

Enfoques modernos del sesgo y la causalidad en la investigación epidemiológica

Modern approaches to bias and causation in epidemiological research

Dr. C. Jorge Bacallao Gallestey

Centro de Investigaciones y Referencia de Aterosclerosis de La Habana (CIRAH). La Habana, Cuba.

RESUMEN

Los conceptos de causalidad y sesgo están en la base de la investigación biomédica moderna, desde el análisis de cientos de factores de exposición, hasta los megaestudios para evaluar intervenciones. Los consumidores de estos productos de la investigación, vemos con desconcierto, que una conclusión que se formula hoy, se pone en duda mañana, y se desecha poco tiempo después, para eventualmente ser retomada en el futuro bajo otras ópticas u otros presupuestos. Aunque no es el único responsable, el sesgo metodológico juega un papel importante como determinante de esta realidad. Este artículo tiene el propósito de destacar el concepto de sesgo, relevante, entre otras posibles acepciones, para la investigación biomédica contemporánea, y su asociación con la definición técnica de "confusión", exponer la visión moderna sobre el significado práctico de una causa y examinar críticamente dos modernos recursos analíticos para afrontar el problema del sesgo y la causalidad: los puntajes de susceptibilidad y las variables instrumentales.

Palabras clave: sesgo, causalidad, confusor, puntajes de susceptibilidad, variables instrumentales.

ABSTRACT

The concepts of causation and bias are crucial to modern biomedical research, ranging from the analysis of hundreds of exposure factors to megatrials, in order to assess the

impact of interventions. As consumers of these research products, we are amazed that a statement made today is put into question tomorrow, discarded afterwards, and eventually retaken in the future from different perspectives or under different assumptions. Although the methodological bias is not the only culprit, it plays an important role as determinant of this reality. This paper intended to clarify the concept of bias, which is relevant, among other possible meanings, to contemporary biomedical research, and its association with the technical meaning of confounding. Other objectives were to present the current vision on the practical meaning of cause in epidemiological causal inference, and to critically review two modern analytical tools to deal with bias and causation such as propensity scores and instrumental variables.

Key words: bias, causation, confounders, propensity scores, instrumental variables.

INTRODUCCIÓN

En un artículo reciente,¹ su autor realizó la siguiente desconcertante afirmación:

Alarmanes grietas están penetrando profundamente en el edificio de la ciencia, y amenazan su estatus y su valor social. Las grietas no pueden atribuirse a los sospechosos habituales: escaso financiamiento, faltas de conducta en la investigación, interferencia política [...] La causa es el sesgo, y su amenaza apunta directamente al corazón de la ciencia.

Siete años antes, en un polémico artículo, *John Ioannidis*,² sostenía que la mayoría de los resultados de investigación publicados, son falsos, en gran medida, debido al sesgo.

El sesgo y la causalidad son dos conceptos clave en el examen crítico que se realiza en este texto. Ambos asoman permanentemente en la investigación epidemiológica, y determinan las bases de credibilidad en ejemplos como el que se inserta a continuación:

LAS PERSONAS DE BAJA ESTATURA TIENEN EL CORAZÓN MÁS FRÁGIL

Las personas de baja estatura tienen un riesgo una vez y media más elevado de desarrollar una enfermedad cardíaca y de morir prematuramente que las personas altas, según un estudio publicado sobre unos tres millones de individuos. Las mujeres que miden menos de 1,53 metros y los hombres de menos de 1,65 metros tienen una tendencia mucho mayor a sufrir de problemas cardiovasculares que las mujeres que superan los 1,66 metros y los hombres de más de 1,73 metros, según el estudio publicado en el *European Heart Journal*. La estatura pequeña debería ser agregada a la lista de los factores conocidos de riesgo cardíaco, como la obesidad y un colesterol elevado, según un estudio de la Universidad de Tampere, en Finlandia.

París, junio 13/2010 (AFP)

El sesgo es una amenaza constante para la investigación epidemiológica, que procura aislar relaciones de causalidad en sistemas complejos en los que las variables

relevantes nunca pueden ser completamente identificadas.¹ Aunque puede definirse formalmente fuera de un contexto de inferencia causal, el sesgo está intrínsecamente ligado a la causalidad.³

SOBRE LAS DEFINICIONES DE SESGO

En este texto, el término *sesgo* se refiere al error en la medición de un efecto causal. La abrumadora mayoría de las modalidades de sesgo asociadas a esta acepción, se deben a diseños o análisis inadecuados, y son formas de *sesgo metodológico*. No obstante, el término tiene otras acepciones.

Para un estadístico matemático, el sesgo es una propiedad de un estimador, que se define como la diferencia entre su valor esperado y el verdadero valor del parámetro que se desea estimar. Para un científico social, el sesgo alude a aspectos de la conducta de los seres humanos, que explican su comportamiento frente a estímulos específicos. Para otros es un prejuicio frente una idea o a un grupo humano.

El sesgo es un defecto estructural que no desaparece con grandes tamaños de muestras. La estadística, excelentemente equipada para el manejo de los sesgos aleatorios, poco puede hacer por sí misma frente a los sesgos sistemáticos.

El problema que centra el interés de este artículo es la posibilidad de alcanzar conclusiones válidas acerca del vínculo causal de un factor A, con una respuesta B, en el ámbito de la salud. ("¿Es cierto que la baja estatura incrementa el riesgo de enfermedad cardíaca y muerte prematura?"). Este es por supuesto, un propósito subsidiario de la práctica de un salubrismo eficiente, difícilmente concebible sin una "inteligencia" epidemiológica que desvele las causas de los problemas, guíe la focalización de las acciones y proporcione las bases científicas para orientar las estrategias y los programas de salud.

Sesgo y confusores

Si dos variables A y B están asociadas y no es debido al azar, la asociación se debe a que una es causa de la otra (el tabaquismo y el cáncer del pulmón), a que ambas comparten una o varias causas comunes (dedos amarillos y cáncer del pulmón) o a las dos razones anteriores. Si A precede a B, la asociación entre A y B tiene dos componentes: una debida al efecto causal de A sobre B y otra, a que comparten causas comunes. El propósito de buena parte de la investigación epidemiológica es estimar la primera componente. La segunda produce "confusión".

Un modo de eliminar una asociación espuria es estratificar para las causas comunes y luego estimar el efecto condicional, (por ejemplo, así desaparece la asociación entre los dedos amarillos y el cáncer de pulmón). Los confusores eliminan o reducen el componente espurio de una asociación, cuando se incluyen en el análisis como criterios de estratificación.

En muchos estudios epidemiológicos, el ingrediente del análisis está constituido por tres tipos de variables: la exposición (A: la baja estatura), el resultado (B: la enfermedad cardíaca o la muerte prematura) y los confusores. Los dos primeros están determinados por la interrogante causal. Los confusores, sin embargo, hay que identificarlos y luego, ajustar para ellos el efecto de A sobre B. Consideremos el ejemplo de la figura.

		C= 1		C= 0	
		B= 1	B= 0	B= 1	B= 0
A= 1		250	400	750	600
A= 0		150	200	150	100

		B= 1	B= 0
A= 1		1 000	1 000
A= 0		300	300

Fuente: El ejemplo ha sido adaptado de un trabajo que data de 1951 y que dio origen a la famosa "paradoja de Simpson". Simpson EH. The interpretation of interaction in contingency tables. J R Stat Soc Ser B. 1951;13:238-41.

Fig. Frecuencias absolutas en una tabla de contingencia múltiple que relaciona una exposición (A) con un efecto (B) y una variable de control (C).

Las tablas de la parte superior de la figura resultan de estratificar los pacientes según un factor C. Si se usa el "odds ratio" como medida de asociación, en las tablas con datos separados por estratos se tendrá que:

$$OR_{AB/C=1} = \frac{250 \times 200}{150 \times 400} = 0,833$$

$$OR_{AB/C=0} = \frac{750 \times 100}{150 \times 600} = 0,833$$

En la tabla con los datos unidos, se tendrá que:

$$OR_{AB} = \frac{1000 \times 300}{1000 \times 300} = 1$$

Consideremos ahora los escenarios siguientes, cada uno de los cuales pudo haber generado la distribución de los datos de la figura.

En el primero, A representa un tratamiento (A= 1 tratados, A= 0 no tratados); B mejoría clínica (B= 1 no mejoría, B= 0, mejoría) y C, la edad estratificada como variable binaria en torno a un punto de corte C_0 (C= 1, edad $\leq C_0$; C= 0, edad $> C_0$). En ambos grupos el tratamiento fue efectivo: la proporción de los que eran mejoría es más alta en los tratados. La asociación relevante es la condicional ($OR_{AB/C}$) ya que en ambos grupos es preferible tratar que no tratar. C es un confusor: con la edad se incrementa la necesidad de tratar y disminuye la probabilidad de mejoría; además, la asociación condicional para C es diferente de la asociación marginal.

En el segundo escenario A designa al mismo tratamiento del ejemplo 1, B representa la mejoría clínica objetiva según una prueba de laboratorio (B=1 no mejoría, B= 0 mejoría), y C, la mejoría clínica subjetiva según percepción del paciente. (C= 1, no percibe mejoría, C= 0, perciben mejoría). En este escenario, la asociación relevante es la marginal. La percepción de mejoría es una consecuencia común del tratamiento y de la mejoría objetiva. El tratamiento provoca una sensación de mejoría, pero no una mejoría objetiva de los indicadores biológicos; cuando estos mejoran, el paciente percibe mejoría subjetiva. C es un "colisor" (un efecto común de A y B) y no un confusor (una causa común). En general, se debe ajustar para un confusor, pero no para un colisor, pero la condición de un factor no puede inferirse de los datos, sino que depende de un conocimiento previo acerca de los vínculos entre las variables.

Aun cuando la asociación condicional es homogénea entre estratos de C (ausencia de interacción), la respuesta adecuada es en algunas ocasiones la asociación marginal, y en otras la condicional. Los datos empíricos y las asociaciones estadísticas deben complementarse con el conocimiento teórico para hacer inferencias causales válidas.

Ha habido notables tentativas de realizar inferencias causales a partir de pura evidencia observacional. *Spirtes, Glymour y Scheines*,⁴ y *Pearl y Verma*,⁵ realizan la sorprendente afirmación de que es posible inferir relaciones causales a partir de asociaciones observadas.

Sin embargo, *Robins y Wasserman*,⁶ mostraron que tal afirmación descansa en el supuesto de que la probabilidad de que no existan causas comunes para X y Y es positiva, y grande con respecto al tamaño muestral. La enorme cantidad de confusores posibles hace que sea extremadamente improbable que no existan causas comunes para X y Y, por lo que a partir solo de la evidencia observada, es imposible descartar la existencia de confusión. Inferir causalidad a partir de las asociaciones, no es una tarea confinada al terreno de la estadística, sino que requiere de un conocimiento teórico previo, para identificar confusores y minimizar la confusión residual. Posteriormente otros autores también refutaron la pretensión de convertir la identificación de confusores en una tarea analítica basada solo en asociaciones observadas.⁷⁻⁹

No obstante, en la práctica se continúa aplicando procedimientos basados en asociaciones estadísticas, mediante dos estrategias fundamentales: 1. La regresión paso a paso, bajo el supuesto de que aunque no todas las variables identificadas serán confusores, todos los confusores importantes serán identificados. 2. La comparación de los efectos no ajustados con los ajustados y si la diferencia relativa es mayor de 10 %, se elige la variable como confusor. Esta estrategia se basa en el supuesto de que toda variable cuyo ajuste provoque un cambio importante en la estimación del efecto causal, debe ser controlada.

Como lo han demostrado diversos autores, las dos estrategias pueden producir sesgo por varias razones.¹⁰⁻¹² La primera es la omisión de confusores no observados. La estimación de efectos causales en estudios observacionales, está sujeta no solo a la incertidumbre del componente aleatorio, sino también a la existencia de confusores ignorados.

La segunda es el ajuste inapropiado para variables no confusoras, que puede ilustrarse muy bien, a partir de los datos de la figura, si suponemos que han sido generados bajo el segundo de los escenarios.

a. Aplicar la estrategia 1, implicaría ajustar un modelo de regresión logística con B como variable (dicotómica) de respuesta, en el que se forzaría la presencia de A y se introduciría a C, como covariante.

En ausencia de interacción, ya que:

$$OR_{AB/C=1} = OR_{AB/C=0}$$

... el modelo correcto sería:

$$\log \text{it } p(B=1/A,C) = \beta_0 + \beta_A * A + \beta_C * C \quad (I)$$

La tabla 1 contiene los resultados de rutina del ajuste del modelo (I). En fuente *itálica* se destacan los valores de la significación, y la estimación del "odds ratio" con su intervalo de confianza.

Tabla 1. Ajuste del modelo I: estimación de los parámetros con intervalo de confianza (IC) al 95% para el odds ratio

	B	EE	Sig.	Exp (B)	IC al 95%	
					inferior	superior
C	- 0,693	0,084	0,000	0,500	0,424	0,590
A	- 0,182	0,097	0,061	0,833	0,689	1,008
Constante	0,405	0,097	0,000	1,500		

Según la estrategia 1, el valor de p y el intervalo de confianza en la tabla 1 sugerirían incluir a C como confusor. La estimación condicional del efecto de A sobre B, vendría dada por:

$$\exp(\beta_A) = 0,833$$

como se obtuvo a partir del cálculo del "odds ratio"estratificado.

b. La estrategia 2 también implica elegir C, ya que la diferencia entre la estimación con el modelo marginal:

$$(OR_{AB} = 1)$$

y el modelo condicional:

$$(OR_{AB/C} = 0,833)(I),$$

es de 16,7 %. Aunque las dos estrategias aconsejarían ajustar para C, dicho ajuste sesga la estimación del efecto causal de A sobre B.

Otros errores comunes asociados con el ajuste estadístico son: 1. Inferir que un factor es causal, si se asocia con el resultado aun después del control estadístico de otros factores. 2. Inferir que un factor no es causalmente importante porque su asociación con la variable de respuesta desaparece o se atenúa luego de la inclusión de covariantes en el proceso de ajuste, lo cual puede deberse a que las covariantes son mediadoras del efecto del factor de exposición.¹³

El enfoque de los resultados potenciales de Rubin

El concepto de causalidad está presente en todas las ramas de la ciencia.¹⁴ Algo similar ocurre en epidemiología, cuyo objeto principal según *Morabia*¹⁵ es la "etiología de la salud poblacional". Para *Swaen y Amelsvoort*,¹⁶ el principal objetivo de la investigación epidemiológica es identificar las causas de las enfermedades, y para *Botti y otros*,¹⁷ el tema central en Epidemiología Ambiental es la evaluación de la naturaleza causal de asociaciones empíricas entre la exposición a agentes ambientales y la ocurrencia de enfermedades. Según algunas perspectivas más moderadas como la de *Weed*,¹⁸ el propósito de la epidemiología no es probar relaciones causales sino profundizar en el conocimiento sobre los determinantes y la distribución de las enfermedades para aplicarlo a mejorar las condiciones de salud.

Durante la segunda mitad del pasado siglo, la aproximación a la causalidad en la investigación epidemiológica seguía las pautas de Sir *Austin Bradford Hill*. En un artículo seminal, *Hill* formuló nueve criterios prácticos, que posteriormente se utilizaron para identificar causalidad a partir de asociaciones observadas, pese a que su autor eludió cualquier referencia a una definición de causalidad.

No obstante, la interpretación más común es que su propuesta constituye un algoritmo de inferencia causal, por lo cual devino el recurso por excelencia para identificar causas, hasta la aparición de enfoques más modernos.¹⁹

Pese al impacto de las consideraciones de *Hill* sobre causalidad, la inferencia causal en epidemiología es mucho más un ejercicio de medición de un efecto, que un algoritmo para decidir si dicho efecto existe o no.

Más recientemente, *Rothman y Greenland*,²⁰ propusieron la noción de "causa suficiente" que definieron como un mecanismo causal completo, compuesto por un conjunto de condiciones y eventos, que inexorablemente producen un resultado.

La más notable implicación de este enfoque es la multicausalidad, ya que cualquier mecanismo causal completo involucra una gran cantidad de factores. La importancia de la multicausalidad radica en que la mayoría de las causas identificadas no son necesarias, ni suficientes para producir el resultado y sin embargo, su eliminación puede implicar una reducción sustancial en la carga poblacional de enfermedad. La comprensión de este hecho es crucial para la práctica de un salubrismo eficiente que debe basarse en buenas estimaciones *a priori* y en buenas evaluaciones *a posteriori* de los impactos de los programas y estrategias de salud.

El también llamado modelo contrafactual fue introducido por *Rubin*.^{21,22} También tiene antecedentes en la literatura económica en los trabajos de *Heckman*.²³ Aunque el modelo contrafactual se ha impuesto como filosofía de inferencia causal, la estimación de efectos presenta dificultades, especialmente en los estudios observacionales. Para definir un efecto causal en un sujeto *i*, supongamos que se desea evaluar el efecto de un factor de exposición *A*, con dos niveles posibles 1 y 0, sobre un resultado *Y*, que puede ser una variable binaria (muerte o sobrevivida) o cuantitativa (nivel de lesión aterosclerótica). La inferencia contrafactual se basa en el supuesto de que el resultado Y_i existe, tanto en caso de que $A=1$ ($Y_{i,1}$) como en caso de que $A=0$ ($Y_{i,0}$)

Para cada sujeto, solo es posible observar uno de los resultados potenciales. Si $A=1$, $Y_{i,0}$ es inobservable, y si $A=0$, lo es entonces $Y_{i,1}$. La condición a la que el sujeto *i* no ha estado expuesto, se denomina condición contrafactual, y el resultado bajo esta condición, resultado contrafactual. Antes de observar la condición o el tratamiento,

hay dos resultados posibles o potenciales. Luego de observar la condición, uno de los resultados es real y el otro contrafactual.

Se dice que *hay un efecto causal* para el sujeto i , si $Y_{i,1}$ es desigual a $Y_{i,0}$. El efecto causal se define como la diferencia $D_i = Y_{i,1} - Y_{i,0}$. Dado que en la estimación de un efecto causal, uno de los valores es siempre contrafactual, es imposible estimar dicho efecto a nivel individual, y en ello consiste el problema fundamental de la inferencia causal. La esencia de la estimación del efecto es que el sujeto de la medición, es idéntico a sí mismo, salvo por la circunstancia de la exposición.

El enfoque contrafactual hace totalmente explícita la definición del parámetro causal, lo cual se aprecia claramente cuando el resultado Y es una variable binaria. Por ejemplo, si D_i es desigual a 0, A causa o protege de Y en el sujeto i ; y si $D_i = 0$, A es irrelevante como fuente de causalidad.

Lo que resulta de interés para la investigación es el efecto causal promedio (ECP) que se define como:

$$ECP = E(D_i) = E(Y_{i,1}) - E(Y_{i,0}) \quad (II)$$

y cuyo estimador sería:

$$E\hat{C}P = \frac{1}{N} \sum_{i=1}^N D_i = \frac{1}{N} \sum_{i=1}^N Y_{i,1} - \frac{1}{N} \sum_{i=1}^N Y_{i,0} \quad (III)$$

El estimador (III) se basa en valores observados y no observados, que se incluyen en (III) como si fuesen conocidos.

En la práctica solo es posible usar el estimador basado en casos observados:

$$E\tilde{C}P = \frac{1}{N_1} \sum_{i=1}^{N_1} Y_{i,1} - \frac{1}{N_0} \sum_{i=1}^{N_0} Y_{i,0} \quad (IV)$$

en donde N_1 es el número de casos expuestos y N_0 el número de casos no expuestos al factor A .

Esta estimación "natural" se basa en supuestos, que casi nunca se cumplen en estudios observacionales, y no son verificables en los datos observados.

El supuesto básico es la intercambiabilidad,^{24,25} que se examina a continuación.

Para simplificar, supongamos que el resultado Y es una variable binaria. En ese caso:

$$\begin{aligned} E(Y_{i,1}) &= \Pr(Y_{i,1} = 1) \quad y \\ E(Y_{i,0}) &= \Pr(Y_{i,0} = 1) \quad (V) \end{aligned}$$

Según (II) el efecto causal es entonces:

$$\Pr(Y_{A=1} = 1) - \Pr(Y_{A=0} = 1) \quad (\text{VI})$$

y el efecto de asociación es:

$$\Pr[Y=1/A=1] - \Pr[Y=1/A=0] \quad (\text{VII})$$

(VI) se refiere a los resultados en todos los sujetos de la población, (VII) se refiere a los resultados reales en sujetos expuestos y no expuestos.

Bajo las condiciones de un experimento ideal en el cual:

- Los sujetos se asignan aleatoriamente a $A=1$ y $A=0$.
- Hay completa adhesión.
- No hay pérdidas por seguimiento.
- El experimento es a ciegas.

... pueden permutarse los casos $A=1$ y $A=0$, y es claro que:

$$\Pr[Y_a = 1/A=1] = \Pr[Y_a = 1/A=0] = \Pr[Y_a = 1] \forall a \quad (\text{VIII})$$

La expresión (VIII) define la condición de "intercambiabilidad" que responde al hecho de que antes de la asignación, los grupos son equivalentes en promedio para todas las posibles covariantes. La aleatorización produce grupos iguales en valor esperado, excepto por la exposición al tratamiento, por lo cual cualquier diferencia puede atribuirse al tratamiento y puede afirmarse que "asociación es causalidad."

En estudios experimentales sin asignación aleatoria, no puede garantizarse la intercambiabilidad debido al sesgo de selección que ocurre cuando la probabilidad marginal de ser asignado a uno de los dos tratamientos, no es igual a la probabilidad condicional al conjunto de las covariantes determinantes del resultado.^{26,27}

En estudios observacionales, tampoco puede garantizarse la intercambiabilidad debido a la existencia de confusores no observados.

En lo que resta se examinan brevemente, dos enfoques analíticos modernos, aplicables en estudios observacionales o *cuasi* experimentales.

Los puntajes de susceptibilidad‡

Los puntajes de susceptibilidad (PS) se usan cada vez más para reducir el sesgo de selección en la estimación de efectos causales.²⁸⁻³³ El PS se define como la probabilidad individual de pertenecer a un nivel dado del factor de exposición o tratamiento, condicional en las covariantes observadas. *Rosenbaum y Rubin* demostraron que, en ausencia de confusores no observados, condicionar para los puntajes de susceptibilidad permite obtener estimaciones no sesgadas del efecto de un tratamiento.^{28,29}

El PS calculado sobre un conjunto de covariantes relevantes, condensa en un escalar único la falta de balance entre niveles del factor de exposición. Si todo el efecto de confusión se concentra en dicho grupo de covariantes, el control del puntaje de susceptibilidad permite estimar efectos causales.²⁸

Obtenido el PS, hay diferentes alternativas para su aplicación. Una es utilizarlo como covariante y remover su efecto para calcular el efecto causal ajustado, como en el análisis de la covarianza clásico.^{32,33} Las otras alternativas son el apareamiento y la estratificación.

El apareamiento consiste en elegir, para cada sujeto tratado o expuesto, uno o más controles dentro de un conjunto cuyo PS se encuentre a una distancia menor que un umbral de tolerancia pequeño (ϵ) elegido *a priori* y su mayor inconveniente es que si hay un escaso solapamiento en la distribución de los PS entre los dos grupos, habrá pocos casos elegibles para el análisis.

Para cada sujeto expuesto i , el conjunto de sus controles posibles estaría definido por:

$$CCP_i = \{j / j \neq i; d(i, j) \leq \epsilon\} \quad (IX)$$

La estratificación se hace del mismo modo que con cualquier otro confusor relevante. Como el PS es una covariante continua, se construyen intervalos de clase y se practica luego el análisis estratificado usual, calculando estimadores ponderados. Si el PS modifica el efecto de la exposición, se estima el efecto de asociación para cada intervalo de clase.

Las variables instrumentales (vi)

Los métodos basados en VI se conocen desde los años 30 del pasado siglo, sin embargo, han tenido muy escasa aplicación en epidemiología.

Su racionalidad consiste en lo siguiente: supongamos que A y B son la exposición y el resultado cuya probable relación causal se investiga, y que podemos medir la asociación de ambas con una variable W (a la que llamaremos (VI) que se asocia con A, pero que solo tiene con B una asociación completamente mediada por A. Bajo las condiciones que se indican más adelante tendremos que:³⁴

$$ASOC_{ZB} = ASOC_{ZA} * ASOC_{AB} \quad (XI)$$

De modo que:

$$ASOC_{AB} = \frac{ASOC_{ZB}}{ASOC_{ZA}} \quad (XII)$$

La relación (XII) es útil cuando:

- La asociación entre A y B está confundida por confusores no observados.
- No es posible medir A directamente, y W es un marcador de A cuya asociación con A puede estimarse, y no depende de otras variables.

Si U es el conjunto de todos los confusores no observados de la relación entre A y B , las condiciones para el empleo de VI son las siguientes:

- W es independiente de U .
- W está asociada con A .
- W está asociada con B , pero dicha asociación está totalmente mediada por A .

Un caso trivial de VI es la asignación aleatoria sin completa adhesión, que da lugar al método de estimación que se conoce como "intención de tratamiento". La asignación aleatoria:

- Está libre de la influencia de cualquier confusor, por definición.
- Se asocia con el tratamiento.
- Su asociación con el resultado está enteramente mediada por el tratamiento, es decir, si el paciente mejora no es a causa de la asignación, sino del tratamiento.

Varios autores han señalado que el éxito del diseño radica en la calidad de la VI , en el sentido de que tenga una asociación alta con A , y de que esté libre del efecto de confusores no observados.³⁴

Un ejemplo reciente se relaciona con el uso de antiinflamatorios no esteroideos (AINES) y el sangrado intestinal.³⁵ La dificultad de este estudio radica en el sesgo de selección, ya que en pacientes con un riesgo elevado de sangrado intestinal, se suele prescribir un nuevo tipo de AINES, llamados "coxibs" que tiene mayor tolerabilidad que los AINES no selectivos. La elección suele hacerse sobre la base de varios factores de riesgo (tabaquismo, alcoholismo, obesidad o antecedentes de úlcera péptica) para el sangrado intestinal. Este hecho genera un sesgo de selección para la medición de la asociación causal entre el uso de uno u otro tipo de AINES y el sangrado.

Sin embargo, el uso de uno u otro tipo de AINES, a menudo depende también de la preferencia del médico, que en casos de mediano o bajo riesgo, opta por uno u otro tipo de medicamento. Este hecho puede utilizarse para construir la VI siguiente:

$W =$ AINES no selectivo si en un paciente anterior elegido al azar se empleó AINES no selectivo.

$W =$ Coxib, si en ese paciente anterior elegido al azar se empleó COXIB.

$W:$

- Está libre de los efectos confusores, porque se trata de otro paciente diferente de aquel en el que se observarán los efectos.
- Se asocia con la prescripción, porque en buena medida la prescripción depende de las preferencias del médico.
- Todo su efecto está mediado por el tratamiento que en realidad recibió el paciente en cuestión.

Los puntajes de susceptibilidad. Un ejemplo de aplicación

El esófago de Barrett (EB) es una complicación frecuente del reflujo gastroesofágico.³⁶ En el hospital "Hermanos Ameijeiras" se realizó un estudio *cuasi* experimental para comparar los tratamientos quirúrgico y no quirúrgico en pacientes con EB. Se anticipaban sesgos de selección, debido a que la mayoría de casos no quirúrgicos fueron los pacientes que rechazaron la opción quirúrgica que se les recomendó.

Se estudiaron 35 pacientes, en 10 de los cuales se practicó cirugía. Las variables de respuesta, medidas antes y después del tratamiento fueron: un puntaje de calidad de vida y la longitud del Barrett, cuyos valores antes del tratamiento fueron utilizados como variables de control, junto con el grado de displasia.

Como muestra la tabla 2, los grupos exhibían diferencias en relación con las tres variables, y en dos de ellas estadísticamente significativas.

Tabla 2. Medias e intervalos de confianza (IC) al 95% para variables clave al inicio

Variables	Medias		IC al 95%	
	Grupo quirúrgico	Grupo medicamentoso	Límite inferior	Límite superior
Calidad de vida	39,6	36,5	0,6	5,5
Longitud del Barrett	3,8	2,4	0,4	2,4
Grado de displasia	1,8	1,5	-0,1	1,7

Debido a la falta de homogeneidad de los grupos y al consecuente sesgo de selección, se resolvió calcular un PS, que condensa en un indicador sintético las desigualdades manifiestas en las tres variables, que "confunden" el efecto del tratamiento.

El PS fue incluido como covariante en un modelo lineal para medir el efecto de la cirugía. La tabla 3 contiene los valores ajustados para el PS del cambio relativo en el puntaje de calidad de vida y la longitud del Barrett.

Tabla 3. Medias e intervalos de confianza (IC) al 95%, para el cambio relativo en el puntaje de calidad de vida y la longitud del Barrett

Variables	Medias		IC al 95%	
	Grupo quirúrgico	Grupo medicamentoso	Grupo quirúrgico	Grupo medicamentoso
Calidad de vida	0,36	0,17	0,29 - 0,43	0,13 - 0,21
Longitud del Barrett	0,47	0,05	0,34 - 0,60	-0,03 - 0,12

Cambio relativo: es la diferencia entre los puntajes antes y después, dividida por los puntajes antes y expresada en porcentaje.

Los intervalos de confianza de las columnas de la derecha constituyen estimaciones ajustadas del efecto de causalidad del tratamiento quirúrgico en relación con el medicamentoso, con respecto a las variables de respuesta consideradas.

El paradigma dominante de inferencia causal en la investigación epidemiológica contemporánea se basa en el enfoque de los resultados potenciales de *Rubin*. Su ventaja principal es que hace explícitos los parámetros que se desean estimar.

Los ensayos clínicos controlados y aleatorizados constituyen la regla de oro para la estimación de los efectos de causalidad, porque bajo condiciones de adhesión completa al tratamiento, ausencia de datos faltantes y diseño a ciegas, los estimadores del efecto de asociación, son estimadores insesgados del efecto de causalidad.

En estudios experimentales sin asignación aleatoria u observacionales, la existencia de confusores no observados constituye una insuperable amenaza de sesgo en la estimación del efecto de causalidad.

La tarea del investigador en escenarios de inferencia causal consiste en la identificación de los confusores, que no se realiza a partir del análisis estadístico de datos empíricos, sino mediante el conocimiento teórico específico del problema que origina la evaluación de la posible relación causal.

Uno de los procedimientos modernos que se emplean para la reducción del sesgo es el empleo de puntajes de susceptibilidad como recurso de apareamiento, estratificación o como covariantes que sintetizan el efecto de los confusores relevantes conocidos. Otro se basa en la identificación de variables instrumentales, que se asocian con el factor de exposición, no se asocian a ninguna variable determinante del resultado, y se asocian con este último solo a través de la mediación del factor de exposición.

La práctica de la salud pública moderna no puede prescindir de la inferencia causal, en al menos dos momentos clave: 1. Al planificar y diseñar las acciones para identificar los criterios de focalización y las dianas más efectivas de las acciones de salud. 2. Al evaluar el impacto de dichas acciones mediante correctos criterios de atribución.

Agradecimientos

Al Dr. *Guillermo Noa*, del servicio de gastroenterología del hospital "Hermanos Ameijeiras" por permitirme usar resultados de su investigación en este artículo.

REFERENCIAS BIBLIOGRÁFICAS

1. Sarewitz D. Beware the creeping cracks of bias. *Nature*. 2012;485: 149.
2. Ioannidis JPA. Why most published research findings are false. *PloS Medicine*. 2005;2:696-701.

3. Weisberg HI. Bias and causation. Models and judgment for valid comparisons. Hoboken NJ: Wiley; 2010.
4. Spirtes P, Glymour C, Scheines R. Causation, prediction and search. New York: Springer-Verlag; 1993.
5. Pearl J, Verma T. A theory of inferred causation. In: Allen JA, Fikes R, Sandewall E, editors. Principles of knowledge representation and reasoning: Proceedings of the 2nd International Conference. San Francisco C,A.: Morgan-Kaufmann Publishers; 1991. p. 441-52.
6. Robins JM, Wasserman L. On the impossibility of inferring causation from association without background knowledge. In: Glymour P, Cooper G, editors. Computation, causation and discovery. Menlo Park CA, Cambridge MA: AAAI Press/MIT Press; 1999. p. 305-21.
7. Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation. An application to birth defects epidemiology. *Am J Epidemiol.* 2002;155:176-84.
8. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10:37-48.
9. Greenland S, Pearl J. Causal diagrams. In: Boslaugh S, editor. Encyclopedia of Epidemiology. Thousand Oaks CA: Sage Publications; 2007. p. 149-56.
10. Weinberg CR. Towards a clearer definition of counfounding. *Am J Epidemiology.* 1993;137:1-8.
11. Robins JM. Data, design and background knowledge in etiologic inference. *Epidemiol.* 2001;12:313-20.
12. Greenland S, Robins JM. Identifiability, exchangeability and epidemiological confounding. *Int J Epidemiol.* 1986;15:413-9.
13. Christenfeld NJS, Sloan RP, Carroll D, Greenland S. Risk factors, confounding and the illusion of statistical control. *Psychosomatic Med.* 2004;66:868-75.
14. Salmon WC. Causality and explanation. New York: NY. Oxford Univ. Press; 1998.
15. Morabia A. Epidemiology: an epistemological perspective. In: Morabia A, editor. A history of epidemiologic methods and concepts. Basel, Switzerland: Birkhauser Verlag; 2004. p. 3-124.
16. Swaen E, van Amelsvoort M. A weight of evidence approach to causal inference. *J Clin Epidemiol.* 2009;62:270-7.
17. Botti C, Comba P, Forastiere F, Settini L. Causal inference in environmental epidemiology: the role of implicit values. *The Science Total Environment.* 1996;84:97-101.
18. Weed DL. Environmental Epidemiology Basics and Proof of Cause-Effect. *Toxicology.* 2002;82:399-403.

19. Phillips CV, Goodman KJ. Causal criteria and counterfactuals; nothing more (or less) than scientific common sense. *Emerging Themes Epidemiol* [Internet]. 2006 [cited 2010 Ago 16]; 3:5. Available from: <http://www.ete-online.com/content/3/1/5>
20. Rothmann KJ, Greenland S. Causation and causal inference in Epidemiology. *Am J Public Health*. 2005; 95: S144-S150.
21. Rubin D. Which ifs causal answers. Comment on "Statistics and Causal Inference" by Paul W. Holland. *J Am Stat Assoc*. 1986; 81: 961-2.
22. Rubin D. Formal models of statistical inference for causal effects. *J Stat Planning Inference*. 1990; 25: 279-92.
23. Heckman JJ. Causal parameters and policy analysis in Economics: a 20th. Century retrospective. *The Quarterly J Economics*. 2000; 115: 45-97.
24. Greenland S, Robins JM. Identifiability, exchangeability and epidemiological confounding. *Int J Epidemiol*. 1986; 15: 413-9.
25. Greenland S, Robins JM. Identifiability, exchangeability and confounding revisited. *Epidemiol Perspectives Innovations* [Internet]. 2009 [cited 2012 Ago 16]; 6: 4. Available from: <http://www.epi-perspectives.com/content/6/1/4>
26. Steyer R. Analyzing individual and average causal effects via structural equation models. *Methodology*. 2005; 1: 39-64.
27. Steyer R, Gabler S, von Davier AA, Nachtigall C. Causal regression models. II: Unconfoundedness and causal unbiasedness. *Methods Psychological Res*. 2000; 5: 55-87.
28. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc*. 1984; 79: 516-24.
29. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983; 70: 41-55.
30. Austin PC, Mamdani MM, Stukel TA, Anderson GM, Tu JV. The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Stat Med*. 2005; 24: 1563-78.
31. Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. *Stat Med*. 2006; 25: 2084-106.
32. Ukoumunne OC, Williamson E, Forbes AB, Gulliford MC, Carlin JB. Confounder-adjusted estimates of the risk difference using propensity score-based weighting. *Statist Med*. 2010; 29: 3126-36.
33. Austin PC. The performance of different propensity score methods for estimating relative risks. *J Clin Epidemiol*. 2008; 61: 537-45.
34. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000; 29: 722-9.
35. Brookhart MA, Schneeweis S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostatistics* [Internet]. 2007 [cited 2012 Ago 16]; 3(14). Available from: <http://www.bepress.com/ijb/vol3/iss1/14>

36. Winters C Jr, Spurling TJ, Chobanian SJ, Curtis DJ. Barrett´s esophagus. A prevalent, occult complication of gastroesophageal reflux disease. *Gastroenterology*. 1987;92: 118-24.

† Otras definiciones, como el cociente, son también posibles, pero la selección de un indicador de efecto puede influir en la interpretación de medidas poblacionales como el efecto promedio, o de la heterogeneidad interindividual.

‡ La expresión original en inglés es "propensity store."

Jorge Bacallao Gallestey. Centro de Investigaciones y Referencia de Aterosclerosis de La Habana (CIRAH). Policlínico "19 de Abril". Calle Tulipán y Panorama. Nuevo Vedado, Plaza. La Habana, Cuba. Correo electrónico: jbacallao@infomed.sld.cu