

¿Qué es una variable de confusión?



Jokin de Irala, Miguel Ángel Martínez-González y Francisco Guillén Grima*

Unidad de Epidemiología y Salud Pública. Facultad de Medicina. Universidad de Navarra.

*Departamento de Ciencias de la Salud. Universidad Pública de Navarra. Pamplona.

El concepto de confusión es probablemente uno de los más importantes de la epidemiología general. En primer lugar, porque gran parte del trabajo que se realiza en este campo de la ciencia consiste precisamente en intentar prevenirlo al diseñar estudios de investigación o controlar su efecto cuando aparece en los trabajos de investigación llevados a cabo. En segundo lugar, y concretamente en lo que se refiere a los profesionales de la salud, porque de la adecuada comprensión de este fenómeno dependerá que puedan interpretar, de manera crítica y correcta, los resultados de los muchos estudios que se publican en la bibliografía científica.

En esta revisión se pretende explicar, de manera didáctica y con la ayuda de varios ejemplos, el concepto de confusión para posteriormente hacer lo mismo con otro concepto importante, el de modificación del efecto (interacción), y finalmente describir las diferencias entre ambos conceptos.

Concepto

Aunque se pueden buscar algunos antecedentes en Francis Bacon, el primer autor que trató explícitamente el tema de la confusión fue el filósofo y economista británico John Stuart Mill (1806-1873)¹. Al referirse a los criterios necesarios para el establecimiento de una relación causal, Mill señalaba la necesidad de asegurarse de que no estuviese presente ningún factor que tuviera efectos susceptibles de confundirse con el agente que se quisiera estudiar. Antes de definir el fenómeno de confusión, es preciso describir el enfoque contrafáctico de los modelos biológicos². Ante la presencia de unos datos referentes a 15 recién nacidos con defectos del tubo neural (DTN) en una muestra de 10.000 mujeres con déficit de ácido fólico, podríamos preguntarnos si la incidencia de malformaciones es debida al déficit de ácido fólico. La pregunta es importante porque, de ser afirmativa la respuesta, tendríamos en nuestras manos una solución sencilla del problema: por ejemplo, el enriquecimiento de alimentos con ácido fólico.

Para responder a dicha pregunta, es necesario comparar este grupo de mujeres con otro que tuviera valores normales de ácido fólico. De ser cierta la hipótesis de que el déficit aumenta la incidencia de DTN en los recién nacidos, sería lógico hallar un número menor de recién nacidos con estas malformaciones, por ejemplo 5 casos con DTN, en otro gru-

po de 10.000 mujeres con concentraciones normales de ácido fólico (tabla 1).

Obtendríamos un riesgo relativo de 3, que se interpretaría afirmando que el déficit de ácido fólico multiplica por tres el riesgo de que un recién nacido nazca con algún defecto del tubo neural. Sin embargo, no podemos descartar que este segundo grupo de mujeres, con valores normales de ácido fólico, presente, además, ciertas características saludables, como una mejor dieta en general, una mejor dotación genética o una baja prevalencia de factores de riesgo como el tabaco o el alcohol. Por ello, sería aceptable concluir que la menor incidencia de DTN en dichas mujeres puede deberse a dos fenómenos diferentes: a los valores normales de ácido fólico, pero también a los hábitos y condiciones más saludables, con lo cual el riesgo relativo de 3 sería una sobrestimación del efecto pernicioso de las concentraciones bajas de ácido fólico en mujeres embarazadas.

Intuitivamente parece lógico que el procedimiento más perfecto para determinar el efecto del déficit en ácido fólico consistiría en comparar a las primeras 10.000 mujeres que presentaban valores insuficientes de ácido fólico con ellas mismas pero suponiendo que tuvieran concentraciones normales de ácido fólico. En este caso, ambos grupos sólo se distinguirían en su exposición al ácido fólico y la medida de asociación sería realmente atribuible al déficit de ácido fólico. Sin embargo, dicho grupo de comparación no es posible en la práctica, no es «factible», va en contra de los hechos (cada mujer tiene o no un valor adecuado de ácido fólico), y por ello ha sido llamado «grupo contrafáctico». De existir una asociación entre el déficit de ácido fólico y los DTN, obtendríamos probablemente una cifra menor de 15 en estas mujeres «contrafácticas» con concentraciones normales de ácido fólico, pero una frecuencia mayor que la cifra hipotética de 5 que presentábamos anteriormente, obtendríamos, por ejemplo, la cifra hipotética de 10 (tabla 2).

TABLA 1

Asociación entre el déficit de ácido fólico y defectos del tubo neural

Ácido fólico	Defectos del tubo neural	Total
Déficit de ácido fólico	15	10.000
Ácido fólico normal	5	10.000

$$\text{Riesgo relativo} = \frac{15/10.000}{5/10.000} = 3.$$

TABLA 2

Asociación entre el déficit de ácido fólico y defectos del tubo neural. Grupo de comparación contrafáctico

Ácido fólico	Defectos del tubo neural	Total
Déficit de ácido fólico	15	10.000
Las mismas mujeres pero con valores normales de ácido fólico	10	10.000

$$\text{Riesgo relativo} = \frac{15/10.000}{10/10.000} = 1,5.$$

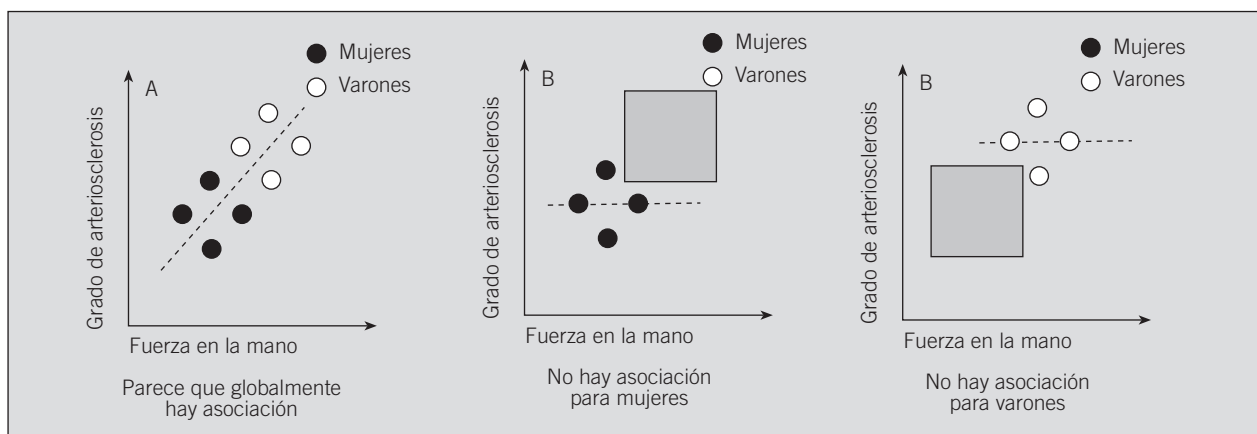
Palabras clave: Ajuste estadístico. Confusión. Epidemiología. Método epidemiológico. Sesgo.

Key words: Bias. Confounding. Epidemiology. Methods. Statistical adjustment.

Correspondencia: Prof. J. de Irala.
Unidad de Epidemiología y Salud Pública.
Facultad de Medicina. Universidad de Navarra.
Irunlarrea, 1. 31008 Pamplona. Navarra.
Correo electrónico: jdeirala@unav.es

Recibido el 14-2-2001; aceptado para su publicación el 21-3-2001

Med Clin (Barc) 2001; 117: 377-385



En el ejemplo de la tabla 2, la medida de asociación sería menor (riesgo relativo [RR] = 1,5). Podemos concluir, por lo tanto, que siempre que estimamos una medida de asociación (RR = 3 de la tabla 1) esta medida es el resultado de la conjunción de dos valores: la medida del efecto real (RR = 1,5 de la tabla 2) y un valor que llamamos «confusión», que se debe al hecho de utilizar un grupo de comparación que difiere del grupo contrafáctico ideal (este valor correspondería a 2 en nuestro caso). De forma resumida:

$$\text{Asociación (3)} = \text{efecto (1,5)} \times \text{confusión (2)}$$

En la medida en que consigamos grupos de comparación similares al contrafáctico ideal, obtendremos medidas de asociación más cercanas a la medida del efecto real que le corresponde a la exposición que se estudia, porque el valor de «confusión» será menor. Esto es lo que se pretende con el método epidemiológico.

En consecuencia, una variable (o factor) de confusión es una variable que distorsiona la medida de la asociación entre otras dos variables. El resultado de la presencia de una variable de confusión puede ser la observación de un efecto donde en realidad no existe o la exageración de una asociación real (confusión positiva) o, por el contrario, la atenuación de una asociación real e incluso una inversión del sentido de una asociación real (confusión negativa).

En el ejemplo ficticio presentado en el panel A de la figura 1, parece evidente que la fuerza de la mano está asociada al grado de arteriosclerosis, porque los 8 círculos presentan un orden claramente ascendente, sistemático y monótono. Al aumentar la fuerza de la mano, aumenta el grado de arteriosclerosis.

Sin embargo, cuando observamos los círculos correspondientes a cada sexo por separado sin modificar los dibujos (paneles B y C de la fig. 1), llama la atención que dicha asociación desaparece (no hay asociación entre la fuerza de la mano y el grado de arteriosclerosis), y se deduce que la aparente asociación que observábamos en el primer panel solamente existe porque hemos incluido ambos sexos en un mismo análisis. Por tanto, las diferencias entre ambos sexos surgen como la única explicación de la asociación observada inicialmente entre la fuerza de la mano y el grado de arteriosclerosis. Esta asociación acaba considerándose errónea y se dice que está confundida por la variable sexo.

Mientras que el sesgo es típicamente un fenómeno ligado a la selección de los sujetos o a la obtención de la información sobre la exposición o la enfermedad, el factor de confusión es, con mayor frecuencia, el resultado de relaciones específicas existentes entre las variables de una base de da-

tos. El factor de confusión está con frecuencia presente a pesar de los esfuerzos que se hacen para evitarlo. Una misma variable (p. ej., el sexo) puede ser un factor de confusión en una base de datos (de un estudio determinado) y no serlo en otra^{2,4}.

Sigamos explorando el fenómeno de confusión con la adaptación de datos clásicos de un análisis propuesto por Norell⁵ en el estudio de la asociación entre el consumo moderado de alcohol y el infarto de miocardio. En este ejemplo, prescindiremos de los intervalos de confianza para facilitar la exposición. En la tabla 3 se presenta la razón de tasas (RT) asociada al consumo moderado de alcohol y obtenemos un valor mayor que el valor nulo (valor nulo de la RT = 1), que indica que el consumo moderado de alcohol estaría asociado a una mayor tasa de infarto de miocardio.

Sin embargo, el consumo de alcohol está frecuentemente asociado al tabaquismo, y no sería aventurado aseverar que el tabaco podría ejercer un efecto de confusión de la asociación protectora entre el consumo moderado de alcohol y el infarto de miocardio. Si exploramos los datos anteriores, separando a los fumadores de los no fumadores, obtendríamos las tablas tetracóricas de la tabla 4.

De los datos presentados en la tabla 4 podemos deducir los siguientes hechos:

- El consumo moderado de alcohol protege del infarto de miocardio en personas fumadoras (estrato 1: RT = 0,80; valor inferior al nulo).
- El consumo moderado de alcohol protege del infarto de miocardio en personas no fumadoras (estrato 2: RT = 0,80; valor inferior al nulo).
- Ambos resultados son idénticos.
- Sin embargo, cuando observamos el efecto global del consumo moderado de alcohol (combinando los datos de ambos grupos de hábito tabáquico), nos sorprende que el efecto protector de este consumo de alcohol no sólo desaparece, sino que cambia de dirección, y encontramos una RT mayor que el valor nulo.

TABLA 3

Datos adaptados sobre la asociación entre el consumo moderado de alcohol y el infarto de miocardio⁵

Consumo moderado de alcohol	Infarto de miocardio (n)	Personas/año	Tasas (por 1.000)	Razón de tasas
No	42	20.400	2,1	1 (referencia)
Sí	187	39.600	4,7	2,2

TABLA 4

Razones de tasas de infarto de miocardio asociadas al consumo moderado de alcohol, por grupos de tabaquismo

Estimación global		
Todos	Consumo moderado de alcohol	
Infarto de miocardio	Sí	No
Casos	187	42
Personas/año	39.600	20.400
Tasas (por 1.000)	4,7	2,1

$$\text{Razón de tasas} = \frac{4,7}{2,1} = 2,2.$$

Estrato 1		
Fumadores	Consumo moderado de alcohol	
Infarto de miocardio	Sí	No
Casos	173	24
Personas/año	21.600	2.400
Tasas (por 1.000)	8	10

$$\text{Razón de tasas} = \frac{8}{10} = 0,8.$$

Estrato 2		
No fumadores	Consumo moderado de alcohol	
Infarto de miocardio	Sí	No
Casos	14	18
Personas/año	18.000	18.000
Tasas (por 1.000)	0,8	1

$$\text{Razón de tasas} = \frac{0,8}{1} = 0,8.$$

Todo esto se debe a que el tabaco actúa como un factor de confusión.

La diferencia entre el valor global y las estimaciones en cada estrato de hábito tabáquico se debe a la conjunción de los siguientes hechos:

1. En primer lugar, el tabaco está asociado a una mayor tasa de infarto de miocardio independientemente del consumo de alcohol o, dicho de otra forma, tanto en personas con o sin consumo moderado de alcohol. Efectivamente, se puede comprobar, al comparar las tasas de infarto de miocardio entre sujetos fumadores y aquellos que no fuman, que las razones de tasas son mayores que el valor nulo en ambos grupos de consumo de alcohol ($RT = 8 \div 0,8 = 10$ y $RT = 10 \div 1 = 10$ en los grupos con y sin consumo moderado de alcohol, respectivamente), y esto constituye la primera condición para que esta variable pueda distorsionar el efecto entre consumo de alcohol e infarto de miocardio; en otras palabras, para que pueda ser una variable de confusión.

2. Por otra parte, el hábito tabáquico no se distribuye de manera homogénea en ambos grupos de consumo de alcohol, ya que observamos que el 55% ($21.600 \div 39.600$) de las unidades de personas-año corresponden a personas con un consumo moderado de alcohol que son fumadores, mientras que solamente el 12% ($2.400 \div 20.400$) de las unidades de personas-año de observación corresponden a personas sin consumo moderado de alcohol que son fumadores. Es la segunda condición para que una variable pueda considerarse un factor de confusión.

El resultado de estos dos hechos es que, al estimar la razón de tasas global del efecto del consumo moderado de alcohol sobre la incidencia del infarto de miocardio, el efecto protector de este consumo moderado de alcohol se ve en cierto modo contrarrestado por el mayor riesgo asociado al tabaco, ya que existe una mayor prevalencia del hábito tabáquico en los sujetos que consumen alcohol.

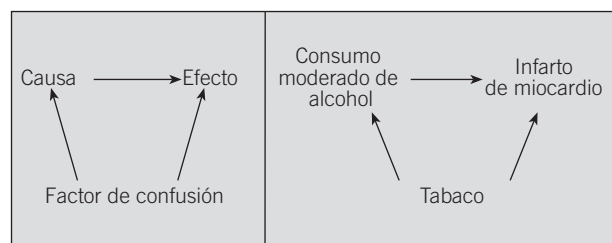


Fig. 2. Gráficos acíclicos dirigidos (DAG).

3. La tercera condición que se debe producir para que tenga lugar el fenómeno de confusión es que la variable de confusión no sea un eslabón intermedio mediante el cual actúa una exposición para producir el desenlace. Ciertamente no es plausible que el consumo moderado de alcohol ejerza su papel protector del infarto de miocardio a través de la modificación del hábito tabáquico de los sujetos del estudio.

El resultado global del primer panel de la tabla 4 es una razón de tasas llamada «bruta» o «cruda», porque no se tiene en cuenta el hábito tabáquico de los sujetos a la hora de estimarse y porque no es posible asegurar que su valor no esté distorsionado por el hábito tabáquico.

Por el contrario, las estimaciones de las razones de tasas obtenidas en cada estrato de tabaco se llaman estimaciones «ajustadas» o «controladas por el hábito tabáquico». Las diferencias en las tasas, entre los grupos de exposición, no pueden ser atribuibles a diferencias en el hábito tabáquico entre los sujetos expuestos y los no expuestos, porque cada estimación de razón de tasas se ha realizado en grupos de sujetos que son todos ellos fumadores o no fumadores. Se puede afirmar que las medidas de asociación han sido estimadas «a igualdad de tabaquismo» o eliminando el efecto distorsionante que pudiera tener el hábito tabáquico. El ajuste por una variable de confusión consiste, por tanto, en la eliminación del efecto distorsionante de dicho factor de confusión.

Este ejemplo facilita la comprensión de la definición de una variable de confusión. Para que una variable sea un factor de confusión de la asociación entre una exposición o factor de riesgo y una respuesta o desenlace, tiene que cumplir las siguientes condiciones²:

- Estar asociada con el desenlace, independientemente de su asociación con la exposición de interés (es decir, en sujetos no expuestos, estar también asociada con el desenlace).
- Estar asociada con la exposición pero no ser un resultado de la misma.
- No ser un eslabón causal intermedio entre la exposición y el desenlace como lo serían, por ejemplo, los valores de colesterol HDL entre una causa como la ingestión de alcohol y su efecto correspondiente, la menor incidencia de infarto de miocardio (el consumo de alcohol produciría un determinado incremento de colesterol HDL que, a su vez, reduce el riesgo de infarto de miocardio).

La relación entre una variable de confusión y las otras dos variables cuya medida de asociación distorsiona se suele representar por el esquema que presentamos a continuación y que autores como Greenland et al⁶ denominan «gráficos acíclicos dirigidos» porque no forman un ciclo cerrado y las variables están unidas por flechas dirigidas (*directed acyclic graphs DAG*; fig. 2, panel de la izquierda).

En el panel de la derecha introducimos las variables del primer ejemplo sobre alcohol e infarto de miocardio donde corresponde (fig. 2).

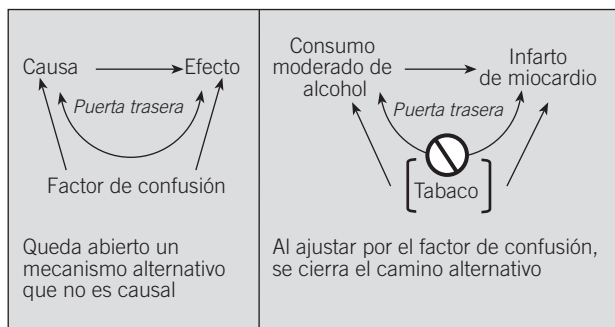


Fig. 3. Concepto de puerta trasera.

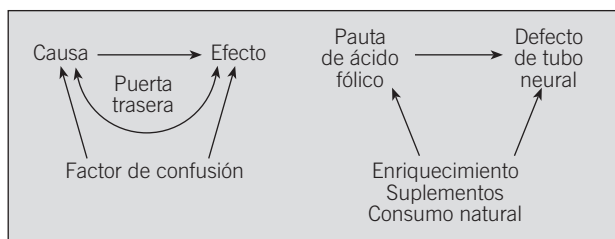


Fig. 4. Factores de confusión en la valoración del efecto protector del ácido fólico.

- El riesgo de un infarto aumenta con el hábito tabáquico (el tabaco está asociado con el riesgo de infarto tanto en consumidores moderados de alcohol como en los que no lo consumen).
- La proporción de fumadores entre quienes consumen alcohol es mayor que en los no consumidores (asociación entre consumo de alcohol y hábito tabáquico).
- Por último, no es plausible que el efecto beneficioso del consumo moderado de alcohol se realice a través del hábito tabáquico.

Recientemente se ha introducido el término de «puerta trasera» para referirse al camino abierto en redes causales, como la esquematizada para el ejemplo del consumo moderado de alcohol y el infarto de miocardio, donde una asociación puede ponerse de manifiesto por otras vías alternativas a la que estamos examinando (panel de la izquierda, fig. 3)⁶. Cuando se ajusta por una variable de confusión, se «cierra» esta puerta trasera, con lo cual se evita la distorsión de la asociación real (o, en su caso, la ausencia de asociación) entre la exposición y el desenlace (panel de la derecha, fig. 3). Cerrar la puerta trasera supone eliminar la confusión.

En muchos estudios epidemiológicos, la edad y el sexo son variables típicamente descritas como factores de confusión, hasta el punto de que son pocos los trabajos que no presentan datos ajustados por edad o por sexo.

En otro trabajo reciente sobre el tabaquismo y las fracturas de cadera, Hoidrup et al⁷ hallaban, en varones fumadores de más de 15 cigarrillos al día, en comparación con varones que nunca fumaron, una razón de tasas de fracturas de cadera de 1,68, que aumentaba hasta 2,2 una vez que se ajustaba por la edad. La diferencia entre la razón de tasas bruta y la ajustada refleja el hecho de que, en el análisis bruto, el efecto nocivo del tabaco quedaba en cierta medida enmascarado por el efecto beneficioso de la menor edad de los varones fumadores, en comparación con los que nunca fumaron.

A diferencia de la edad o el sexo, los factores de confusión pueden ser muy específicos de los estudios que se estén realizando. Por ejemplo, una vez que se ha difundido en la población que la toma de ácido fólico periconcepcional pre-

viene los defectos del tubo neural, y ante la evidencia de que muchos embarazos son imprevistos, se han propuesto diferentes medidas para mejorar la prevención primaria de estas malformaciones en mujeres en edades de reproducción. Estas medidas incluyen el enriquecimiento de alimentos básicos (cereales), el incremento de consumo de folatos a través de suplementos (p. ej., en forma de multivitaminas) o el aumento de consumo de alimentos ricos en folatos naturales (como las frutas y verduras)^{8,9}. Estos hechos tienen implicaciones en el caso de realizar estudios observacionales para valorar el efecto protector de alguna pauta determinada de ácido fólico, por ejemplo, porque tendremos que considerar cada una de estas medidas de prevención primaria como posibles variables de confusión (fig. 4).

Los alimentos pueden comportarse como factores de confusión de la asociación entre factores medioambientales y el cáncer de pulmón¹⁰⁻¹⁵. Koo y Ho¹² detectaron dicha confusión en un estudio de contaminación del aire por humos y cáncer de pulmón. Las mujeres expuestas a los humos de la cocina presentaban riesgos mayores de cáncer de pulmón. Sin embargo, aquellas con menor exposición a los humos de la cocina también presentaban hábitos dietéticos más beneficiosos con mayor frecuencia, con lo que se producía una posible exageración del efecto cancerígeno de la exposición al humo de la cocina.

Identificación

En términos generales, se habla de confusión cuando existen diferencias importantes entre las estimaciones brutas de una asociación y las ajustadas por los posibles factores de confusión¹⁶. Estas diferencias se pueden valorar siguiendo varios criterios, aunque existe un cierto consenso en la importancia de valorar el efecto que tiene el ajuste sobre la magnitud de los cambios de las medidas de asociación. De este modo, un factor puede considerarse de confusión cuando su ajuste es responsable de un cambio de al menos un 10% en la magnitud de la diferencia entre las estimaciones ajustadas y las brutas^{2,17,18}.

Antes de llevar a cabo estas comparaciones es preciso estimar los valores ajustados. El método más clásico para obtener valores ajustados de medidas de asociación es el que hemos presentado anteriormente, y que consiste en volver a calcular nuevas estimaciones dentro de cada estrato de la variable posiblemente de confusión. Es fácil comprender que, cuando queremos valorar varios factores de confusión simultáneamente (p. ej., la edad categorizada en dos grupos, el sexo y la ingestión de un alimento determinado como variable dicotómica), llegamos rápidamente a la situación donde nos faltan efectivos en los estratos para poder llegar a una estimación válida en cada estrato (obtendríamos 8 estratos con las variables que acabamos de enumerar) (fig. 5).

Las medidas de asociación que obtendríamos en la columna B de la figura anterior estarían ajustadas por edad e ingestión de alimento. Las medidas de la columna C estarían ajustadas por edad, por ingestión del alimento y por sexo, pero los sujetos del estudio se tienen que repartir entre 8 estratos y quedarán muy pocos (o ninguno) dentro de alguno de estos estratos.

Una opción más eficiente para poder considerar el papel de confusión de diversas variables simultáneamente es el análisis multivariante. El análisis multivariante es un procedimiento complejo que realizan más o menos automáticamente los programas estadísticos y que consiste en obtener, a partir de un número inicialmente importante de variables, el conjunto de variables (llamadas variables independientes, covariables o variables predictoras) que están más intensa-

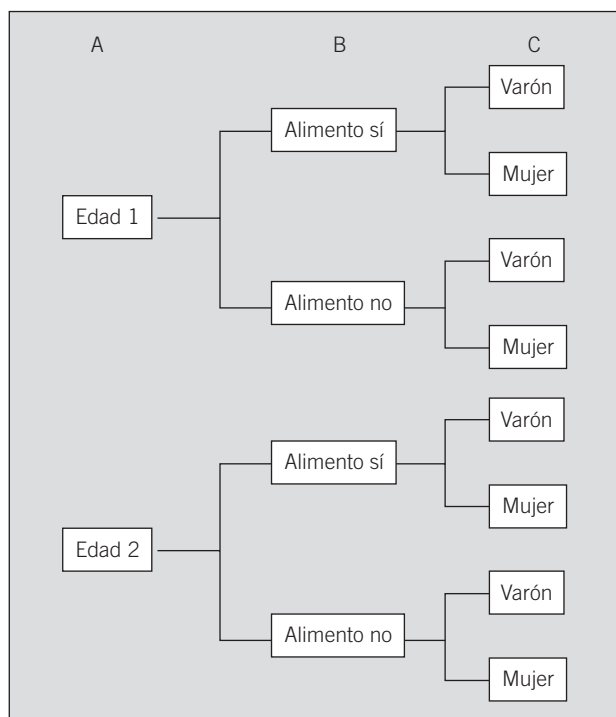


Fig. 5. Estratificación por grupos de edad, ingestión de alimento y sexo.

mente asociadas con el desenlace de interés (variable dependiente). Este conjunto de variables constituye lo que llamamos el «modelo estadístico multivariante». A partir de este modelo obtenemos las medidas de asociación de las diferentes variables que la componen, pero con la ventaja adicional de que cada una de estas estimaciones está ajustada por las demás variables que integran el modelo (fig. 6). Según la escala de la variable que cuantifica el desenlace, se utilizan diferentes tipos de modelos multivariantes: regresión lineal múltiple (desenlace cuantitativo), regresión logística (desenlace dicotómico) o regresión múltiple de Cox (función de supervivencia como desenlace de interés), regresión de Poisson (desenlace en forma de tasas)¹⁹.

La ventaja principal del análisis multivariante frente al análisis estratificado es que los modelos multivariantes son más eficientes. Es decir, a igualdad del tamaño muestral utilizado, se obtienen estimaciones más precisas y con un número mayor de variables que lo que sería admisible en un análisis estratificado¹⁹.

Cuando se estiman estos modelos para identificar variables de confusión, se recomienda escoger toda variable que, a la vez que cumple con los criterios generales de variables de confusión (criterios resumidos en los diagramas acíclicos), sea responsable de cambios de más del 10% entre medidas de asociación brutas (sin dicha variable en el modelo) y ajustadas (con dicha variable incluida en el modelo) y que presente un nivel de significación (valor de p) conservador, aproximadamente menor² de 0,20.

El sobreajuste

A veces, algunas variables pueden considerarse equivocadamente variables de confusión. El ajuste innecesario por estas variables se llama «sobreajuste», y puede tener como consecuencia precisamente aquello que pretendemos eliminar con el ajuste, es decir, la distorsión de la medida de asociación (p. ej., la eliminación o atenuación de una asociación real):

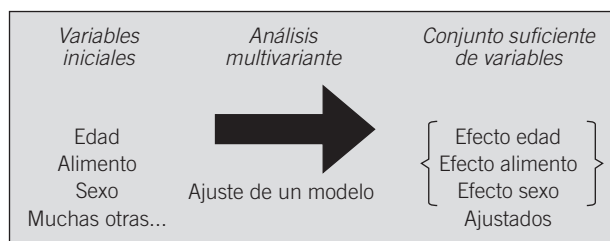


Fig. 6. Esquema del análisis multivariante.

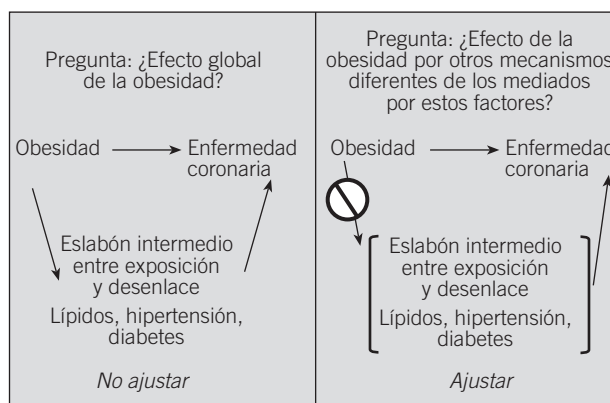


Fig. 7. Ajuste innecesario por variables del eslabón causal entre exposición y desenlace.

– Es posible que se ajuste inadvertidamente por variables que, en realidad, están en el camino causal entre una exposición y un desenlace. Por ejemplo, parte de la polémica respecto al efecto de la obesidad sobre la enfermedad coronaria se ha debido a la atenuación de las medidas de asociación entre estas variables en estudios que han ajustado por factores de riesgo como los lípidos en sangre, la diabetes y la hipertensión, que empeoran y/o son producidos por la obesidad³. El ajuste por los lípidos en sangre, la diabetes y la hipertensión, que no son más que eslabones intermedios, sería adecuado en el caso de querer valorar si la obesidad aumenta el riesgo de enfermedad coronaria por otros mecanismos diferentes de los mediados a través de estos factores de riesgo (fig. 7)²⁰.

– Algunas variables son marcadores de la exposición (pueden ser producidos por la exposición) y sólo están asociados al desenlace a través de la asociación que existe entre la exposición y dicho desenlace. Estos marcadores no estarían, por tanto, asociados con el desenlace independientemente de la exposición (esta asociación estaría ausente en sujetos no expuestos) y no cumplirían con una de las características de las variables de confusión (línea discontinua de la fig. 8). El control de este tipo de variables también daría lugar a un sobreajuste. En el ejemplo siguiente, presentado por Rothman y Greenland², si quisiéramos valorar el riesgo

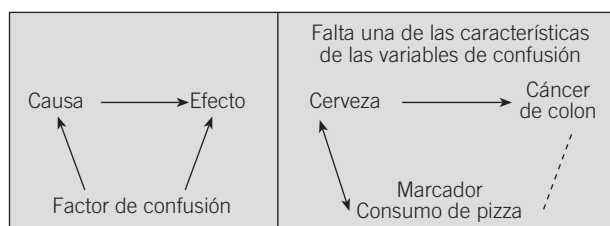


Fig. 8. Ajuste innecesario por variables marcadoras de la exposición. La flecha bidireccional indica una asociación no necesariamente causal.

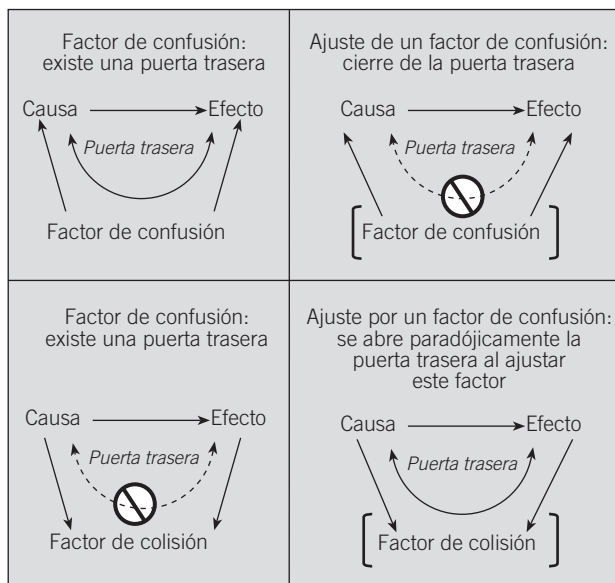


Fig. 9. Factor de colisión.

de cáncer de colon asociado al consumo de cerveza y existiese una asociación positiva entre ambos, encontraríamos probablemente una asociación estadística entre el consumo de pizza y el de cerveza (ambos se consumen simultáneamente con frecuencia) y también una asociación entre el consumo de pizza y el cáncer de colon. Sin embargo, la variable «consumo de pizzas» no estará asociada al cáncer de colon en personas que no consumen cerveza y no habría que ajustar por ella (fig. 8).

– Existen variables en las redes causales representadas por diagramas acíclicos dirigidos que se caracterizan por el hecho de que dos caminos confluyen en la misma variable (dos cabezas de flecha llegan a ella). Estas variables han sido denominadas *colliders* por Greenland et al⁶ y su uso en castellano ha sido infrecuente por ser un concepto reciente. Proponemos traducirlo por «factor de colisión» por su semejanza con el concepto de factor de confusión (recientemente se ha utilizado la terminología «factor colisionador»²¹) (fig. 9).

La relevancia de estos factores de colisión radica en el hecho de que se asemejan a factores de confusión por su asociación con la exposición y el desenlace, aunque no deben considerarse como tales porque no distorsionan la medida de asociación. Un factor de colisión bloquea el camino que pasa por él; sin embargo, si ajustáramos por un factor de colisión, paradójicamente causaríamos la apertura de una puerta trasera e introduciríamos un sesgo de confusión, allá donde no existía previamente⁶.

Un ejemplo que podría ayudarnos a entender el concepto de factor de colisión es el del uso de drogas por vía parenteral (UDVP), que no tiene por qué estar asociado a la homosexualidad (la proporción de usuarios de drogas no es necesariamente mayor en homosexuales en comparación con la población general), aunque dicha asociación podría ponerse de manifiesto en el caso de ajustar por infección por el virus de la inmunodeficiencia humana (VIH) (fig. 10).

La infección por el VIH sería un factor de colisión porque ambos colectivos tienen una mayor prevalencia de infección por el VIH en comparación con la población general. Si ajustáramos por el VIH (es decir, si valoráramos la asociación entre ser UDVP y homosexual mediante un análisis estratificado por VIH, separando los sujetos seropositivos de

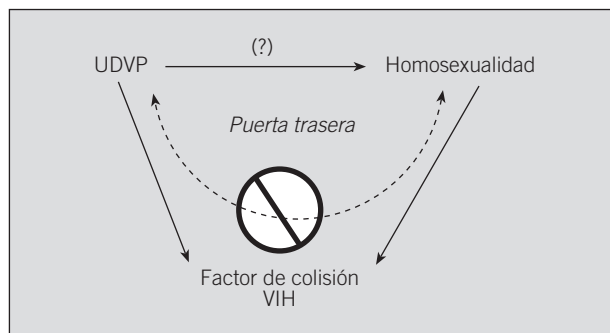


Fig. 10. Infección por el virus de la inmunodeficiencia humana (VIH): variable de colisión. UDVP: usuario de drogas por vía parenteral.

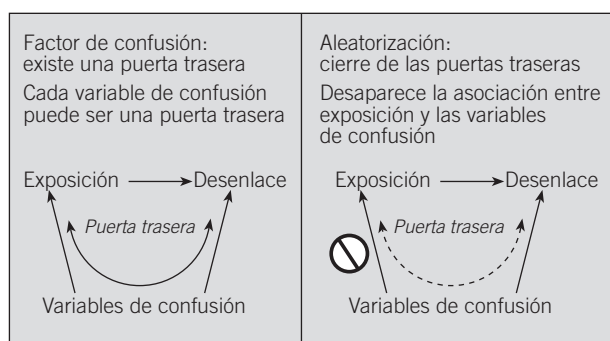


Fig. 11. Efecto de la aleatorización sobre la puerta trasera del diagrama causal acíclico.

los seronegativos), acabaríamos encontrando asociaciones entre la homosexualidad y el uso de drogas dentro de los dos estratos. Obsérvese que, en el análisis global (primer panel de la tabla 5), la prevalencia de UDVP es idéntica en ambos grupos (10% en homosexuales y no homosexuales).

Prevención y control del sesgo de confusión

Debido a que los factores de confusión introducen, por definición, un sesgo en las medidas de asociación, es evidente que se debe intentar prevenir y controlar dicho efecto antes de presentar los resultados definitivos de una investigación. Los factores de confusión pueden prevenirse en la fase de diseño o bien eliminarse en la fase de análisis de un estudio epidemiológico.

Fase de diseño

Se suelen utilizar tres variantes metodológicas para evitar el efecto de confusión de variables que tradicionalmente se comportan como tales.

TABLA 5

Ejemplo hipotético de variable de colisión. Usuarios de drogas por vía parenteral (UDVP), homosexualidad (HMSX) y positividad para el virus de la inmunodeficiencia humana (VIH)

Análisis global	Muestra total		VIH +		VIH -	
	HMSX	No HMSX	HMSX	No HMSX	HMSX	No HMSX
UDVP	10	90	8	60	2	30
No UDVP	90	810	60	10	30	800
Total	100	900	68	70	32	830
UDVP (%)	10	10	11,8	85,7	6,3	3,6

1. *Aleatorización*. Consiste en asignar al azar a los sujetos de un estudio, a los grupos que se pretenden comparar de modo que cualquier posible factor de confusión, conocido o desconocido, se puede considerar homogéneamente distribuido entre dichos grupos, con lo que desaparece una de las condiciones para que cause confusión (la asociación entre dicha variable y la exposición estudiada). Siguiendo con el concepto de la puerta trasera, el resultado de la aleatorización sería el bloqueo de la puerta trasera que pase por la variable aleatorizada (fig. 11). Este método es el utilizado en estudios clínicos o de intervención. La aleatorización consigue eliminar mejor la confusión cuando el tamaño muestral es grande. Sin embargo, nunca debe considerarse una garantía absoluta de ausencia de confusión, y esta posible explicación de los resultados debe discutirse pertinentemente a la hora de publicarlos. Por otra parte, la aleatorización no es siempre posible por problemas éticos²² o por el diseño observacional del estudio. Incluso cuando es posible, pueden surgir problemas, como la pérdida diferencial de individuos entre los grupos o la no adherencia al protocolo, que podrían introducir diferencias entre los grupos. Se ha intentado soslayar este problema mediante el análisis según «intención de tratar»²³, que considera a cada individuo en el grupo al que fue asignado aunque posteriormente no haya cumplido el tratamiento o exposición que se le asignó. Sin embargo, esto no evita la existencia de factores de confusión, ya que la adherencia puede ser diferencial dependiendo de otras variables. Por otra parte, desde Fisher²⁴ se ha considerado que teóricamente la aleatorización permite el control tanto de los factores conocidos como de los desconocidos o no medidos, pero ha existido polémica sobre la presencia o no de factores de confusión en estudios clínicos aleatorios, ya que la afirmación de Fisher se basaba en la ley de los grandes números, con grandes tamaños muestrales, por lo que un factor de confusión podría surgir por el azar al producirse desequilibrios entre los grupos durante la aleatorización.

Siguiendo con el concepto del modelo «contrafáctico» que observábamos con anterioridad, la aleatorización ayuda a obtener dos grupos de comparación que se asemejarían a los grupos del modelo contrafáctico porque se pretende que la única diferencia entre ambos sea la exposición estudiada. Cuando la aleatorización cumple su misión, puede decirse que ninguna cabeza de flecha de variables posiblemente de confusión apunta a dicha exposición, porque existe una distribución homogénea de estas variables en los dos grupos comparados.

2. *Restricción*. Consiste en admitir, en el estudio, sólo a aquellos sujetos que estén en una de las categorías del factor de confusión (p. ej., todos serán no fumadores para prevenir la confusión por tabaco), y con ello se consigue romper también la distribución heterogénea del factor de confusión entre los grupos comparados. La restricción no permite estudiar el efecto de la variable restringida. Es posible utilizar la restricción en estudios de cohortes, casos y controles y estudios experimentales. La restricción tiene varios inconvenientes desde el punto de vista práctico al reducir el número de individuos en el estudio, lo que implica una pérdida de potencia. También puede ser importante la amenaza a la validez externa, ya que puede ser difícil la generalización de los resultados.

3. *Emparejamiento* (matching). Es una modalidad de restricción que sólo se utiliza para ajustar por factores de confusión en estudios experimentales o de cohortes. Cada sujeto del grupo de expuestos se empareja con uno o más sujetos del grupo de no expuestos que se encuentran en la misma categoría del factor de confusión. Por ejemplo, ante

un sujeto fumador que consuma alcohol de manera moderada, buscaremos a otro también fumador pero que no consuma alcohol. Cualquier diferencia entre los grupos de consumo no podrá, en principio, atribuirse al tabaco, y diremos que la asociación observada está «ajustada por el tabaco». El problema, tanto de la restricción como del emparejamiento, es que no pueden controlarse los factores de confusión desconocidos o no medidos.

Fase de análisis

En la fase de análisis existen dos metodologías fundamentales de ajuste:

1. *La estratificación*. Consiste en realizar estimaciones de las medidas de asociación en cada subgrupo de la variable de confusión. Es precisamente lo que hemos hecho con el ejemplo del consumo moderado de alcohol y el infarto de miocardio (tabla 4). Las estimaciones ajustadas son aquellas que se consiguen dentro de cada estrato (categoría) de la variable de confusión.

2. *Análisis multivariante*. Cualquier estimación que produzca un modelo multivariante se puede considerar ajustada por las demás variables que constituyen el modelo. Para ajustar por una variable de confusión utilizando modelos multivariantes, basta con introducir el factor de confusión en el modelo.

En el ajuste multivariante, es preciso llegar a un equilibrio entre la necesidad de ajustar por muchas variables, introduciendo el máximo número de variables en el modelo, con la necesidad de obtener estimaciones precisas (se pierde precisión cuantas más variables se introducen en un modelo multivariante). Autores como Kleinbaum et al²⁵ abogan por la búsqueda del «grupo de confusión suficiente», que definen como el grupo mínimo de variables que pueden eliminar sustancialmente la confusión a través de su inclusión en los modelos, pero con la menor pérdida de precisión de las estimaciones obtenidas por este modelo. La palabra «mínima» se refiere a que ninguna variable se podría eliminar del análisis sin que ello afectara a la validez de las estimaciones.

Existen otros métodos para ajustar por factores de confusión, aunque su explicación detallada excedería al ámbito de esta revisión. Algunos son clásicos, como las tasas de mortalidad ajustadas o estandarizadas por edad, y se utilizan para comparar la mortalidad entre poblaciones ajustando por diferencias en las distribuciones de edades². Otros son más recientes, como el algoritmo de cálculo generalizado (*G-estimation*)²⁶ o la ponderación inversa al tratamiento y los modelos estructurales marginales (*marginal structural models*)²⁷.

Los diagramas acíclicos dirigidos pueden ser más complejos que lo visto anteriormente²⁸.

Los DAG ayudan a comprender también el concepto de factores de confusión no medidos (se representan como NM) como se recoge en la figura 12. Los factores de confusión que sí se han medido se representan por C. La exposición de interés cuyo efecto se intenta valorar se representa como E y el desenlace como D.

En la figura 13 se representa un ejemplo parcial de uno de estos diagramas adaptado al estudio de tratamientos con antihistamínicos y el asma⁶.

Este ejemplo fue presentado a un grupo de estudiantes y profesores a los que se preguntó: «¿Cuál es el mínimo conjunto de variables que se deben controlar al estudiar el efecto de los antihistamínicos sobre la aparición de asma?». La respuesta más frecuente fue que bastaría controlar por reactividad bronquial, porque parece que así se bloquean todos

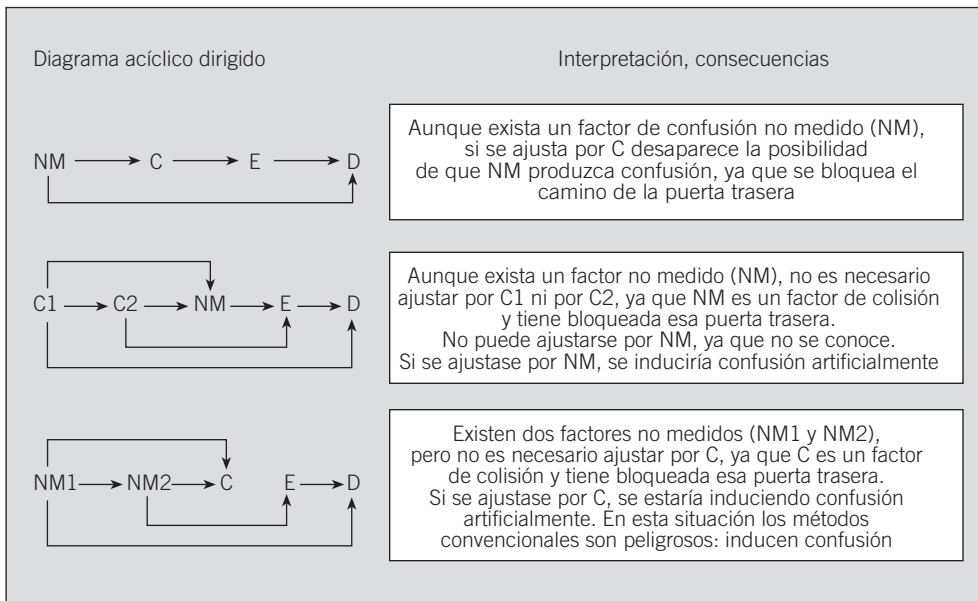


Fig. 12. Varios ejemplos de situaciones de investigación representadas mediante diagramas acíclicos dirigidos.

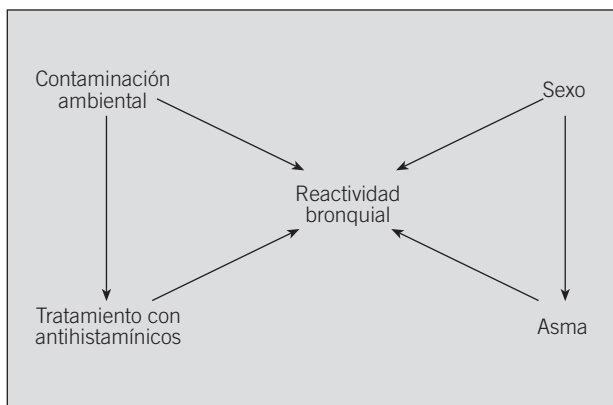


Fig. 13. Diagrama acíclico dirigido. Tratamiento con antihistamínicos y asma. (Adaptado de Greenland et al¹⁶.)

Presentación de medidas de asociación en presencia de confusión

Teniendo en cuenta que una estimación adecuadamente ajustada es más válida que otra bruta, se deben presentar siempre las estimaciones ajustadas. Cuando las estimaciones en cada estrato de la variable de confusión difieren ligeramente entre sí, es preciso combinarlas en una estimación global (ajustada), y a tal efecto se procede a calcular la media ponderada de las estimaciones en cada estrato para llegar a una cifra única que represente la medida de asociación global. El estimador de Mantel y Haenszel realiza, mediante un promedio ponderado, este tipo de estimación global². Por ejemplo, si tuviéramos que hallar una *odds ratio* ajustada global a partir de varias *odds ratios* que provienen de un número *k* de estratos de una variable de confusión en un estudio de casos y controles, utilizaríamos el procedimiento siguiente:

$$OR_{MH} = \frac{\sum_{i=1}^k a_i d_i / n_i}{\sum_{i=1}^k b_i c_i / n_i}$$

donde *a* y *b*, *c* y *d* corresponden a las celdas de las tablas tetracóricas que se construyen en los *k* estratos de la variable de confusión y *n_i* los efectivos totales en cada estrato *i*. Es importante detallar en el texto, o a pie de tabla en su caso, todas las variables por las cuales están ajustadas las estimaciones y cómo se codificó cada una de ellas. Sin embargo, tampoco es incorrecto presentar las estimaciones brutas junto con las ajustadas, porque permite al lector valorar los cambios existentes entre los valores brutos y los ajustados. En este caso, sin embargo, las conclusiones de nuestro trabajo siempre tienen que referirse a los valores ajustados por ser éstos los que están más libres de sesgos.

REFERENCIAS BIBLIOGRÁFICAS

1. Mill JS. Of the plurality of causes and the intermixture of effects. En: Robson JM, McRae RF, editores. A system of logic ratiocinative and inductive. Collected Works of John Stuart Mill. Toronto: Toronto University Press, 1974; 434-453.

los caminos de puerta trasera entre antihistamínicos y asma. Sin embargo, ésta es una solución ingenua y equivocada, ya que al controlar por reactividad bronquial se abre un camino entre antihistamínicos y asma en el que la reactividad bronquial es un «factor de colisión». Este camino se compondría de los siguientes eslabones: antihistamínicos-contaminación-reactividad bronquial-sexo-asma. Es importante tener en cuenta que la reactividad bronquial sólo es factor de colisión en este camino, pero no en otros. Por ejemplo, en el camino antihistamínicos-contaminación-reactividad bronquial-asma no es factor de colisión. Al controlar por reactividad bronquial se cierran algunos caminos, pero se abre el del factor de colisión. Lo correcto es controlar simultáneamente por contaminación y reactividad bronquial, o bien controlar simultáneamente por sexo y reactividad bronquial. Por otra parte, independientemente de la adecuada selección de las variables que trataremos como de confusión, es importante medirlas bien, porque su ajuste estadístico nunca elimina los cambios de la estimación atribuibles a errores de medición, y nos encontraríamos con medidas de asociación distorsionadas a pesar del ajuste estadístico, fenómeno que lleva el nombre de «confusión residual»^{29,30}.

2. Rothman KJ, Greenland S. *Modern Epidemiology*. Washington: Lippincott-Raven, 1998.
3. Friedman GD. *Primer of epidemiology* (4.ª ed.). Nueva York: McGraw-Hill, 1994.
4. Last JM. *A dictionary of epidemiology*. Nueva York: Oxford University Press, 1995.
5. Norell SE. *Workbook of epidemiology*. Nueva York: Oxford University Press, 1995.
6. Greenland S, Pearl J, Robins JM. Causal Diagrams for epidemiologic research. *Epidemiology* 1999; 10: 37-48.
7. Hoidrup S, Prescott E, Sorensen TIA, Gottschau A, Lauritzen JB, Schroll M et al. Tobacco smoking and risk of hip fracture in men and women. *Int J Epidemiol* 2000; 29: 253-259.
8. Berry RJ, Li Z, Erickson JD, Li S, Moore CA, Wang H et al. Prevention of neural-tube defects with folic acid in China. *N Engl J Med* 1999; 341: 1485-1490.
9. Botto LD, Moore CA, Khoury MJ, Erickson JD. Neural-tube defects. *N Engl J Med* 1999; 341: 1509-1519.
10. Axelsson G, Liljeqvist T, Andersson L, Bergman B, Rylander R. Dietary factors and lung cancer among men in west Sweden. *Int J Epidemiol* 1996; 25: 32-39.
11. Nyberg F, Agrenius V, Suartengren K, Svensson C, Pershagen G. Dietary factors and risk of lung cancer in never-smokers. *Int J Cancer* 1998; 78: 430-436.
12. Koo LC, Ho JH. Diet as a confounder of the association between air pollution and female lung cancer: Hong Kong studies on exposures to environmental tobacco smoke, incense, and cooking fumes as examples. *Lung Cancer* 1996; 14: S47-S61.
13. Ascherio A, Rimm EB, Hernan MA, Giovannucci E, Kawachi I, Stampfer MJ et al. Relation of consumption of vitamin E, vitamin C and carotenoids to risk for stroke among men in the United States. *Ann Intern Med* 1999; 130: 963-970.
14. The Heart Outcomes Prevention Evaluation Study Investigators (HOPES). Vitamin E supplementation and cardiovascular events in high-risk patients. *N Engl J Med* 2000; 342: 154-160.
15. Lichtenstein AH. Soy protein, isoflavones and cardiovascular disease risk. *J Nutr* 1998; 128: 1589-1592.
16. De Irala J, Fernández-Crehuet R. *Apuntes de epidemiología general: conceptos teóricos y ejercicios prácticos*. Pamplona: Newbook Ediciones, 1999.
17. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol* 1993; 138: 923-936.
18. Sonis J. A closer look at confounding. *Fam Med* 1998; 30: 584-588.
19. De Irala Estévez J, Díaz Molina C, Fernández-Crehuet Navajas R. *Regresión logística aplicada a las ciencias de la salud*. Pamplona: Newbook Ediciones, 1999.
20. Hu FB, Stampfer MJ, Rimm EB, Manson JAE, Ascherio A, Colditz GA et al. A prospective study of egg consumption and risk of cardiovascular disease in men and women. *JAMA* 1999; 281: 1387-1394.
21. Delgado M, Llorca Díaz, J. Sesgos. En: Galvez Vargas R, Sierra López A, Sáenz González MC, Gómez López LI, Fernández-Crehuet Navajas R, Salleras Sanmartí L et al. *Piédrola Gil medicina preventiva y salud pública* (10.ª ed.). Barcelona: Masson, 2001.
22. Ruiz-Canela M, Martínez-González MA, Gómez Gracia E, Fernández-Crehuet Navajas J. Informed consent and approval by institutional review board in published clinical trials. *N Engl J Med* 1999; 340: 1114-1115. [Erratum *N Engl J Med* 1999; 341: 460.]
23. Ruiz-Canela M, Martínez-González MA, De Irala-Estévez J. Intention to treat analysis in randomised clinical trials. *Br Med J* 2000; 320: 1007.
24. Fisher RA. *The design of experiments*. Edimburgo, Oliver & Boyd, 1935.
25. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research*. Nueva York: Van Nostrand Reinhold, 1982.
26. Witteman JC, D'Agostino RB, Stijnen T, Kannel WB, Cobb JC, De Ridder MA et al. G-estimation of causal effects: isolated hypertension and cardiovascular death in the Framingham Heart Study. *Am J Epidemiol* 1998; 148: 390-401.
27. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in Epidemiology. *epidemiology* 2000; 11: 550-560.
28. Psaty BM, Koepsell TD, Lin D, Weiss NS, Siscovich DS, Rosendaal FR et al. Assessment and control for confounding by indication in observational studies. *J Am Geriatr Soc* 1999; 47: 749-754.
29. Becher H. The concept of residual confounding in regression models and some applications. *Stat Med* 1991; 11: 1747-1758.
30. Willett W. *Nutritional epidemiology*. Nueva York: Oxford University Press, 1998.