



II

LA OBSERVACIÓN CLÍNICA Y EL RAZONAMIENTO CAUSAL

Un profesional sanitario puede hacerse muchos tipos de preguntas sobre relaciones de causalidad. Puede estar interesado en conocer los determinantes de una enfermedad que le permitan hacerse una idea del pronóstico. También puede interesarle conocer si para el tratamiento de un problema determinado es mejor un tratamiento *A* o uno *B*, puede que desee identificar variables que permitan predecir la eficacia de cada uno de los tratamientos disponibles. Análogamente, puede preguntarse si la mejoría o el empeoramiento de un cuadro clínico pueden atribuirse al tratamiento administrado.

De manera consciente o inconsciente, en medicina la toma decisiones reposa sobre un razonamiento inductivo: a partir de las observaciones clínicas se establecen conjeturas que van siendo descartadas o confirmadas con sucesivas observaciones ordenadas de manera más o menos sistemática. Por el contrario, la lógica de la investigación usa sobre todo razonamientos deductivos. Cuando un paciente que se encuentra en un estado *A* progresa a un estado diferente (estado *B*) tras tomar un medicamento, interesa saber si este cambio se debe a la intervención objeto de estudio. El estado *A* podría ser una enfermedad y el estado *B*—el desenlace— podría ser la curación o simplemente un alivio sintomático. (En ocasiones las opciones son más complejas y las elecciones más difíciles; por ej., el estado *A* también podría ser la presencia de una enfermedad y el estado *B* su ausencia, pero a cambio de una reacción adversa grave).

En un ensayo clínico terapéutico se efectúa una comparación, y la cuestión clave consiste en determinar si hay diferencias entre las intervenciones comparadas para hacer que los pacientes en un estado de salud *A* pasen (o no pasen) a un nuevo estado *B* más favorable. El hecho de que la intervención preceda en el tiempo al estado *B* no implica necesariamente que la intervención sea la causa de *B*.^{1,2} Si se sacara esta conclusión se cometería lo que se denomina una falacia *post hoc ergo propter hoc* (véase el Glosario).

El estudio de las relaciones de causalidad en medicina ha estado muy influido por los postulados de Koch (véase *Postulados de Koch* en el Glosario). Estos postulados se basaban en la creencia de que para enunciar una afirmación inequívoca de causalidad, era necesario que la causa propuesta fuera a la vez necesaria y suficiente (relación unicausal y unívoca). Sin embargo, la expansión de los conocimientos médicos obliga a considerar más bien la noción de causa contributiva. En general, para considerar que una intervención terapéutica es contributiva (ha contribuido a modificar el estado de salud inicial *A*), aquélla debe haber precedido al efecto, y en algunos casos (si no son de esperar efectos irreversibles) la modificación o eliminación de la intervención debe modificar consecuentemente la magnitud o incluso la naturaleza del efecto. Sin embargo, debido a la variabilidad interindividual comentada en el Capítulo I, a menudo no es posible aislar una sola variable y estudiarla de manera individualizada, de manera que en biología la causa contributiva es identificada como un valor promedio, que no es necesariamente observable en todos los individuos y a veces ni tan solo en la mayoría de los individuos. Así, no todos los pacientes sometidos a una causa contributiva manifestarán el efecto esperado, o, dicho de otro modo, no todos los pacientes tratados con un fármaco de "eficacia demostrada" responderán favorablemente a su administración.

La variabilidad inter e intraindividual obliga a plantear la evaluación de los efectos de una intervención médica (por ej., de un medicamento), en una perspectiva de grupo y no individual. (También hay métodos para la evaluación rigurosa de la respuesta de un paciente concreto a diferentes modalidades de tratamiento; véanse el Capítulo VII y *Ensayo clínico de $n = 1$* en el Glosario).

Esta evaluación debe comprender por lo menos dos aspectos: la identificación del efecto —que se basa en el binomio experiencia clínica + rigor metodológico— y su cuantificación, bien en términos de medición de la intensidad de la respuesta en uno o en varios pacientes, o bien en términos de establecimiento de un porcentaje de pacientes que responden de una determinada manera (véanse, respectivamente, *Variable continua* y *Variable categórica* en el Glosario).

En la medicina contemporánea se utilizan diversas estrategias de investigación para proponer, investigar y comprobar relaciones de causa a efecto. En ciertas ocasiones poco comunes, una observación en un solo paciente puede sugerir la posibilidad de una nueva indicación de un fármaco, o de un efecto indeseado hasta entonces desconocido. Si se registran observaciones similares en varios pa-

cientes, la "fuerza" de la posible relación de causalidad es mayor. En los próximos apartados se revisan los elementos a tener en cuenta cuando se evalúa la posibilidad de que una asociación de dos factores se deba a una relación causal.

EVALUACION DE LOS RESULTADOS DE ESTUDIOS DE CASOS Y SERIES DE CASOS

La publicación —por ej., en forma de Carta al Director de una revista médica— de observaciones clínicas anecdóticas sirve para generar hipótesis de relaciones causales, por ejemplo entre la administración de un medicamento y la aparición de un posible efecto adverso. Los siguientes criterios contribuyen a que una observación anecdótica de asociación entre la exposición a un factor y la aparición de un nuevo estado clínico pueda considerarse sospechosa de ser una verdadera relación de causalidad:¹

- Que se asegure una secuencia temporal adecuada, es decir, que el fármaco haya sido realmente administrado antes de que comenzara la enfermedad, y no sólo antes de que se manifestaran sus síntomas o de que fuera diagnosticada.
- Que los dos hechos asociados en la observación anecdótica sean excepcionales. Sería difícil creer que un analgésico de uso muy común como el paracetamol puede producir un síntoma frecuente como fatiga, sólo sobre la base de notificaciones anecdóticas sin grupo de comparación, porque uso de paracetamol y fatiga son fenómenos frecuentes, que por lo tanto pueden coincidir casualmente (por otra parte, el paracetamol puede haber sido tomado para el tratamiento de un cuadro asociado a fatiga).
- Que, tras cesar la exposición al factor sospechoso, el nuevo estado clínico desaparezca (en caso de que sea de naturaleza reversible).
- Que los conocimientos biológicos sobre el factor sospechoso y el estado clínico permitan suponer un determinado mecanismo de acción (plausibilidad biológica).
- Que, tras un detenido examen, se puedan descartar otras causas (como enfermedades asociadas, exposición a otros factores, etc.).
- Que la misma asociación haya sido apreciada por otros autores, en circunstancias diferentes.

Estos criterios no aseguran en modo alguno que la observación de una asociación sea realmente de causalidad, pero por lo menos ayudan a descartar la posibilidad de coincidencia entre la exposición a un factor y un acontecimiento clínico determinado.

LIMITACIONES DE LOS ESTUDIOS NO CONTROLADOS

El valor preventivo o terapéutico general de una determinada intervención médica sólo se puede juzgar si se la compara con intervenciones alternativas, o bien con ninguna intervención. Por naturaleza, un estudio clínico controlado implica comparación entre grupos de pacientes. Se dice que un estudio es no controlado si describe el curso de una enfermedad en un solo grupo de pacientes que han sido expuestos a la intervención de interés. Los estudios en los que se compara el estado clínico de un grupo de pacientes antes y después de la aplicación de una determinada intervención no permiten sacar conclusiones sobre relaciones de causalidad, por varios motivos:²

- En primer lugar, porque la mayoría de las enfermedades tienen un curso imprevisible. Si el curso clínico de la enfermedad es previsible —lo que ocurre raramente—, el examen de un grupo de comparación es menos necesario. Así por ejemplo, la endocarditis bacteriana subaguda no tratada es catastrófica. No obstante, en la mayoría de las situaciones, el médico toma sus decisiones terapéuticas frente a enfermedades de curso menos previsible, y en estas circunstancias (que son las más habituales) no sería fiable evaluar los efectos de un posible tratamiento sólo a través de la observación del curso clínico de uno o varios pacientes. Si el curso espontáneo de una enfermedad es habitualmente autolimitado (por ej., faringoamigdalitis vírica o bacteriana, gastroenteritis), la aplicación de medidas con finalidad terapéutica puede coincidir con su mejoría, pero no por ello se puede suponer que constituyan la causa de la mejoría. Además, muchas enfermedades graves no autolimitadas pueden cursar a brotes, con remisiones espontáneas.
- En segundo lugar, porque las personas tienen tendencia a modificar su comportamiento por ser objeto de interés y eventualmente a responder a la atención que se les da, independientemente de la naturaleza de esta intervención (véase *Efecto Hawthorne* en el Glosario).
- En tercer lugar, porque en general los pacientes que presentan valores extremos de una distribución (por ej., están enfermos), tenderán, por término medio, a presentar valores menos extremos en mediciones subsiguientes. Este fenómeno se denomina regresión a la media (véase *Regresión a la media* en el Glosario). Dicho de otro modo, en un momento determinado el número de episodios o la intensidad de las molestias producidas por una enfermedad es elevado (este momento suele ser el de la primera visita al médico); sin embargo, en algún momento ulterior, el número de episodios o la intensidad de las molestias habrán disminuido y serán más próximos al valor promedio de los demás pacientes que padecen esta misma enfermedad.¹ Así por ejemplo, aunque la historia natural de la angina vasoespástica sea variable, el número de episodios disminuye con el tiempo, de modo que dos años después del diagnóstico

inicial la gran mayoría de los pacientes no presentan episodios de angina.³ Algo parecido sucede con la artritis reumatoide: en hospitales universitarios, en donde se visitan los casos más graves de esta enfermedad, se registran tasas de remisión espontánea de un 25 a 50%. [En algunos libros de texto se recomienda probar primero un antiinflamatorio no esteroide (AINE) durante unas semanas; si esto no se sigue de una mejoría, se recomienda cambiar a un segundo AINE durante unas semanas, y eventualmente a un tercero en caso de que no haya respuesta al segundo; si el paciente responde al tercer fármaco, se dice que éste es el fármaco idóneo para este paciente en particular, cuando en realidad lo más probable es que haya habido tiempo suficiente para que la enfermedad remita espontáneamente; véanse *Falacia post hoc, ergo propter hoc* y *Regresión a la media* en el Glosario]. También es típica la regresión a la media en caso de dolor agudo, como el cólico renal: en tres ensayos clínicos en los que se comparaban diuréticos tiacídicos con placebo para saber si éstos disminuían el número de cólicos, se registró una notable disminución de la formación de cálculos con el tiempo, tanto en el grupo tratado con el tiacídico como en el grupo tratado con placebo.⁴⁻⁶

TIPOS DE ESTUDIOS CONTROLADOS

Los estudios epidemiológicos con grupo de control o referencia pueden ser de dos tipos: experimentales u observacionales. En un **estudio experimental** (*ensayo clínico controlado* o bien *ensayo comunitario*; véanse estos términos en el Glosario), los sujetos participantes son seleccionados a partir de una población y distribuidos al azar en tantos grupos como el protocolo prevea; generalmente se forman dos grupos: el tratado con el fármaco estudiado y otro, de control, que recibe el tratamiento habitual o, si no lo hay, un placebo. La asignación al azar propia del ensayo clínico permite que, si el número de participantes es suficiente (y si no hay mala suerte), las variables pronósticas relevantes (como por ej., edad, grado de evolución de la enfermedad estudiada, presencia de otras patologías, uso concomitante de otros fármacos, etc.) se distribuyan en la misma proporción en los dos o más grupos que van a ser comparados (véanse *Asignación aleatoria* y *Asignación aleatoria por bloques* en el Glosario). Nótese que, si la asignación es aleatoria, las variables pronósticas conocidas, pero también las desconocidas, se deben distribuir de manera similar entre los grupos. Así, cualquier diferencia que se registre podrá ser atribuida al diferente tratamiento recibido por uno y otro grupo. Por eso se dice que el ensayo clínico es el patrón oro de los métodos epidemiológicos usados para establecer relaciones de causalidad.

No siempre es posible hacer un estudio experimental, por razones organizativas o de índole ética, y en ocasiones el investigador tiene que limitarse a «observar» la realidad. La diferencia fundamental entre un estudio

experimental y un **estudio observacional** es que en el primero el investigador determina (mediante la asignación aleatoria) la distribución de los pacientes entre los grupos, y en el segundo no. En consecuencia, en un estudio observacional no se puede asegurar que otros factores, distintos del tratamiento o variable en estudio, no difieran entre uno y otro grupo. Cuando el investigador no puede planificar una asignación aleatoria a las diferentes intervenciones estudiadas, tiene dos alternativas: el estudio de cohortes o bien el estudio de casos y controles (véase la figura 2).²

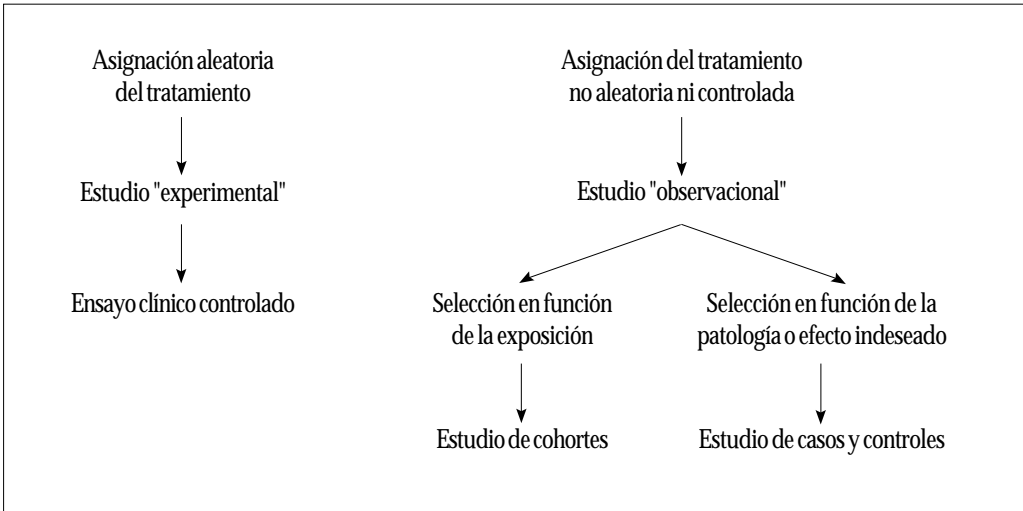


Figura 2. Tipos de diseños epidemiológicos analíticos aplicados en la evaluación de relaciones de causalidad.

Estudios de cohortes

En un estudio de cohortes el investigador identifica una población de individuos expuestos al fármaco de interés. Simultáneamente, debe identificar otra población de individuos no expuestos. Cada una de estas poblaciones, que difieren entre sí por su exposición, se denomina cohorte. A continuación el investigador supervisa el estado de salud a intervalos regulares durante un determinado período de tiempo y mide y compara la frecuencia de las enfermedades de interés en uno y otro grupo (véanse *Cohorte* y *Estudio de cohortes* en el Glosario). Cualquier diferencia de incidencia de una enfermedad entre uno y otro grupo podría ser debida a la diferente exposición que precisamente diferencia a una cohorte de la otra.^{7,8} Por lo tanto, en los estudios de cohortes se evalúa un fármaco o grupo farmacológico, en relación con el riesgo de tantas enfermedades o síntomas como se desee estudiar. [Por el contrario, en un estudio de casos y controles se evalúa una enfermedad, en relación con tantos

factores de riesgo como se desee estudiar (véase más adelante)]. Un estudio de cohortes tiene la ventaja de que los grupos comparados son supervisados en el tiempo tras la exposición al factor estudiado, es decir que sigue la secuencia natural de los hechos; además, permite conocer la incidencia de las enfermedades de interés en los grupos estudiados. Los estudios de cohortes tienen la desventaja de que son muy costosos, en términos de personal y de tiempo, cuando se desea estudiar el riesgo de una enfermedad con una tasa de incidencia baja,⁹ y cuando transcurre un período largo entre la exposición al factor de riesgo y la aparición de la enfermedad. Sin embargo, pueden ser útiles para evaluar la efectividad de una determinada medida terapéutica (véanse *Eficacia* y *Efectividad* en el Glosario). Los datos obtenidos en un estudio de cohortes se pueden disponer en una tabla de 2×2 (véase la tabla 1), en la que las filas se forman en función de la exposición y las columnas en función de la aparición o no de la enfermedad en cuestión. Obsérvese que en un estudio de cohortes (y también en un ensayo clínico) la "vía de entrada" en la tabla de 2×2 es desde la izquierda. Los cuatro valores de la tabla de 2×2 (tabla 1) permiten calcular el riesgo relativo (véase *Riesgo relativo* en el Glosario).

Tabla 1. Un ejemplo simple y común de tabla de contingencia: la tabla de 2×2 .

	Presencia de la enfermedad		
	Sí	No	Total
Individuos expuestos	a	b	a + b
Individuos no expuestos	c	d	c + d
Totales	a + c	b+d	a+b+c+d

En un estudio de cohortes (o en un ensayo clínico), la incidencia (I_e) entre los individuos expuestos es $I_e = a/a+b$, la incidencia entre los no expuestos (I_p) es $I_p = c/c+d$ y por lo tanto el riesgo relativo es $RR = I_e/I_p = [a/(a+b)]/[c/(c+d)]$. Dado que para enfermedades raras $a+b \approx b$ y $c+d \approx d$, $RR \approx (a \times d)/(b \times c)$ (véanse *Incidencia*, *Riesgo relativo* y *Tasa de incidencia* en el Glosario).

En todo estudio epidemiológico —incluso en el ensayo clínico— se pueden cometer errores sistemáticos o sesgos (véase el Glosario). Los principales sesgos a los que hay que prestar atención en un estudio de cohortes son el sesgo de selección, el sesgo de migración, el sesgo de observación y el sesgo por presencia de factores de confusión.⁸

El **sesgo de selección** es el error sistemático que determina que las cohortes no sean comparables en todas las variables pronósticas y se debe a la naturaleza observacional del estudio: si el investigador tiene que contentarse con observar la realidad, sin determinar (idealmente al azar) la asignación a los tratamientos o exposiciones comparados, es posible que las personas incluidas en el estudio tiendan a recibir una u otra modalidad de tratamiento por algún motivo. Esto impide que se pueda considerar, como se hace en el ensayo clínico, que ambos grupos son comparables en todas las variables pronósticas excepto en la estudiada. Así por ejemplo, en un estudio de cohortes cuyo objetivo era evaluar la seguridad de diferentes formas de contracepción, se observó que, en comparación con las usuarias de diafragma o las de DIU, las usuarias de contraceptivos hormonales eran más jóvenes, incluían una proporción mayor de nulíparas y de fumadoras y presentaban menos antecedentes de enfermedad mamaria benigna, toxemia preeclámpsica, venas varicosas, diabetes y enfermedades de tiroides (véase la tabla 2).⁹ En este ejemplo el sesgo de selección determinaría una infraestimación del riesgo relativo de enfermedad mamaria benigna, venas varicosas, diabetes o enfermedades del tiroides entre las usuarias de contraceptivos, pero también determinaría una sobreestimación del riesgo de cualquier enfermedad que se asociara simultáneamente al hábito de fumar y al uso de contraceptivos hormonales (por ej., enfermedades tromboembólicas).

El **sesgo de migración** es el que se debe al abandono, por algunos sujetos del estudio, de la cohorte de la que formaban parte inicialmente. Así por ejemplo, en el estudio antes referido sobre seguridad de métodos contraceptivos, cinco años después de su comienzo sólo un 33% de las usuarias de contraceptivos hormonales, un 15% de las de diafragma y un 17% de las de DIU seguían usando el método inicial.⁹ El sesgo de migración limita el poder estadístico del estudio (porque se pierden efectivos) y, si no se toma en consideración, tiende a infraestimar el valor del riesgo relativo. También puede alterar la magnitud de la asociación (e incluso su dirección), debido a que los individuos que migran de una cohorte a otra pueden tener características diferentes de los que no migran.

El **sesgo de observación** es el derivado de que cuando un médico efectúa observaciones clínicas, las hace con conocimiento del tratamiento que cada sujeto recibe. Eso determina que, de manera consciente o inconsciente, el observador tienda a buscar los diferentes acontecimientos clínicos con intensidad diferente según el grupo o cohorte al que pertenezca el sujeto (véanse *Ensayo clínico a ciego simple* y *Ensayo clínico a doble ciego* en el Glosario, y el apartado *Enmascaramiento* en el Capítulo III). Así por ejemplo, en un estudio de cohortes en el que se comparen contraceptivos hormonales con otro método contraceptivo, el médico que actúa como observador puede tender a registrar una mayor incidencia de venas varicosas entre las usuarias de contraceptivos hormonales, en comparación con las mujeres de los demás grupos, por temor a que los contraceptivos pudieran incrementar especialmente el riesgo de trombosis venosa en las portadoras de varices.

Tabla 2 Estudio sobre contracepción de la *Oxford Family Planning Association*: algunas características de las tres cohortes (Tomada de Vessey⁹).

	Cohorte (método contraceptivo inicial)		
	Contraceptivo		
	oral (n= 9.653)	Diafragma (n= 4.217)	DIU (n=3.162)
Antecedentes			
Enfermedad mamaria benigna	3,1%	4,4%	4,2%
Toxemia pre eclámpsica	12,6%	16,3%	16,1%
Venas varicosas ^a	1,6%	2,3%	3,9%
Diabetes	0,1%	0,2%	0,3%
Enfermedad del tiroides ^a	0,7%	0,9%	1,1%
Características sociodemográficas			
Edad ≥30 años	44%	65%	65%
Nulípara	29%	15%	<1%
Fuma ≥15 cigarrillos al día	17%	7%	12%

^aDe gravedad suficiente para necesitar hospitalización.

El **sesgo por presencia de factores de confusión** aparece cuando una variable (llamada variable de confusión o factor de confusión) se encuentra en relación simultánea, pero independiente, con la exposición y con la enfermedad estudiada (véanse *Factor de confusión* y los diferentes tipos de sesgo en el Glosario).

En conclusión, la diferencia esencial entre un estudio de cohortes y un ensayo clínico es que en el primero no hay asignación aleatoria a las diferentes modalidades de tratamiento. Esta característica determina que los grupos formados no sean iguales o comparables en todas las variables pronósticas excep-

to en el tratamiento que reciben. Esto invalida considerablemente los diseños de cohortes para evaluar la eficacia de los tratamientos farmacológicos. No obstante, los resultados de estudios de cohortes con criterios amplios de inclusión pueden dar información muy útil sobre la historia natural de los problemas y sobre la efectividad de las intervenciones en la práctica habitual (véanse el Capítulo V y *Eficacia y Efectividad* en el Glosario).

Estudios de casos y controles

Una segunda alternativa de estudio epidemiológico observacional controlado es el estudio de casos y controles. En un estudio de casos y controles se seleccionan sujetos con la enfermedad de interés y a continuación se averiguan sus exposiciones previas a los factores de riesgo de interés; simultáneamente, se identifica una muestra de individuos sin la enfermedad objeto de estudio y se averiguan también entre ellos las exposiciones previas a las variables de interés.^{10,11} Nótese que los sujetos con la enfermedad en estudio ($a + c$ en la tabla 1) se dividen entre los expuestos (a) y los no expuestos (c) y que los sujetos sin la enfermedad objeto de estudio ($b + d$) también se dividen en dos subgrupos de expuestos (b) y no expuestos (d), de modo que en un estudio de casos y controles la "vía de entrada" a la tabla de 2×2 es desde arriba. En un estudio de casos y controles no se pueden calcular incidencias, y por lo tanto un riesgo relativo, pero sí un parámetro sustancialmente equivalente, la razón de ventajas (denominada *odds ratio* en inglés), que es el producto cruzado de los cuatro elementos de la tabla de 2×2 . Si en un estudio de cohortes la fuerza de la asociación se mide con el riesgo relativo, que es $RR = I_e/I_c = [a/(a+b)]/[c/(c+d)]$, y para enfermedades raras es $RR = (a \times d)/(b \times c)$, en uno de casos y controles la razón de ventajas es el producto cruzado de los cuatro elementos de la tabla de 2×2 , es decir $OR = (a \times d)/(b \times c)$ [a menudo los resultados de ensayos clínicos también se expresan con el valor de la razón de ventajas] (véanse la tabla 1, y *Estudio de casos y controles* y *Razón de ventajas* en el Glosario).

Los estudios de casos y controles se emplean sobre todo para explorar la etiología de acontecimientos raros. Supóngase que se desea saber si el uso de contraceptivos hormonales incrementa el riesgo de espina bífida en caso de embarazo (no deseado). Dado que la incidencia de la espina bífida es de alrededor de 1,5 casos por cada 1.000 nacidos vivos, es fácil imaginar que con un estudio de cohortes será necesario estudiar un número muy elevado de embarazos para obtener cifras algo "estables" en la tabla de 2×2 . Así por ejemplo, el estudio de 30.000 embarazos identificaría un total de unos 45 casos de espina bífida ($a+c=45$, véase la tabla 1). El número necesario de participantes sería tanto más elevado cuanto menor fuera la prevalencia de la exposición al factor de riesgo estudiado. Así, si la prevalencia de uso de los contraceptivos hormonales es de un 5% entre las mujeres en edad fértil, el valor a en la tabla de 2×2

para un riesgo relativo de 3 sería de 6, lo que conferirá escaso poder estadístico al estudio. En un estudio de casos y controles en el que se incluyeran unos 350 casos de espina bífida ($a+c=350$) y un número similar de controles ($b+d=350$), y suponiendo que el valor de la razón de ventajas sea de 3, el valor de a sería aproximadamente de 51 (17×3) y el valor de c de 299; entre las controles se identificarían 17 expuestas (b) y 333 no expuestas (d). Nótese que en este ejemplo las cifras de la tabla de 2×2 obtenidas con un estudio de casos y controles son estadísticamente más "estables", con lo que la estimación del riesgo es más precisa que con un estudio de cohortes, con la enorme ventaja añadida de que con la estrategia de cohortes se habrían tenido que estudiar 30.000 embarazos y con la de casos y controles sólo 700.

Los tres principales tipos de sesgo en un estudio de casos y controles son el sesgo de selección, el sesgo de información y el sesgo por presencia de factores de confusión.¹² El **sesgo de selección** se produce cuando los casos y los controles son seleccionados de manera sistemáticamente diferente en relación con la posibilidad de que hayan estado expuestos al factor de riesgo de interés. Puede haber sesgo de selección de los casos o de los controles. Cuando no hay garantía de que se incluye a todos los casos que ocurren en una población y período determinados, existe la posibilidad de que los casos incluidos en el estudio no sean representativos (en términos de factores de riesgo, gravedad de la enfermedad o exposiciones previas) de la totalidad de los casos. También pueden obtenerse resultados diferentes por incluir los casos prevalentes (nuevos y "viejos" casos de la enfermedad), en lugar de limitar el estudio a los casos incidentes (es decir, los nuevos casos de la enfermedad de interés) El método de selección de los controles también puede introducir sesgos. Los controles deben ser idealmente personas que hubieran tenido la misma probabilidad que los casos de exponerse al factor de riesgo de interés.

El **sesgo de información** se origina cuando la información sobre las exposiciones previas se obtiene de manera sistemáticamente diferente en los casos, en comparación con los controles. Por ejemplo, los pacientes con enfermedades en las que se sospecha la implicación de un medicamento (por ej., agranulocitosis, síndrome de Stevens-Johnson) pueden haber sido repetidamente interrogados sobre uso de fármacos, por lo que, en comparación con los controles, su memoria está más "sensibilizada". El sesgo de información puede ser debido pues a diferente sensibilización de la memoria; también puede ser debido a que la información sobre exposiciones previas a los factores de riesgo de interés y sobre otras circunstancias haya sido obtenida con interrogatorios de diferente intensidad entre casos y controles. El uso de cuestionarios estructurados limita los sesgos de información (véase *Cuestionario estructurado* en el Glosario).

El **sesgo por presencia de factores de confusión** se describe en el apartado anterior sobre estudios de cohortes.

Los estudios de casos y controles se usan a menudo en farmacovigilancia porque, afortunadamente, la mayoría de los efectos indeseados graves produci-

dos por medicamentos tienen una tasa de incidencia muy baja. No obstante, no son adecuados para evaluar la efectividad de los tratamientos farmacológicos, porque están sujetos a un inevitable sesgo de selección.

Estudios caso-población

Una limitación frecuente de los estudios de casos y controles es que cuando la prevalencia de la exposición al factor de riesgo de interés es baja, su poder estadístico es muy débil. [Esta baja prevalencia se traduce en un valor bajo, y por lo tanto estadísticamente inestable (es decir, que da lugar a menor precisión del estimador de la casilla c en la tabla de 2×2).

En farmacovigilancia, uno de los campos de interés es la identificación de los posibles efectos adversos raros de los medicamentos de más reciente introducción en terapéutica. Esto se debe a que, en general, estos son los menos conocidos, porque la experiencia con su utilización clínica es escasa. Antes de su comercialización, estos fármacos han sido probados en ensayos clínicos y otros estudios que han incluido a un número limitado de pacientes (del orden de unos miles). Con este número relativamente bajo de personas expuestas, la probabilidad de detectar un efecto adverso raro (con una incidencia inferior a $1/10.000$) es muy baja.

La farmacovigilancia se fundamenta generalmente en las señales generadas por los sistemas de notificación espontánea de sospechas de efectos adversos de los medicamentos. Estos sistemas son útiles para detectar efectos adversos anteriormente no descritos, pero sus características (infranotificación, notificación selectiva) suelen conferirles limitaciones para cuantificar el riesgo (expresado, por ej., en términos de número de acontecimientos por millón de pacientes tratados). Cuando el medicamento es sospechoso de producir un efecto indeseado potencialmente grave y es innovador, interesa tener una estimación de la probabilidad de aparición de este efecto.

Desde un punto de vista epidemiológico, este problema suele caracterizarse por una variable de interés (la enfermedad) que es de baja incidencia, lo que tiende a aconsejar una estrategia de casos y controles para cuantificar el riesgo. Pero también suele caracterizarse por una baja prevalencia de la exposición al factor de riesgo de interés (el nuevo medicamento, que por ser nuevo suele tener una baja prevalencia de uso). En estas situaciones se ha propuesto otra estrategia epidemiológica, el estudio caso-población.

En un estudio caso-población se identifica a todos los casos de la enfermedad de interés en una población predeterminada. Estos pacientes son incluidos en el estudio (con su consentimiento), independientemente de sus exposiciones previas. Se investigan con detalle, mediante entrevista directa con el paciente (o con un familiar si aquel ha fallecido o no puede responder)

y examen de la historia clínica, sus exposiciones previas a fármacos. Al mismo tiempo, se determina la prevalencia de uso del fármaco de interés en la población general o en un estrato de la población general de características demográficas equivalentes a las de los pacientes; este cálculo se basa en los datos de consumo de fármaco de interés. A continuación se compara la incidencia de la enfermedad entre los usuarios, con la incidencia en la población no expuesta. Alternativamente, la información obtenida de este modo puede ser tratada como en un estudio de casos y controles, y se puede comparar la prevalencia de exposición entre los casos con la prevalencia entre los controles. Una mayor prevalencia de uso del fármaco entre los casos sugerirá un incremento del riesgo; una menor prevalencia de uso sugerirá un efecto protector.¹³ Comparado con los estudios de casos y controles, en el caso-población es más difícil y en ocasiones imposible ajustar por sesgos y factores de confusión; no obstante, esta estrategia (y sus variaciones) tiene la ventaja de que permite establecer el orden de magnitud del problema en menos tiempo.

Estudios transversales

En un estudio transversal se registran simultáneamente diversas características de una población en un momento determinado. A continuación se examina si existe alguna asociación entre una y otra característica (véase *Estudio transversal* en el Glosario). Así por ejemplo, en una población determinada se pueden recoger datos sobre la relación peso/altura, cifras de presión arterial, colesterol plasmático, hábito de fumar, antecedentes de cardiopatía isquémica y otros factores de riesgo cardiovascular. Se puede determinar si un factor se asocia con otro, pero dado que no se conoce la secuencia temporal de los acontecimientos, sólo se pueden establecer relaciones de causa a efecto con variables que no se modifican con el tiempo (por ej., sexo, grupo sanguíneo, etc.).

¿ASOCIACIÓN O CAUSALIDAD?

Los diferentes métodos disponibles para la investigación causal en medicina se pueden clasificar según el valor de cada uno de ellos para establecer, confirmar o refutar hipótesis sobre relaciones de causa a efecto (véase la figura 3).

En cada uno de los estudios descritos anteriormente se pueden registrar asociaciones entre dos variables. Sin embargo, el hecho de que dos variables estén asociadas (en un grado "estadísticamente significativo") no implica que una de ellas sea causa de la otra. Para ello es preciso considerar otras circunstancias.

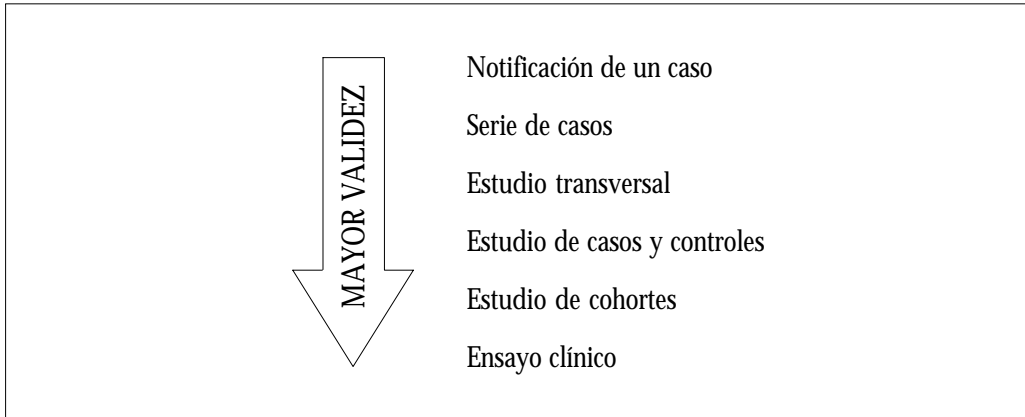


Figura 3. Las diferentes estrategias empleadas para establecer relaciones de causalidad en medicina: el ensayo clínico es el «patrón oro» de los métodos de investigación clínica, pero no siempre es factible. Las notificaciones o series de casos no incluyen grupo de control, y por lo tanto en principio pueden ser útiles para generar hipótesis, pero no para confirmarlas.

En 1945 Sir Austin Bradford Hill, considerado el padre de la estadística médica moderna, enunció nueve criterios para examinar si una asociación entre dos variables implica una relación de causalidad.¹⁴

- El primero es la **fuerza de la asociación**, que es cuantificada por el valor del riesgo relativo (RR) o de la razón de ventajas (OR). Cuanto mayor es la fuerza de la asociación, más probable es que ésta refleje una verdadera relación de causalidad, de modo que los posibles sesgos pueden haber influido sobre la cuantía de la asociación, pero es menos probable que hayan influido sobre su dirección.
- El segundo es la **consistencia** de los resultados, es decir, la correspondencia y homogeneidad de la asociación en diferentes estratos de la población estudiada (por ej., en hombres y mujeres, en mayores y menores de 50 años, etc.), en diferentes circunstancias y momentos y en diferentes estudios. Parece paradójico el hecho de que resultados diferentes obtenidos con estudios de diseño diferente no refutan necesariamente las hipótesis de un primer estudio; sin embargo, los mismos resultados obtenidos precisamente con el mismo tipo de estudio no refuerzan demasiado la hipótesis original.¹⁴ Cuando se desea confirmar una hipótesis, se debería dar prioridad a los resultados similares obtenidos en estudios de diseño diferente.
- El tercero es la **especificidad** del efecto. Cuando se describió por primera vez (con un estudio de casos y controles) que el tabaco daba lugar a cáncer de pulmón, Bradford Hill razonaba que el tabaco producía cáncer de pulmón y sólo cáncer de pulmón, y que este hecho le daba credi-

bilidad a la posible relación de causa a efecto. Él mismo matizó el criterio de la especificidad al cabo de unos años, cuando, con sus discípulos, describió también por vez primera que el hábito tabáquico es un factor de riesgo para padecer otros cánceres, bronconeumopatía obstructiva crónica y cardiopatía isquémica.

- El cuarto es la **secuencia temporal**, es decir, la garantía de que la exposición al factor de riesgo fue anterior a la enfermedad. Aunque este criterio parezca obvio, no siempre lo es, sobre todo cuando se estudian enfermedades crónicas, con un curso subclínico prolongado antes de dar lugar a algún síntoma, como la arteriosclerosis, la enfermedad de Parkinson o la demencia senil. Así por ejemplo, en un estudio sobre la relación entre la dieta y una de estas enfermedades crónicas, cabe preguntarse si la dieta da lugar a la enfermedad, o si las primeras fases de la enfermedad condicionan hábitos dietéticos particulares.
- El quinto es el **gradiente biológico** o relación dosis-respuesta: la frecuencia o la gravedad de la enfermedad en estudio debe guardar relación con la dosis o con la duración de la exposición al factor de riesgo considerado. Aunque para la mayoría de las relaciones de causalidad esta consideración puede ser válida, hay que recordar que la frecuencia o la intensidad de algunos efectos (por ej., reacciones de hipersensibilidad alérgica) no guardan relación con la dosis administrada.
- El sexto es la **plausibilidad biológica**, es decir la concordancia de la asociación encontrada con los conocimientos biológicos y clínicos sobre la exposición de interés. No obstante, éste es un criterio de importancia relativa, porque lo biológicamente plausible depende de los conocimientos biológicos que se tienen en cada momento sobre un problema determinado. En otras palabras, la asociación registrada puede ser nueva para la ciencia o la medicina y por lo tanto, en caso de "ignorancia biológica" no se puede descartar su plausibilidad.
- El séptimo es la **coherencia**: la interpretación causal de una asociación no debe entrar en conflicto con los conocimientos sobre la historia natural y la biología de la enfermedad.
- El octavo es la **experimentación**, es decir la posibilidad de aplicar una intervención y comprobar si da lugar al efecto sospechoso de ser producido por ella. (En este contexto, el término "intervención" equivale en sentido amplio a "ensayo clínico"). Este criterio es éticamente aplicable cuando un primer estudio observacional ha sugerido que una determinada intervención es preventiva o curativa: si un estudio sugiere que una determinada intervención podría prevenir una enfermedad, se puede probar a ponerla en práctica y comprobar si realmente tiene el efecto preventivo sospechado.

- El noveno y último es el **razonamiento por analogía**. En ciertas circunstancias se pueden identificar asociaciones causales de naturaleza similar, que pueden ayudar a establecer una relación de causalidad en el caso presente. Así por ejemplo, si la talidomida produce graves malformaciones congénitas, es más fácil aceptar que otros fármacos también las puedan producir.

Bibliografía

1. Spector R, Park GD. Regression to the mean: a potential source of error in clinical pharmacological studies. *Drug Intell Clin Pharm* 1985;19:916-9.
2. Laporte JR, Carné X. Metodología epidemiológica básica en farmacovigilancia. En: JR Laporte, G Tognoni (eds) *Principios de epidemiología del medicamento*, 2ª ed. Barcelona: Ediciones Científicas y Técnicas, 1993:111-30.
3. Severi S, Davies G, Maseri A, Marzullo P, L'Abbate A. Long-term prognosis of "variant" angina with medical treatment. *Am J Cardiol* 1980;46:226-32.
4. Brocks P, Dahl C, Wolf H, Transbøl IB. Do thiazides prevent recurrent idiopathic renal calcium stones? *Lancet* 1981;2:124-5.
5. Laerum E. Thiazide prophylaxis of urolithiasis: a double-blind study in general practice. *Scand J Urol Nephrol* 1983;80(suppl):70-7.
6. Scholz D, Schwille PO, Sigel A. Double-blind study with thiazide in recurrent calcium lithiasis. *J Urol* 1982;128:903-7.
7. Laporte JR, Carné X. Estudios de cohortes en farmacovigilancia. En: JR Laporte, G Tognoni (eds) *Principios de epidemiología del medicamento*, 2ª ed. Barcelona: Ediciones Científicas y Técnicas, 1993:171-97.
8. Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology - the essentials*. Baltimore: Williams & Wilkins, 1982:134-6.
9. Vessey MP, Lawless M. Estudio sobre contracepción de la Oxford-Family Planning Association. En: J Laporte, JR Laporte (eds) *Avances en terapéutica*, vol 13. Barcelona: Salvat, 1985:39-51.
10. Ibrahim MA (ed). *The case control study. Consensus and controversy*. Oxford: Pergamon Press, 1979.
11. Shapiro S, Kaufman D, Laporte JR. La estrategia de casos y controles en farmacovigilancia. En: JR Laporte, G Tognoni (eds) *Principios de epidemiología del medicamento*, 2ª ed. Barcelona: Ediciones Científicas y Técnicas, 1993:199-218.
12. Sackett DL. Bias in analytic research. *J Chron Dis* 1979;32:51-63.

13. Capellà D, Laporte JR, Vidal X, Wiholm BE, Bégaud B, Langman MJS, Rawlins M. European network for the case-population surveillance of rare diseases (Euronet). A prospective feasibility study. *Eur J Clin Pharmacol* 1998;53:299-302.
14. Bradford Hill A, Hill ID. *Principles of medical statistics*, 12^a ed. Londres: Edward Arnold, 1991.