitud biológica en la hipótesis de la asociación causal?
- ¿Hubo analogía de la asociación causal postulada con otras asociación semejantes ya probadas entre una exposición de la misma clase y el mismo efecto?

¿CUÁLES SON LOS RESULTADOS?

¿Cuál fue la fuerza de asociación entre la exposición y el efecto?

- Fuerza de asociación como criterio de causalidad
- Medidas de efecto: razón de riesgos y razón de tasas
- Medidas de efecto: razón de posibilidades (odds ratios)
- Número necesario para hacer daño

¿Qué precisión tuvo la estimación del riesgo?

- Intervalos de confianza

¿SERÁN ÚTILES LOS RESULTADOS PARA LA ASISTENCIA DE LOS PACIENTES?

¿Pueden extrapolarse los resultados del estudio a este paciente?

¿Cuál es el riesgo de que este paciente sufra el efecto nocivo?

¿Cuáles son las preferencias, preocupaciones y expectativas del paciente con respecto a este tratamiento?

¿Cuáles son los tratamientos alternativos disponibles?

REFERENCIAS BIBLIOGRÁFICAS

1. García López F, Quereda Rodríguez-Navarro C. Apuntes para la lectura crítica de un ensayo clínico. *Nefrología* 2007;27(Suppl 1):60-6.
2. García López F, Quereda C. Apuntes para la lectura crítica de una revisión sistemática. *Nefrología* 2007;27:417-24.
3. Pfeffer MA, Burdmann EA, Chen CY, Cooper ME, De Zeeuw D, Eckardt KU, et al. A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N Engl J Med* 2009;361:2019-32.
4. Walker AM, Stampfer MJ. Observational studies of drug safety. *Lancet* 1996;348:489.
5. Grobbee DE, Hoes AW. Confounding and indication for treatment in evaluation of drug treatment for hypertension. *BMJ* 1997;315:1151-4.
6. Bradbury BD, Wang O, Critchlow CW, Rothman KJ, Heagerty P, Keen M, et al. Exploring relative mortality and epoetin alfa dose among hemodialysis patients. *Am J Kidney Dis* 2008;51:62-70.
7. Schlesselman JJ. Case-control studies: design, conduct, analysis. New York: Oxford University Press, 1982.
8. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. *Am J Epidemiol* 1992;135:1019-28.
9. Perneger TV, Whelton PK, Klag MJ. Risk of kidney failure associated with the use of acetaminophen, aspirin, and nonsteroidal antiinflammatory drugs. *N Engl J Med* 1994;331:1675-9.
10. Swartz RD, Crofford LJ, Phan SH, Ike RW, Su LD. Nephrogenic fibrosing dermatopathy: a novel cutaneous fibrosing disorder in patients with renal failure. *Am J Med* 2003;114:563-72.
11. Shabana WM, Cohan RH, Ellis JH, Hussain HK, Francis IR, Su LD, et al. Nephrogenic systemic fibrosis: a report of 29 cases. *AJR Am J Roentgenol* 2008;190:736-41.
12. Mellemaard A, Moller H, Olsen JH. Diuretics may increase risk of renal cell carcinoma. *Cancer Causes Control* 1992;3:309-12.
13. Fayad A, El-Sheikh MG, Abdelmohsen M, Abdelraouf H. Evaluation of renal function in children undergoing extracorporeal shock wave lithotripsy. *J Urol* 2010;184:1111-4.
14. Ibrahim HN, Foley R, Tan L, Rogers T, Bailey RF, Guo H, et al. Long-term consequences of kidney donation. *N Engl J Med* 2009;360:459-69.
15. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.
16. Rothman KJ, Greenland S. Hill's criteria for causality. In: Gail MH, Benichou J, editors. *Encyclopedia of epidemiologic methods*. Chichester, England: John Wiley & Sons, Ltd., 2000;428-31.

17. Chien KL, Hsu HC, Chen PC, Su TC, Chang WT, Chen MF, et al. Urinary sodium and potassium excretion and risk of hypertension in Chinese: report from a community-based cohort study in Taiwan. *J Hypertens* 2008;26:1750-6.
18. Weiner DE, Tighiouart H, Levey AS, Elsayed E, Griffith JL, Salem DN, et al. Lowest systolic blood pressure is associated with stroke in stages 3 to 4 chronic kidney disease. *J Am Soc Nephrol* 2007;18:960-6.
19. Weiss NS. Inferring causal relationships: elaboration of the criterion of «dose-response». *Am J Epidemiol* 1981;113:487-90.
20. Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am J Epidemiol* 1991;133:635-48.
21. Alderson P, Bunn F, Lefebvre C, Li WP, Li L, Roberts I, et al. Human albumin solution for resuscitation and volume expansion in critically ill patients. *Cochrane Database Syst Rev* 2004;CD001208.
22. Taubes G. Epidemiology faces its limits. *Science* 1995;269:164-9.
23. Humar A, Lebranchu Y, Vincenti F, Blumberg EA, Punch JD, Limaye AP, et al. The efficacy and safety of 200 days valganciclovir cytomegalovirus prophylaxis in high-risk kidney transplant recipients. *Am J Transplant* 2010;10:1228-37.
24. Barrio V, García López F. Apuntes para la lectura crítica de un estudio sobre pronóstico. *Nefrología* 2007;27:537-41.
25. Zamora J, Abraira V. Analisis de la calidad de los estudios de evaluación de pruebas diagnósticas. *Nefrología* 2008;28 (Suppl 2):42-5.

Referencias básicas

- Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence Based Medicine Working Group. *JAMA* 1994;271:1615-9.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Medicina basada en la evidencia: cómo ejercer y enseñar la MBE*. Madrid: Churchill Livingstone, 1997.

Referencias complementarias en la red (todas ellas consultadas el 16 de octubre de 2010):

- CASPe (Critical Appraisal Skills Program en España): programa de habilidades en lectura crítica, España. <http://www.redcaspe.org/herramientas/index.htm>.
- Centre for Evidence-Based Medicine, Oxford-Centre for Evidence Based Medicine. <http://www.cebm.net/>
- Enseñar MBE: herramientas y calculadoras interactivas. <http://www.infodoctor.org/rafabravo/herramientas.htm>