

LUCES Y SOMBRAS DE LOS ESTUDIOS DE CASOS Y CONTROLES. IMPACTO SOBRE EL CONOCIMIENTO DE LAS MALFORMACIONES CONGÉNITAS

C. Ochoa Sangrador

Servicio de Pediatría. Hospital Virgen de la Concha. Zamora.

Introducción

El estudio de la etiología de las malformaciones congénitas constituye uno de los objetivos prioritarios en investigación epidemiológica. De los distintos diseños epidemiológicos empleados, los estudios de casos y controles reúnen ciertas características que los hacen especialmente apropiados. Otros diseños analíticos, metodológicamente más robustos, como los estudios de cohortes, resultan menos factibles por las dificultades existentes para detectar y monitorizar ciertas exposiciones y la escasa frecuencia de muchas malformaciones congénitas. Asimismo, recurrir en este escenario a diseños experimentales con humanos (ensayos clínicos), suele resultar éticamente injustificable. Por otra parte, los estudios descriptivos, útiles en vigilancia epidemiológica, nos han permitido, en algunas situaciones dramáticas de fuerte asociación entre exposición y efecto, la identificación de agentes causales y la puesta en marcha de alarmas epidemiológicas [Lenz W, 1961; McBride WG, 1961].

En los últimos 40 años la utilización creciente de los estudios de casos y controles ha aportado importantes avances al conocimiento de las malformaciones congénitas, permitiendo la identificación de múltiples factores ambientales y genéticos implicados en su etiología. Sin embargo, por sus características metodológicas, en estos estudios es frecuente cometer errores en su diseño, ejecución o análisis. El conocimiento y control de dichos errores nos ha de permitir el aprovechamiento de esta eficiente herramienta epidemiológica, compensando sus luces y sus sombras.

Metodología de los estudios de casos y controles

El estudio de casos y controles está diseñado para ayudar a determinar si una exposición está asociada a un efecto, observado como enfermedad u otra condición de interés (p.e. malformación congénita). Su diseño puede des-

cribirse de forma simple. Se parte de dos grupos de sujetos: unos que tienen la anomalía o enfermedad estudiada (casos) y otros que carecen de ella (controles). A continuación investigamos, hacia atrás en el tiempo, la existencia de los factores de exposición o riesgo presumiblemente relacionados con la anomalía o enfermedad, comparando la frecuencia de la exposición observada en ambos grupos.

Por definición son estudios observacionales (no hay intervención), analíticos (hay grupo de comparación), de direccionalidad retrógrada (primero se conoce el efecto y después se investiga la exposición) y temporalidad mixta (el investigador recaba datos históricos acaecidos previamente al inicio del estudio, mientras que suele asistir a la aparición o diagnóstico de la enfermedad). Se conocen también como estudios de "casos-testigo" o "retrospectivos", refiriéndose este término a la direccionalidad retrógrada del estudio (efecto-exposición) y no al momento en el que el investigador recopila la información.

Los estudios de casos y controles tienen algunas ventajas sobre otros diseños epidemiológicos. Son comparativamente más rápidos, baratos y fáciles, especialmente cuando hay que evaluar brotes epidémicos o enfermedades raras. Uno de los escenarios que mejor se adapta a su diseño es la investigación clínica de las malformaciones congénitas. La mayoría de ellas tienen muy baja frecuencia, por ello, si tuviéramos que realizar un estudio de seguimiento, el tamaño muestral tendría que ser extremadamente grande para poder recopilar un suficiente número de casos. Asimismo, el hecho de que la población susceptible de estudio (fetos y recién nacidos) entre necesariamente en contacto con el sistema sanitario en el momento del diagnóstico, facilita un acceso eficiente y fiable a muestras representativas de casos y controles. Otra ventaja de estos estudios es que permiten analizar simultáneamente el papel de múltiples factores de riesgo, genéticos y ambientales, que puedan intervenir como agentes causales. Además, su diseño facilita el control de los procedimientos diagnósticos y la reclasificación de los casos, en función de la información médica de la que se va disponiendo.

El aspecto más crucial de los estudios de casos y controles es la adecuada selección de los controles. Éstos serán sujetos sin la anomalía o enfermedad objeto de estudio, aunque en riesgo de tenerla (p.e. una niña no debería ser un control para el estudio del hipospadias), que sea representativa del nivel de exposición (del factor o factores de riesgo) en la población de la que surgen los casos. Un buen control será aquel que de haber estado enfermo hubiera formado parte de los casos. Debemos tratar de buscar controles similares a los casos excepto en la existencia de enfermedad, para ello se suelen emplear criterios sociodemográficos de emparejamiento. Es habitual utilizar como controles sujetos afectados de otras enfermedades, por su accesibilidad y predisposición, o controles poblacionales, menos accesibles y predisuestos, aunque, en principio, más representativos.

Como hemos comentado previamente, en el estudio de las malformaciones congénitas contamos con una fuente muestral accesible y representativa, ya que tanto los casos como los posibles controles se encuentran disponibles en el momento del nacimiento. Un estudio de casos y controles que incluya los recién nacidos malformados, así como una muestra aleatoria de recién nacidos no malformados, es mucho más eficiente, considerando el coste por unidad de información, que un estudio de cohortes. Esta estrategia requiere, sin embargo, que la exposición en los casos sea razonablemente prevalente, si no llega a ser una causa necesaria para la mayoría de los neonatos con la malformación en estudio. En esta situación, un estudio de casos y controles con un factor de exposición raro en la población resulta muy informativo. Sin embargo, esta circunstancia, lejos de ser la regla, es poco habitual [Olsen J y cols., 2002].

Un estudio de casos y controles de una malformación congénita con una base poblacional definida, habitualmente una región o país, debería incluir como casos todos los neonatos con la malformación objeto de estudio. Los controles deberían obtenerse de una muestra aleatoria de todos los recién nacidos en la región y en el periodo de tiempo de estudio, aunque habitualmente resulta adecuado usar los neonatos no malformados, considerando la rareza de esta enfermedad.

Si realizamos un muestreo apropiado de casos y controles podremos calcular estimadores de asociación no sesgados entre exposición y efecto. El cociente entre la odds de exposición en los casos (proporción de casos expuestos dividido por su complementario), y la odds de exposición en los controles, conocido como odds ratio (OR), nos proporciona una estimación del riesgo de malformaciones entre los expuestos con respecto a los no expuestos. Teniendo en cuenta la baja frecuencia de gran parte de las malformaciones congénitas, la OR calculada se aproximará mucho al riesgo relativo que se hubiera obtenido con un estudio de co-

hortes [Schlesselman JJ, 1982]. En la tabla 1 vemos, con un ejemplo figurado, cómo se puede calcular la OR. Esta medida, por estar obtenida de muestras poblacionales, debe ser acompañada del cálculo de sus intervalos de confianza. Una OR mayor de 1 expresa que el factor de exposición estudiado incrementa el riesgo de la malformación, tanto más cuanto más alejada esté la OR de 1; pero sólo cuando en el intervalo de confianza no esté incluido el 1 podremos descartar que la asociación encontrada se deba al azar. En el ejemplo empleado en la tabla 1, la OR de nuestra exposición (2,2) lo que expresa aumento del riesgo (riesgo algo mayor del doble), aunque su intervalo de confianza (0,70-6,88), que comprende el 1, sugiere que puede ser explicable por azar. Cuando la exposición sea un factor de protección la OR se situará por debajo de 1.

El talón de Aquiles de los estudios de casos y controles lo constituye la reconstrucción de la secuencia temporal de la exposición, que habitualmente requiere recuperar información histórica de los 6 a 10 meses previos. Esta información debe estar accesible y además ser situada correctamente en el tiempo. Veamos este aspecto con un ejemplo. En un estudio sobre el efecto de los inhibidores de la dihidrofolato reductasa (trimetoprim, triamterene, sulfasalazina) en los defectos cardiovasculares y hendiduras orales congénitas [Hernandez-Diaz S y cols., 2000], no tiene las mismas consecuencias tomar estos medicamentos antes, durante o después del periodo más sensible (segundo-tercer mes de gestación), por lo que resulta de gran importancia situar cada exposición correctamente en el tiempo. Por

TABLA 1
ANÁLISIS DE ESTUDIOS DE CASOS Y CONTROLES.
ODDS RATIO

FACTOR EXPOSICIÓN (O RIESGO)	EFECTO (ENFERMEDAD)		
	CASOS	CONTROLES	
Exposición presente	a	b	N1
Exposición ausente	c	d	N0
	M1	M0	T

$$\begin{aligned} \text{Odds de enfermedad en expuestos} &= a/b \\ \text{Odds de enfermedad en no expuestos} &= c/d \\ \text{Odds ratio} = \text{OR} &= \frac{a/c}{b/d} = \frac{axd}{bxc} \end{aligned}$$

Ejemplo:

FACTOR EXPOSICIÓN (O RIESGO)	MALFORMACIÓN CONGÉNITA		
	CASOS	CONTROLES	
Exposición presente	10	5	15
Exposición ausente	50	55	105
	60	60	120

$$\text{Odds ratio} = \text{OR} = \frac{10/50}{5/55} = \frac{10 \times 55}{5 \times 50} = 2,20$$

ello, en estos estudios de casos y controles interesa no retrasar la recogida de los antecedentes de riesgo, para evitar errores relacionados con la cronología de la exposición.

Asimismo, resulta de gran importancia que la recogida de datos sobre factores de riesgo en el pasado sea objetiva, especialmente si quien la recoge ya sabe si está entrevistando a un caso o a un control. En algunos estudios, es posible acceder a datos cumplimentados prospectivamente durante la gestación, disponibles en registros médicos, que nos proporcionarán información no sesgada, al estar recopilada antes del diagnóstico. Sin embargo, cierta subjetividad en el recuerdo de antecedentes por parte de madres o familiares parece inevitable. Clásicamente se ha interpretado que los controles tienen una menor participación y una menor predisposición a recordar antecedentes de riesgo que los casos [Rockenbauer M y cols., 2001]. Una opción para minimizar este problema es emplear como controles, para el estudio de una determinada malformación, los casos de otras malformaciones no relacionadas (por ejemplo, niños con Síndrome de Down para el estudio del efecto de una exposición tóxica) [Czeizel AE y cols., 1998; Lief S y cols., 1999].

Hemos visto las peculiares características que tienen los estudios de casos y controles, que los hacen apropiados para la investigación de las malformaciones congénitas. Sin embargo, la interpretación de sus resultados debe ser cautelosa porque, como veremos más adelante, estos estudios presentan también limitaciones relacionadas con su direccionalidad retrógrada: dificultades para la selección de controles, problemas para la recopilación histórica de ciertos antecedentes y potencial susceptibilidad a los sesgos. No obstante, son la forma más eficiente de identificar asociación entre factores de riesgo y malformaciones congénitas (que no causalidad), y en ocasiones la única vía éticamente posible para hacerlo.

Errores en los estudios de casos y controles

Como cualquier otro tipo de estudio epidemiológico, los estudios de casos y controles deben ser entendidos como un ejercicio de medición, en el que ha de perseguirse la exactitud, esto es, la realización de estimaciones con el mínimo error. Existen dos posibles fuentes de error: los errores aleatorios y los errores sistemáticos.

Los **errores aleatorios** son debidos a la variabilidad que el azar introduce en el proceso de medición. Estos errores no suelen afectar a la validez interna de los estudios, pero sí limitan su potencia. En epidemiología, la principal forma de reducir el error aleatorio es aumentar el tamaño de la muestra. De esta manera, minimizamos el riesgo de obte-

ner resultados distorsionados por azar. En los estudios sobre malformaciones congénitas, esta cuestión viene a menudo condicionada por la rareza de muchas de ellas. Por ello, los estudios colaborativos son una buena estrategia para obtener muestras suficientes y representativas. En relación con la cuantificación del error aleatorio, es preciso señalar que en estudios en los que se evalúan simultáneamente múltiples factores de riesgo y malformaciones congénitas, la asunción del margen arbitrario de error del 5% ($p < 0,05$) puede resultar arriesgada, ya que, por puro azar, puede darse en 1 de cada 20 contrastes [Czeizel AE y cols., 1998].

Mayor importancia tienen los **errores sistemáticos**, especialmente frecuentes en los estudios de casos y controles. Por error sistemático entendemos la falta de validez en la medición. También se le denomina sesgo (bias). La validez se refiere al concepto de "medir lo que se quiere realmente medir". Los errores sistemáticos o sesgos son atribuibles a errores metodológicos que, a diferencia de los errores aleatorios, no se reducen aumentando el tamaño muestral. Podemos clasificar los sesgos en sesgos de selección, sesgos de información (o clasificación) y sesgos de análisis (confusión e interacción).

1. Sesgos de selección

Existe un sesgo de selección cuando se produce una distorsión del efecto estimado como consecuencia de errores en el proceso de selección de los sujetos. Si la elección de los grupos a comparar no es apropiada se producirá un sesgo de selección. En el estudio sobre malformaciones congénitas, es frecuente encontrar sesgos de selección por no incluir las gestaciones con malformaciones que terminan en aborto. Cualquier factor que incremente el riesgo de aborto podrá mostrarse como factor de protección, al verse infrarrepresentado en las gestaciones que llegan a término. Este sesgo de selección prenatal, es una variante del conocido como sesgo de prevalencia de Neyman, y ha sido observado, por ejemplo, en relación al Síndrome de Down con distintos factores como el tabaco [Hook EB, Cross PK, 1985] y el consumo de café [Torfs CP, Christianson RE, 2000]. Pero no sólo los abortos espontáneos pueden provocar sesgos de selección de casos, también los abortos inducidos. Aquellas anomalías que son más susceptibles de desencadenar abortos espontáneos o interrupciones voluntarias, serán más sensibles a estos sesgos. Por ejemplo, la no inclusión de abortos tendrá más repercusión en estudios sobre defectos del tubo neural que en los de defectos del tracto urinario [Ray JG y cols., 2004].

Si bien es difícil ponerlo en práctica, lo ideal sería que los estudios sobre anomalías con alta probabilidad de pérdida

prenatal incluyeran los casos diagnosticados en nacidos vivos, nacidos muertos y pérdidas prenatales tempranas.

Otra fuente de error, propia de las malformaciones congénitas, es la pérdida de casos cuyas manifestaciones no son aparentes al nacimiento. Pensemos, por ejemplo, en algunas cardiopatías congénitas que frecuentemente son diagnosticadas a lo largo del primer año de vida, o incluso más adelante, por lo que no son incluidas entre los casos, o que incluso pueden ser incorporadas a los controles. Para controlar este error es aconsejable establecer mecanismos de seguimiento y notificación de nuevos hallazgos en casos y controles, y procedimientos de recaptura a partir de fuentes alternativas de datos (p.e. servicios de cardiología infantil).

La utilización de criterios de restricción para la selección de controles puede dar lugar también a sesgos de selección. Si optamos por escoger como controles, recién nacidos con malformaciones distintas a la que es objeto de estudio, también llamados pseudocontroles, podemos enmascarar el efecto de algún factor de riesgo que esté relacionado con esas malformaciones control [Swan SH y cols., 1992]. Este error, denominado "sesgo de teratogenicidad no específica" [Hokok EB, 2000], debe ser controlado en el diseño, pudiendo utilizarse, para ello, la comparación entre los resultados de grupos control alternativos. Lieff y cols. [Lieff S y cols., 1999] estudiaron el riesgo de labio leporino ligado al consumo de tabaco materno, a través de estudios de casos y controles con varios grupos control: 1) excluyendo malformaciones relacionadas con el tabaco, 2) excluyendo aquellas que tuvieran una prevalencia de tabaquismo 1 o más desviaciones estándar por encima o debajo de la media de todos los malformados, 3) incluyendo sólo malformaciones ligadas a herencia mendeliana y 4) excluyendo sólo los síndromes asociados a labio leporino. El estimador de riesgo (OR cercana a 1,5) no mostró variaciones al utilizar distintos grupos control. Este resultado sugería la inexistencia de un sesgo de teratogenicidad no específica. Sin embargo, no siempre disponemos de información etiológica fiable, de manera que algunos factores de riesgo que tengan un bajo grado de asociación con otras malformaciones pueden originar pequeños sesgos de selección, por lo que el uso de grupos control restrictivos podría ser inapropiado.

Otro entorno en el que es posible incurrir en sesgos de selección es el del estudio de la agregación familiar o riesgo de recurrencia de malformaciones congénitas. Es práctica habitual en estudios de casos y controles, comparar el antecedente de familiares afectos en ambos grupos, como forma de identificar la existencia de agregación familiar. No obstante, resulta más adecuado calcular el riesgo de recurrencia a partir de la reconstrucción de un estudio de cohortes, en el que comparamos el riesgo de recurrencia en la co-

horte de familiares de los casos con el de la cohorte de familiares de los controles [Bai Y y cols., 2000]. Sin embargo, el riesgo de recurrencia estimado va a variar en función de que incluyamos o no más de una vez a las familias con múltiples casos. Si sólo incluimos cada familia una vez, obtendremos estimadores sesgados de la verdadera recurrencia, con una infraestimación directamente proporcional a la probabilidad de selección de casos y controles en nuestro estudio; sólo si estas probabilidades son bajas (para los casos malformados es poco probable) obtendremos estimadores válidos. Por ello, resulta más apropiado incluir los familiares tantas veces como casos o controles de su familia sean seleccionados. Llama la atención que esta estrategia, en principio más válida, haya sido utilizada en pocas ocasiones [Sánchez-Cascos A, 1978].

2. Sesgos de información

Los sesgos de información o clasificación consisten en la distorsión del efecto estimado por errores en la medición y/o clasificación de los sujetos en una o más variables. En los estudios de casos y controles suelen estar relacionados con problemas para identificar antecedentes históricos, condicionados, consciente o inconscientemente, por los sujetos encargados de recopilar la información, investigadores o familiares.

Estos errores pueden distorsionar nuestros resultados, especialmente si son diferenciales. Entendemos por error de clasificación diferencial el producido cuando el criterio de clasificación en uno de los grupos que se comparan es diferente del criterio utilizado en el otro grupo, pudiéndose clasificar a los sujetos de un grupo incorrectamente y a los del otro correctamente, o a los dos incorrectamente pero en distinta dirección o magnitud. La estimación del efecto puede verse aumentada o disminuida, produciéndose una distorsión en contra o a favor de la hipótesis nula, respectivamente.

El error de clasificación no diferencial es el cometido cuando se clasifica incorrectamente a los dos grupos por igual. La clasificación de la exposición es errónea para iguales proporciones de sujetos en los dos grupos de comparación. Este error tiende a producir una subestimación del efecto, esto es, una distorsión a favor de la hipótesis nula y, en consecuencia, genera resultados más conservadores. Por ello el error no diferencial ocasiona menos problemas que el error diferencial.

En el estudio de las malformaciones congénitas es frecuente que se produzcan sesgos diferenciales con relación al recuerdo de antecedentes de riesgo. Se ha sugerido que las madres de niños malformados realizan un mayor esfuer-

zo para recordar que las madres de los controles, lo que podría llegar a incriminar como teratogénicos factores que no lo son [Swan SH y cols., 1992]. Sin embargo este sesgo, conocido como **sesgo de recuerdo**, largamente discutido en la literatura, ha sido poco documentado, fundamentalmente, por las dificultades existentes para encontrar un patrón de referencia válido a la hora de medir algunas exposiciones [Lief S y cols., 1999; Swan SH y cols., 1992]. Diversos autores prefieren denominar a este sesgo "sesgo de declaración" para expresar que, en ocasiones, la información podría ser recordada pero no manifestada por temor, si la exposición conlleva rechazo social, o por aturdimiento materno, relacionado con el anuncio del diagnóstico de la malformación de su hijo [Swan SH y cols., 1992].

En 1989 Werler y cols. [Werler MM y cols., 1989] compararon los datos obtenidos en una entrevista postparto, con los existentes en registros médicos, encontrando que el grado de recuerdo de varios factores de riesgo (consumo de medicamentos, infecciones urinarias o por hongos, uso de anticonceptivos tras la concepción, antecedente de aborto o infertilidad) era mayor entre las madres con hijos malformados que entre las que tuvieron hijos normales. Sin embargo, algunos autores cuestionan la validez de los registros médicos, por ser incompletos y poder presentar errores [Hewson D, Bennett A, 1987], aunque tienen la ventaja de no estar sesgados, si se cumplimentan antes de conocer el resultado gestacional [Rockenbauer M y cols., 2001].

Fraser y Warburton [Fraser FC, Warburton D, 1964; Fraser FC, 2000] recopilaron información sobre el antecedente de episodios de estrés materno durante la gestación, en 187 casos con labio leporino (con o sin fisura palatina), en 59 con fisura palatina aislada y en sus hermanos sanos. Encontraron antecedente de estrés materno en la gestación en el 18% y 15% de los casos con labio leporino y fisura palatina, respectivamente, pero sólo en un 5% y 3% de sus hermanos. Este hallazgo podría ser erróneamente interpretado como existencia de teratogenicidad; sin embargo, al buscar la misma información en un grupo de 90 casos con trastornos de herencia Mendeliana (albinismo, hemofilia, fibrosis quística, etc.), no atribuibles a incidentes durante la gestación, y en sus hermanos sanos, encontraron resultados similares (14 y 4% respectivamente). Las diferencias observadas podrían ser explicadas por un sesgo de recuerdo, debido en parte a una mayor motivación para recordar en los casos, y en parte, particularmente en este estudio, a un mayor intervalo entre la gestación y la entrevista en los controles. Aunque lo habitual es que el sesgo de recuerdo presente una tendencia diferencial hacia una exageración del efecto, en contra de la hipótesis nula, para algunos factores de riesgo se han observado tendencias contrarias [Feldman Y y cols., 1989; Delgado Rodríguez M y cols., 1993].

En 1989, Feldman y cols. [Feldman Y y cols., 1989], compararon la información recogida en el primer trimestre de la gestación, en un servicio de consulta prenatal por exposición a medicamentos, drogas, productos químicos y radiaciones, con la obtenida en una entrevista postparto. De media, las madres sólo recordaban el 62% de los productos a que habían estado expuestas, mientras que la dosis y el momento de la gestación sólo del 24% y 37% de los productos, respectivamente. Sin embargo, comparando el grado de recuerdo de las madres que tuvieron una gestación con un resultado adverso con el de las que lo tuvieron normal no había diferencias, excepto en el recuerdo del consumo de alcohol, que fue menor en las primeras [Feldman Y y cols., 1989]. Es posible que las mujeres incluidas en este estudio fueran una muestra sesgada, por considerarse tanto casos como controles bajo un riesgo especial.

Rockenbauer y cols. [Rockenbauer M y cols., 2001] analizaron la exposición a drogas de 22.865 casos con malformaciones congénitas y 39.151 controles sanos, comparando los datos registrados en una cartilla maternal, cumplimentada por los obstetras (no incluía prescripciones de otros facultativos o automedicación), con los declarados por la madre a través de un cuestionario por correo en los primeros meses tras el nacimiento. Para los casos, el 84% de los tratamientos de corta duración declarados por las madres y el 75% de los de larga duración no estaban reflejados en la cartilla maternal. Para los controles, esas cifras descendían a 72% y 54%, expresando una menor declaración no registrada en este grupo. Asimismo, encontraron que los casos acertaban más el periodo de tiempo desde la toma del medicamento hasta la encuesta que los controles. Como resultado de un posible sesgo de recuerdo, los datos basados en la entrevista mostraban OR por encima de 1, indicadoras de riesgo, que al considerar los datos de la cartilla maternal descendían, alcanzando incluso cifras propias de factores protectores.

Un estudio de casos y controles multinacional [Kallen B y cols., 1992] que analizaba el papel de distintos factores de riesgo sobre el hipospadias, cuantificaron la magnitud del posible sesgo de recuerdo, comparando las OR calculadas con pares de casos-controles discordantes (caso expuesto con su control no expuesto y viceversa). Examinando las OR obtenidas para un numeroso conjunto de factores, diferenciando si la exposición era antes, durante o después del periodo más sensible a la formación de hipospadias (semanas 8 a 16 de gestación), observaron que la mayoría eran superiores a 1, con un valor medio cercano a 1,3. Este resultado parece reflejar la existencia de un sesgo de recuerdo, pudiendo cuantificar su nivel en dicha cifra, de manera que el valor nulo para contrastar el límite inferior de los intervalos de confianza debería ser 1,3 y no 1. Sin embargo,

ante la ausencia de un patrón de referencia totalmente válido para clasificar la exposición, resulta arriesgado cuantificar el sesgo de recuerdo diferencial, aunque se han descrito modelos matemáticos para estimar su impacto [Flanders WD y cols., 1995].

A pesar de que la mayoría de los autores asumen que el sesgo de recuerdo tiende siempre a exagerar la estimación del efecto (en contra de la hipótesis nula), esta asunción no parece generalizable. La magnitud y dirección del sesgo de recuerdo va a depender de la metodología de recogida de información en cada estudio; fundamentalmente, del momento en que se hace, quién la recopila, de qué fuentes y cómo. No es lo mismo recabar información mediante cuestionario por correo, meses después del nacimiento, que entrevistar a las madres en los primeros días tras el parto y por personal adecuadamente formado. De hecho, se ha sugerido que, en condiciones ideales, la pérdida de información podría ser mayor en los casos que en los controles, por encontrarse las madres de los primeros conmovidas por la mala noticia, e influenciadas por un instintivo sentimiento de culpabilidad [Martínez-Frías ML, 1993]. En ocasiones, se ha recurrido a argumentos similares para explicar algunos resultados paradójicos [Mills JL y cols., 1989].

Como se ha comentado anteriormente, es habitual recurrir a controles con malformaciones no relacionadas con la que es objeto de análisis (seudocontroles). Con esta estrategia se espera minimizar el sesgo de recuerdo, aunque podemos incurrir en un sesgo de selección de "teratogenicidad no específica". Debemos tener en cuenta que ambos sesgos suelen tener una diferente dirección, el primero hacia una sobreestimación del efecto y el segundo hacia una infraestimación, aunque son posibles otras situaciones. Se han desarrollado modelos para estimar la dirección e intensidad de estos sesgos [Swan SH y cols., 1992], en los que los factores a considerar son el porcentaje de clasificación errónea en casos y en controles y la probabilidad de incluir en el grupo control, pseudocontroles con otras malformaciones.

Prieto y Martínez-Frías [Prieto L, Martínez-Frías ML, 1999] estudiaron la asociación entre la exposición materna a ácido valproico y espina bífida, usando como controles, por separado, recién nacidos normales y malformados. Las OR obtenidas fueron alrededor de 50 y 15 respectivamente, ambas indicadoras de riesgo, aunque diferentes. Resulta, pues, interesante emplear ambos tipos de controles para identificar efectos teratogénicos. Hook ha sugerido [Hook EB, 2000; Hook EB, 1993 (a,b); Hook EB, Regal RR, 1993] que la verdadera estimación del efecto se encontrará entre ambas estimaciones, sin embargo esta asunción resulta arriesgada, ya que en ocasiones podemos encontrar OR de riesgo (por encima de 1), a partir del cociente de 2 riesgos relativos protectores [Prieto L, Martínez-Frías ML, 2000]. En esta situa-

ción, el valor obtenido a partir de controles malformados no será el límite inferior del verdadero riesgo. Por ello, el uso separado de diferentes controles no sólo facilita el control de sesgos [Czeizel AE y cols., 2001 (a)], si no que además nos permite distinguir si un factor concreto es teratogénico o protector [Prieto L, Martínez-Frías ML, 2000]. No obstante, debemos recordar que, conceptualmente, los controles deberían ser representativos del nivel de exposición a estudio, en la población de la que proceden los casos; el cumplimiento o no de este criterio va a modificar la cuantificación del riesgo resultante.

En investigación epidemiológica es frecuente que se obtenga información sobre exposición a agentes ambientales, a partir del domicilio de residencia. Si asumimos como válido el domicilio en el momento del parto, las gestantes que se hayan desplazado después del primer trimestre de embarazo, periodo de alta susceptibilidad, podrán ser incorrectamente clasificadas. Este hecho podría tener trascendencia si el traslado de domicilio es más frecuente en algún grupo concreto de población (mujeres más jóvenes, de distinta formación académica o de distinta raza) [Schulman J y cols., 1993].

3. Sesgos de información

Mientras los sesgos de selección o clasificación, sólo pueden ser previstos y corregidos en la fase de diseño, otros errores pueden ser controlados en la fase de análisis. De estos sesgos de análisis, los principales son la confusión y la interacción.

Cuando el efecto estimado de la relación exposición - enfermedad está distorsionado por la presencia de algún factor extraño en el análisis de los datos, decimos que existe **confusión**. El estimador del efecto de interés está distorsionado por estar mezclado con el efecto producido por un factor extraño que llamamos factor (o variable) de confusión. La presencia de tal factor de confusión altera los resultados y puede producir un aumento o disminución del efecto, o incluso cambiar la dirección del mismo.

Para que un factor sea de confusión necesita cumplir unos requisitos: I) debe ser factor de riesgo independiente de la enfermedad tanto para los expuestos como para los no expuestos en la población, II) estar asociado al factor de exposición que estudiamos en la población de donde provienen los casos, y III) no ser una consecuencia de dicha exposición; es decir, no ser un paso intermedio en la cadena o secuencia de causalidad entre la exposición y la enfermedad [Schlesselman JJ, 1982].

Un sesgo de confusión puede corregirse previamente al inicio del estudio durante el diseño del mismo, empleando

técnicas de restricción, pero también puede corregirse en el análisis, mediante técnicas de ajuste, fundamentalmente el análisis estratificado y el análisis multivariante. Es preciso advertir que la utilización de criterios de apareamiento restrictivos, usando posibles factores de confusión, puede dar lugar a sesgos que dirigen la estimación del riesgo hacia el valor nulo, produciéndose un fenómeno de sobreapareamiento [Pelz J, 1996]. Este fenómeno es más probable con factores influenciados que con factores no influenciados (como edad y sexo); cuando están involucrados los primeros es preferible recurrir a técnicas de ajuste.

Un ejemplo de sesgo de confusión es el observado con la edad materna en el estudio de factores de riesgo de Síndrome de Down. La edad materna, factor de riesgo claramente identificado, se comporta como factor de confusión con otras variables relacionadas con ella como el tabaquismo materno. Una asociación inversa entre éste y la edad materna (a mayor edad menor consumo de tabaco) ha motivado que el tabaquismo pueda aparecer como factor de protección [Chen CL y cols., 1999], sino se realiza un ajuste en función de la edad de la madre. Hay que tener en cuenta que en ocasiones las técnicas de ajuste no llegan a controlar la influencia de los factores de confusión, esto podría explicar algunos hallazgos llamativos, como la asociación entre el uso de ampicilina en el segundo-tercer mes de gestación y la aparición de hendidura palatina congénita (por control insuficiente de enfermedades o medicamentos concomitantes) [Czeizel AE y cols., 2001 (b)]. Ante resultados no refrendados en otros estudios o sin una plausible explicación biológica, debemos sospechar que pueden ser fruto de error aleatorio o sistemático.

También podemos encontrar errores al analizar variables relacionadas con la consanguinidad. La consanguinidad incrementa la probabilidad de homocigosis de rasgos genéticos recesivos deletéreos y a su vez se asocia con numerosos factores demográficos, culturales, religiosos y geográficos, que pueden comportarse como factores de confusión.

La repercusión de los factores de confusión sobre la estimación del efecto de una determinada exposición va a depender de su prevalencia. En el estudio multinacional de Källén y cols. [Kallen B y cols., 1992], sobre el papel de las hormonas exógenas en el hipospadias, los distintos programas participantes mostraban grandes diferencias en la presencia de potenciales factores de riesgo, relacionados o no con la ingesta de hormonas sexuales (amenaza de aborto, abortos inducidos o espontáneos, antecedente de infertilidad, infecciones y tratamientos); estas diferencias eran en parte debidas a diferencias reales, y en parte a variaciones en la metodología de captación de casos e información (programa rutinario vs específico; entrevista en el momento del nacimiento vs diferida). Los países con mayor pres-

cripción de hormonas eran más proclives a encontrar otros factores relacionados con el hipospadias. A pesar de que el análisis conjunto mostró asociación estadísticamente significativa, los autores del trabajo no pudieron consensuar la interpretación de sus resultados [Kallen B y cols., 1992].

El otro sesgo de análisis que debemos controlar es el de modificación del efecto o **interacción**. Este sesgo ocurre cuando el efecto producido por una determinada variable se ve modificado en presencia de una tercera variable, produciéndose efectos de distinta intensidad, en la misma o distinta dirección. Este fenómeno debe ser distinguido del de confusión y también controlado en el análisis. Un ejemplo de interacción, en distinta dirección, puede verse en un estudio sobre el efecto de la toma de inhibidores de la dihidrofolato reductasa sobre defectos cardiovasculares congénitos, en el que el consumo de polivitamínicos (conteniendo ácido fólico) reducía el riesgo estimado para los inhibidores, hasta hacerlo no significativo [Hernandez-Diaz S y cols., 2000].

Considerando los potenciales errores que pueden ocurrir en estos estudios, es frecuente hallar trabajos publicados en los que a pesar de encontrar asociación estadísticamente significativa entre un factor de exposición y alguna malformación congénita, sus autores no defienden su validez, argumentando que probablemente son fruto de errores aleatorios o sistemáticos (sesgos de selección, recuerdo y confusión fundamentalmente) [Czeizel AE y cols., 1998; Czeizel AE y cols., 2001]. Son más fácilmente defendibles aquellos hallazgos con referencias concordantes en la literatura, biológicamente plausibles, estables ante diferentes tipos de controles y técnicas de ajuste, con clara relación temporal y gradiente biológico.

Conclusiones

Hemos revisado las características metodológicas de los estudios de casos y controles y valorado el grado de adecuación de las mismas al estudio de las malformaciones congénitas. A pesar de no ser metodológicamente los estudios más robustos, por criterios de factibilidad, son la forma más eficiente de identificar asociación entre factores de riesgo y malformaciones congénitas, y en ocasiones la única vía éticamente posible para hacerlo. Su contribución al conocimiento de las malformaciones congénitas ha sido de gran importancia.

También hemos visto que en estos estudios es habitual cometer errores en su diseño, ejecución o análisis. Estos errores pueden limitar la validez interna o externa de los resultados, por lo que es frecuente que muchos autores interpreten con grandes reservas hallazgos formalmente significativos, pero inesperados.

A partir del conocimiento de las principales limitaciones de estos estudios, podemos dictar algunas recomendaciones para reducir al máximo sus errores:

- Ante la escasa prevalencia de muchas malformaciones congénitas se requiere un esfuerzo de colaboración, que permita obtener muestras de casos suficientes, con las que reducir el error aleatorio y aumentar la potencia estadística de nuestros resultados. Esta colaboración precisa de apoyo institucional a escala nacional, para el desarrollo y sostenimiento de programas de vigilancia y estudio de las malformaciones congénitas, así como de iniciativas de cooperación a escala internacional.
- El estudio simultáneo de múltiples factores de riesgo y malformaciones congénitas diferentes, sin hipótesis de trabajo predeterminadas, puede requerir la asunción de un margen de error aleatorio más exigente.
- En el estudio de malformaciones congénitas con alto riesgo de pérdidas antenatales (por aborto espontáneo o inducido) debemos tratar de incluir, en la medida de lo posible, los casos identificados entre fetos y mortinatos, o al menos estimar su frecuencia, para evitar o controlar posibles sesgos de selección.
- Los criterios de inclusión de controles han de considerar que el nivel de exposición de los mismos sea representativo del de la población origen de los casos. Si incluimos pseudocontroles malformados en el grupo control, éstos no deberían estar relacionados con el factor de exposición a estudio.
- El estudio de la agregación familiar o riesgo de recurrencia de las malformaciones requiere la construcción de estudios de cohortes retrospectivos, a partir de los familiares de casos y controles.
- La metodología empleada en la recopilación de antecedentes de riesgo debería ser objetiva (idealmente enmascarada), lo más cercana posible al diagnóstico, realizada con procedimientos estandarizados y precisos y por personal adecuadamente preparado. La utilización de información complementaria recogida antes del diagnóstico, en registros médicos o estudios prospectivos, debe ser valorada, sopesando la posible falta de datos con su menor susceptibilidad a los sesgos.
- Cuando se sospeche que los antecedentes de riesgo puedan ser declarados de forma diferente en casos y controles, resulta útil contrastar los resultados con los obtenidos con grupos control alternativos. Resultados discordantes deben ser interpretados sobre la base de los conocimientos previos y su plausibilidad biológica. Resulta arriesgado estimar a priori la dirección de los sesgos de recuerdo o declaración, o introducir correcciones cuantitativas en las medidas de riesgo resultantes.
- Los factores de riesgo cuyo efecto esté relacionado con el momento de la exposición, deben ser situados correctamente en el tiempo, constituyendo el análisis de la exposición en periodos diferentes, una buena forma de contrastar su efecto.
- A la hora de caracterizar la asociación entre una determinada exposición y una malformación congénita, es preciso controlar el impacto de otros factores, relacionados con exposición y efecto, que pudieran producir confusión o interacción. El sobreapareamiento en la selección de controles, con variables influenciadas, ha de ser evitado, pudiendo en estos casos recurrirse a técnicas de ajuste.
- Cualquier atribución causal sugerida en los resultados de estudios de casos y controles, aún siendo metodológicamente válidos, debería ser concordante con los conocimientos previos y tener fundamento biológico; habitualmente habrá de requerir su confirmación con otros estudios. La consideración de estas recomendaciones pueden permitirnos aprovechar esta eficiente herramienta epidemiológica en el estudio de las malformaciones congénitas, para aportar luz a las múltiples lagunas existentes en el conocimiento de sus causas, iluminando las sombras de sus limitaciones metodológicas.

Referencias

- Bai Y, Sherman S, Khoury MJ, Flanders WD (2000): Bias associated with study protocols in epidemiologic studies of disease familial aggregation. *Am J Epidemiol* 151(9):927-937.
- Chen CL, Gilbert TJ, Daling JR (1999): Maternal smoking and Down syndrome: the confounding effect of maternal age. *Am J Epidemiol* 149:442-446.
- Czeizel AE, Rockenbauer M, Sorensen HT, Olsen J (2001,a): Use of cephalosporins during pregnancy and in the presence of congenital abnormalities: a population-based, case-control study. *Am J Obstet Gynecol* 184(6):1289-1296.
- Czeizel AE, Rockenbauer M, Sorensen HT, Olsen J (2001,b): A population-based case-control teratologic study of ampicillin treatment during pregnancy. *Am J Obstet Gynecol* 185(1):140-147.
- Czeizel AE, Rockenbauer M. A (1998): population based case-control teratologic study of oral metronidazole treatment during pregnancy. *Br J Obstet Gynaecol* 105(3):322-327.
- Delgado Rodriguez M, Jimenez Jimenez C, Lardelli Claret P, Lopez-Moratalla M, Galvez R (1993): Possible recall bias in a case-control study of congenital hip dysplasia. *Gac Sanit* 7(39):274-281.
- Feldman Y, Koren G, Mattice K, Shear H, Pellegrini E, MacLeod SM (1989): Determinants of recall and recall bias in studying drug and chemical exposure in pregnancy. *Teratology* 40(1):37-45.
- Flanders WD, Drews CD, Kosinski AS (1995): Methodology to correct for differential misclassification. *Epidemiology* 6(2):152-156.
- Fraser FC, Warburton D (1964): No association of emotional stress or vitamin supplement during pregnancy to cleft lip or palate in man. *Plast Reconstr Surg* 33: 295-299.

- Fraser FC (2000): Recall bias in case control studies of malformed infants. *Teratology* 62(6):371.
- Hernandez-Diaz S, Werler MM, Walker AM, Mitchell AA (2000): Folic acid antagonists during pregnancy and the risk of birth defects. *N Engl J Med* 343(22):1608-1614.
- Hewson D, Bennett A (1987): Childbirth research data: medical records or women's reports? *Am J Epidemiol* 125:484-491.
- Hook EB, Cross PK (1985): Cigarette smoking and Down syndrome. *Am J Hum Genet* 37:1216-1224.
- Hook EB (2000): What kind of controls to use in case control studies of malformed infants: recall bias versus 'teratogen nonspecificity' bias. *Teratology* 61(5):325-326.
- Hook EB (1993, a): Normal or affected controls in case-control studies of congenital malformations and other birth defects: reporting bias issues. *Epidemiology* 4(2):182-184.
- Hook EB (1993, b): Estimation and the nature of controls in birth defect studies. *Epidemiology* 4(6):558-559.
- Hook EB, Regal RR (1993): Relation of the crude relative risk of a disorder to relative risks in strata of a 'susceptible' variable. *Epidemiology* 4(6):524-529.
- Kallen B, Castilla EE, Robert E, Lancaster PA, Kringelbach M, Mutchnick O, Martínez-Frías ML, Mastroiacovo P (1992): An international case-control study on hypospadias. The problem with variability and the beauty of diversity. *Eur J Epidemiol* 8(2):256-263.
- Lenz W (1961): Thalidomide and congenital abnormalities. *Lancet* 1:45.
- Lieff S, Olshan AF, Werler M, Savitz DA, Mitchell AA (1999): Selection bias and the use of controls with malformations in case-control studies of birth defects. *Epidemiology* 10(3):238-241.
- Martínez-Frías ML (1993): Interviewer bias and maternal bias. *Teratology* 47(6):531-532.
- McBride WG (1961): Thalidomide and congenital abnormalities. *Lancet* 1:1358.
- Mills JL, Rhoads GG, Simpson JL, Cunningham GC, Conley MR, Lassman MR, Walden ME, Depp OR, Hoffman HJ (1989): The absence of a relation between the periconceptional use of vitamins and neural-tube defects. National Institute of Child Health and Human Development Neural Tube Defects Study Group. *N Engl J Med* 321:430-435.
- Olsen J, Czeizel A, Sorensen HT, Nielsen GL, de Jong van den Berg LT, Irgens LM, Olesen C, Pedersen L, Larsen H, Lie RT, de Vries CS, Bergman U (2002): How do we best detect toxic effects of drugs taken during pregnancy?. A EuroMap paper. *Drug Saf* 25(1):21-32.
- Pelz J (1996): Natural family planning and Down syndrome—matching and confounding in a case control study. *Am J Med Genet* 66(3):363-364.
- Prieto L, Martínez-Frías ML (1999): Case-control studies using only malformed infants: are we interpreting the results correctly? *Teratology* 60:1-2.
- Prieto L, Martínez-Frías ML (2000): Response to "What kind of controls to use in case control studies of malformed infants: recall bias versus 'teratogen nonspecificity' bias". *Teratology* 62(6):372-373.
- Ray JG, Vermeulen MJ, Meier C, Wyatt PR (2004): Risk of congenital anomalies detected during antenatal serum screening in women with pregestational diabetes. *Qjm* 97(10):651-653.
- Rockenbauer M, Olsen J, Czeizel AE, Pedersen L, Sorensen HT, EuroMAP Group (2001): Recall bias in a case-control surveillance system on the use of medicine during pregnancy. *Epidemiology* 12:461-466.
- Sánchez-Cascos A (1978): The recurrence risk in congenital heart disease. *Eur J Cardiol* 7:197-210.
- Schlesselman JJ (1982): Case-control studies: design, conduct, analysis. New York: Oxford University Press.
- Schulman J, Selvin S, Shaw GM, Malcoe LH (1993): Exposure misclassification due to residential mobility during pregnancy in epidemiologic investigations of congenital malformations. *Arch Environ Health* 48(2):114-119.
- Swan SH, Shaw GM, Schulman J (1992): Reporting and selection bias in case-control studies of congenital malformations. *Epidemiology* 3(4):356-363.
- Torfs CP, Christianson RE (2000): Effect of maternal smoking and coffee consumption on the risk of having a recognized Down syndrome pregnancy. *Am J Epidemiol* 152:1185-1191.
- Werler MM, Pober BR, Nelson K, Holmes LB (1989): Reporting accuracy among mother of malformed and nonmalformed infants. *Am J Epidemiol* 129:415-421.