



COLABORACIÓN
ESPECIAL

La causalidad y la confusión

De la Macorra García JC.
La causalidad y la confusión.
Cient. Dent. 2025; 22; 2; 57-60



Prof. José Carlos de la Macorra García
Licenciado en Medicina,
Especialista en Estomatología,
Especialista en Medicina y
Trabajo, Doctor en Medicina y
Cirugía (1988) en la Universidad
Complutense de Madrid.
Profesor Emérito (Universidad
Complutense de Madrid).

LA CAUSALIDAD

Para ponernos en situación veamos muy brevemente en primer lugar el concepto de causalidad. En 1965 Austin Bradford Hill¹ sugirió que la pregunta decisiva para conocer si existe una posible relación causal es “*si la frecuencia con la que un evento indeseable, B (Figura 1), sería influenciada por un cambio en una característica del ambiente, A*”. En este mismo trabajo listó las características de dicha posible asociación entre A y B (*robustez, consistencia, especificidad, temporalidad, curva dosis-respuesta, plausibilidad biológica, coherencia, evidencia experimental y analogía*) que deberían valorarse, que son relevantes para dar credibilidad a esa asociación causal.

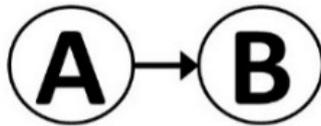


Figura 1. Una relación probablemente causal.

Debe, en este momento, señalarse que el hecho de encontrar una posible asociación no supone la demostración de la causalidad. Esta solo podría establecerse a través de un estudio específico.

LA CONFUSIÓN Y LA COMPARABILIDAD

La confusión en la causalidad no parece, así de entrada, algo que queramos tener en una investigación. Además de los si-

nónimos claramente negativos que le atribuye la RAE^a, en una investigación su presencia es aciaga: hay confusión cuando el efecto de la exposición – *el efecto del pretendido agente causal cuya acción se está estudiando, A en el ejemplo anterior de la Figura 1-* sobre el resultado – *el evento indeseable, B en el mismo ejemplo-* se confunde con el efecto de otro factor de riesgo u otro factor protector^{b 3}. Puede comprenderse que se atribuirían a A unos efectos que no se deben, total o parcialmente, a su acción o presencia.

La confusión aparece cuando no puede garantizarse la comparabilidad adecuada de los grupos, de manera que absolutamente todos los factores que afectan a los grupos que se cotejan sean idénticos. La situación ideal –de máxima comparabilidad– se daría cuando se compara el efecto de una exposición (A) sobre un grupo, con lo que ocurre con ese mismo grupo pero no expuesto. Esto es, a todos los efectos, imposible.

Para acometer esta dificultad, se busca comparar lo que ocurre en dos o más grupos contruidos de manera que sean lo más comparables posible; uno se somete a la exposición del pretendido agente causal y el otro o los otros a otros agentes o a ninguno. Pero esta comparabilidad entre los grupos nunca es perfecta, pues los sujetos que los forman son diferentes por más que se intente que sean similares. Y, además, los grupos pueden estar sometidos a influencias diferentes.

^a. Confusión: sinónimo de desorden, desbarajuste, caos, embrollo, alboroto, barahúnda, guirigay, lío, batiburrillo, desquicio, berrodo (RAE)

^b. Factor protector: algo que puede disminuir la probabilidad de sufrir una cierta enfermedad ². Debe considerarse como lo contrario de un factor de riesgo, el cual aumenta dicha probabilidad.

Correspondencia:
Prof. José Carlos de la Macorra García
Departamento de Odontología
Conservadora y Prótesis.
Facultad de Odontología.
Plaza Ramón y Cajal s/n
Ciudad Universitaria. 28040 Madrid.
macorra@ucm.es

Si no son suficientemente comparables los grupos, podría haber en alguno de ellos algún factor que provoque una respuesta adicional, que no conocemos previamente o que no hemos valorado suficientemente, que confunda el resultado que se estudia: un factor de confusión. Su presencia puede incrementar la respuesta *-haciendo que parezca que el efecto es mayor-* o disminuirla *-enmascarando el efecto-*.

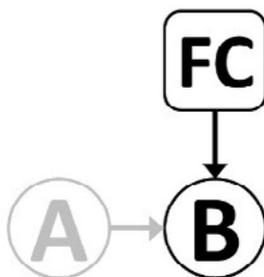


Figura 2. Un posible factor de confusión.

Para poder considerar con propiedad algo como un factor de confusión se considera desde hace tiempo que debe cumplir simultáneamente, tres requisitos⁴:

1. **Ser una causa de la enfermedad** (como sería en la Figura 1) **o una medida substitutiva de una causa**, en gente expuesta (como en el ejemplo de la Figura 2). Es decir: **ser, por sí mismo, un factor de riesgo de la enfermedad**.

De lo anterior se desprende que un factor de confusión puede ser per se un predictor del resultado, aún en la ausencia de la exposición que se está estudiando⁵. Como se indica en la Figura 2, el posible factor de confusión (FC) podría causar B, aún en ausencia de A -el pretendido agente causal-

Este es un requisito clásico para estudios sobre enfermedad, pero es válido también para otros estudios que evalúan otras relaciones. Por ejemplo: la pertenencia a un grupo social determinado (un equipo de fútbol, por ejemplo), ¿se asocia con una conducta antisocial?. Claramente, en este estudio la edad es muy probablemente un factor de confusión: es muy probable que cuanto ma-

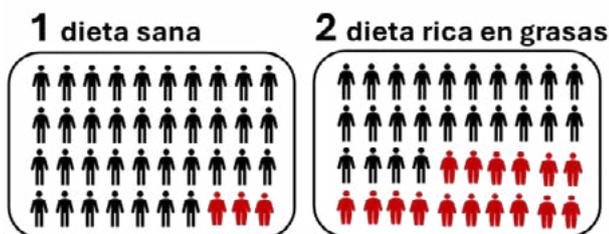


Figura 3. Factor de confusión relacionado con la exposición.

yor sea el individuo (cualquiera que sea el equipo al que se pertenezca), menos antisocial será la conducta.

2. **Estar relacionado** (positiva o negativamente) con la exposición en la población de estudio.

Si estamos estudiando la relación entre la dieta (la exposición) y la enfermedad coronaria (la respuesta), y la población de estudio está dividida en expuestos, con una dieta sana, y no expuestos, con una grieta rica en grasas (grupos 1 y 2 en la Figura 3), el posible factor -por ejemplo, la obesidad- debe tener diferente prevalencia en ambos grupos para ser considerado un factor de confusión. En dicha figura, afecta a los miembros de los grupos coloreados en rojo. En este caso, es más prevalente en el grupo 2.

3. **No ser un factor intermedio en la vía etiopatogénica entre la exposición y la respuesta**.

En la Figura 4 vemos un posible factor (FC) intermedio entre la exposición (A) y la respuesta (B). En esta situación, **no sería un factor de confusión**.

En el ejemplo del estudio de la relación entre la enfermedad coronaria y la dieta, la respuesta (la enfermedad coronaria, B en la Figura 4) podría ser explicada midiendo el nivel de colesterol. La colesterolemia elevada no es un factor de confusión, puesto que puede ser un factor causal intermedio entre la dieta y la enfermedad coronaria.



Figura 4. Factor intermedio entre la exposición y la respuesta.

4. A los tres requerimientos clásicos anteriores actualmente se suma el que el posible factor de confusión **no debe estar afectado por la exposición**⁶. Este sería el caso, que se muestra en la Figura 5, en el que tanto el posible factor de confusión (FC) como la respuesta (B)

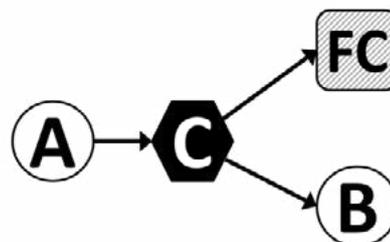


Figura 5. Factor afectado por la exposición.

tienen una causa común (C), que depende de la exposición.

EL CONTROL DE LA CONFUSIÓN

Siendo potencialmente ubicuos, la posible presencia de factores de confusión en una investigación es una de las principales preocupaciones de los investigadores en el diseño o análisis de estudios. Para controlarlos y excluir en lo posible su acción, habitualmente se siguen unos pasos.

El primero, evidente, es que debe dominar muy bien el proceso, los posibles agentes causales (o exposiciones) y sus interacciones y el proceso biológico (o social, en el caso de los equipos de fútbol) que lo gobierna: conocer lo que se sabe. Para esto es obligado estudiar la literatura pertinente actual y pasada y buscar y dejarse aconsejar por quien tenga conocimientos y experiencia.

Además, hay una posible lista de estrategias o acciones que nos ayudarán a controlar la confusión. Estas pueden aplicarse antes de realizar el estudio (durante la fase de su diseño) o después (durante la fase de evaluación de los datos obtenidos). Todas ellas tienen ventajas y desventajas, y algunas de ellas son difíciles de implementar. Son^{4, 7}:

Durante la fase de diseño:

- Aleatorización

- Aunque la aleatorización de grandes grupos de pacientes resultará normalmente en una distribución equilibrada de los factores de confusión conocidos y de los desconocidos entre los grupos, es improbable que este equilibrio se consiga para todas las características de los pacientes y, además, exige que los grupos sean grandes: la aleatorización no puede funcionar bien en grupos pequeños. Y conseguir grupos grandes es difícil y costoso.

- Restricción

- Si se conoce un factor de confusión (por ejemplo: la edad), se trata de limitar el acceso al estudio a un rango de casos limitado (por ejemplo: solo casos entre 50 y 55 años).

- Por esta razón se limita el acceso a los estudios, habitualmente, de acuerdo a los factores de confusión más comunes (edad, sexo, estado de salud, hábitos del paciente y otros)

- Los problemas son que restringe el reclutamiento, pues es más difícil encontrar los casos, ya que deben cumplir los requisitos, y que la validez externa del estudio disminuye: sus conclusiones serán aplicables solo a la población a la que representa.

- Emparejamiento (o *matching*)

- Consiste en admitir en los grupos a pacientes emparejándolos, de manera que -si se hiciera por edad, por ejemplo- si en el grupo experimental se admite un varón de 55 años, se admita a otro en el grupo control con esa misma edad.

- Puede ser complicado si se empareja por varios factores de confusión: no se encontrarán casos fácilmente.

- Estratificación

- Se preparan los grupos de manera que se estratifiquen los casos, de acuerdo a los factores de confusión conocidos. Se conformarán, dentro de los grupos, varios subgrupos de acuerdo a dichos factores.

- También puede hacerse durante la fase de análisis, cuando haya concluido la investigación, si se encuentra(n) alguna(s) variable(s) de confusión desconocida(s) hasta ese momento y que se hacen patentes a la vista de los resultados, pero deberá haber casos suficientes.

- El principal inconveniente es que requiere muchos casos (en todos los grupos), de manera que los datos tengan significado para todos y cada grupo.

Durante la fase de análisis

- Técnicas multivariantes.

- Es la técnica estadística que examina el efecto de un factor controlando el efecto de varios otros

- Pueden controlarse así más factores que en la estratificación, pero los resultados son más difíciles de interpretar.

- Estratificación

- Hay otras técnicas (estandarización, puntajes de propensión (*propensity scores*), análisis de sensibilidad, ponderación de probabilidad inversa)^{3, 8, 9} cuya exposición escapa a la intención de este escrito.

Como puede verse en las estrategias que pueden utilizarse durante el diseño del estudio, siempre que se gana algo, se pierde algo: si se disminuye la confusión el estudio será más costoso (la aleatorización y la estratificación exigen más casos) y/o los casos serán más difíciles de encontrar (si se utiliza la restricción o el emparejamiento).

Esta breve visión de la causalidad y de la confusión ahonda en la afirmación hecha en publicaciones anteriores de esta serie acerca de que algunas de las reglas del método científico no son simples ni intuitivas, y acerca de que el aprendizaje de estos métodos es largo y complejo.



BIBLIOGRAFÍA

1. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58(5):295-300.
2. National Cancer Institute. Protective factor: NIH; 2025 [Disponible en: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/protective-factor>].
3. Howards PP. An overview of confounding. Part 1: the concept and how to address it. *Acta Obstet Gynecol Scand.* 2018;97(4):394-399.
4. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet.* 2002;359(9302):248-352.
5. Skelly AC, Dettori JR, Brodt ED. Assessing bias: the importance of considering confounding. *Evid Based Spine Care J.* 2012;3(01):9-12.
6. McNamee R. Confounding and confounders. *Occup Environ Med.* 2003;60(3):227-234.
7. Jager KJ, Zoccali C, MacLeod A, Dekker FW. Confounding: What it is and how to deal with it. *Kidney Int.* 2008;73(3):256-260.
8. Schober P, Vetter TR. Propensity score matching in observational research. *Anesth Analg.* 2020;130(6):1616-1617.
9. Van Smeden M, Penning de Vries BBL, Nab L, Groenwold RHH. Approaches to addressing missing values, measurement error, and confounding in epidemiologic studies. *J Clin Epidemiol.* 2021;131:89-100.