

Sesgos y Confusores

Nigel Paneth M.D., MPH

College of Human Medicine Michigan State Univ.

paneth@msu.edu

Nicolás Padilla, M.D.

Universidad de Guanajuato, México

padilla@celaya.podernet.com.mx

Inicialmente en el sitio: www.pitt.edu/~super1/

Formulación de Hipótesis y Errores en la Investigación

Todos los estudios analíticos deben iniciar con una hipótesis claramente formulada. La hipótesis deberá ser *cuantificable y específica*. Deberá predecir una relación de un específico tamaño.
begin with a clearly formulated hypothesis.

- **Por ejemplo:**

“Bebés que son alimentados al seno materno tienen menos enfermedades que los bebés alimentados con fórmula.”

¿Cuáles enfermedades? ¿Cómo es definido el tipo de alimentación? ¿Cuán grande es la diferencia en riesgo?

- **Un mejor ejemplo:**

“Bebés exclusivamente alimentados al seno materno por 3 meses o más tendrán una reducción en la incidencia de hospitalizaciones por gastroenteritis de al menos del 30% durante el primer año de vida.”

Sólo predicción específica permite extraer legítimas conclusiones de un estudio que prueba una hipótesis. Pero aún con mejores hipótesis formuladas, dos tipos de errores pueden ocurrir.

- Tipo 1 - observando una diferencia cuando no hay ninguna.**
- Tipo 2 - fracasando para observar una diferencia cuando realmente hay.**

Esos errores son generalmente producidos por uno o más de los siguientes:

- Error aleatorio**
- Misclasificación al azar**
- Sesgo**
- Confusores**

Error Aleatorio

Desviación de resultados e inferencias de la verdad, ocurriendo como resultado del efecto del azar. Puede producir errores tipo 1 o tipo 2.

Misclasificación al Azar (no diferencial)

Error al azar aplicado a la **medición** de una exposición o resultado. Errores en clasificación pueden sólo producir errores de tipo 2, excepto si aplica a un *confusor* o a un *gradiente de exposición*.

Sesgo

Sistemática, no aleatoria desviación de resultados e inferencias de la verdad, o **procesos que dan lugar** a tal desviación. Cualquier tendencia en la colección, análisis, interpretación, publicación o revisión de datos que puedan dar lugar a conclusiones que son sistemáticamente diferentes de la verdad. (Diccionario de Epidemiología, 3a ed.)

Más Acerca de Sesgo

Note que en sesgo, el enfoque es sobre un *artefacto* de alguna parte del *proceso* de investigación (reunión de sujetos, colección de datos, análisis de datos) que producen un resultado espurio. Sesgo puede producir errores tipo 1 o tipo 2, pero usualmente nos enfocamos en el tipo 1 debido a sesgo.

Mós sobre Sesgo

Sesgo puede ser conciente o inconciente. En epidemiología, la palabra sesgo no implica, como en su uso común, perjuicio o desviación deliberada de la verdad.

Confusores

Un problema resultante de que un hecho de los sujetos en estudio *no ha sido separado* de un segundo hecho, y ha sido *confundido* con él, produciendo un resultado espurio. El resultado espurio surge del efecto del primer hecho siendo erróneamente atribuido al segundo hecho. Confusión puede producir un error tipo 1 o tipo 2, pero usualmente se enfoca en errores tipo 1.

La Diferencia entre Sesgo y Confusor

Sesgo crea una asociación que no es verdad, pero confusores describe una asociación que es verdad, pero potencialmente erróneo.

Ejemplos de Error Aleatorio, Sesgo, Misclasificación y Confusor en el Mismo Estudio:

Estudio: en un estudio cohorte, bebés de mujeres quienes alimentaron con fórmula y quienes alimentaron al seno, fueron comparadas, y se encontró que la incidencia de gastroenteritis, registrada así en registros médicos, es más baja en los bebés alimentados al seno.

Ejemplo de Error Aleatorio

Por azar, hay más episodios de gastroenteritis en el grupo alimentado con fórmula en la muestra del estudio, produciendo un error tipo 1. (Si no se encuentran diferencias, un error tipo 2 podría haber ocurrido si **por azar** habían sido pocos episodios de gastroenteritis en el grupo alimentado al seno.)

Ejemplo de Misclasificación Aleatoria

Ausencia de buena información sobre la historia de alimentación resulta en que algunas madres que alimentan al seno siendo aleatoriamente clasificadas como que alimentan con fórmula, y viceversa. Si sucede, el estudio encontrará *subestimaciones* del verdadero RR, cualquier modalidad de alimentación está asociada con alta incidencia de enfermedad, produciendo un error tipo 2.

Ejemplo de Sesgo

Los registros médicos de bebés alimentados con fórmula **sólo** son *menos completos* que aquellos de los alimentados al seno, y así, registrar **sólo** pocos episodios de gastroenteritis entre ellos.

Ejemplo de Confusor

Las madres de bebés alimentados al seno son de clase social alta, y los bebés tienen *mejor higiene, menos hacinamiento* y quizá otros factores que protegen contra la gastroenteritis. Hacinamiento e higiene son **verdaderamente** protectores contra gastroenteritis, pero erróneamente atribuimos sus efectos a la alimentación al seno.

Protección contra Error Aleatorio y Misclasificación Aleatoria

Error aleatorio puede falsamente producir una asociación (error tipo 1) o falsamente no producir una asociación (error tipo 2).

Nos protegemos contra la misclasificación aleatoria produciendo error tipo 2 eligiendo la más *precisa y segura* medición de la exposición y del resultado.

Protección contra Error Tipo 1

Protegemos nuestro estudio contra errores aleatorio tipo 1 estableciendo que el resultado deberá ser improbable que haya ocurrido por azar (v.gr. $P < 0.05$). Valores de p son establecidas para proteger contra errores tipo 1 debido al azar y *no se garantiza protección contra errores tipo 1 debido a sesgos o confusores*. Por esto decimos que estadísticos demuestran *asociación* pero no *causalidad*.

Protección contra Error Tipo 2

Protegemos nuestro estudio contra errores tipo 2 por:

- **proveyendo tamaño de muestra adecuado y**
- **haciendo hipótesis de grandes diferencias.**

Entre más grande el tamaño de muestra, más fácil se detectará una verdadera diferencia, y las más grandes diferencias serán más fácilmente detectadas. (Imagine cuán difícil será detectar un riesgo aumentado del 1% de gastroenteritis en los alimentados con fórmula)

Dos Formas de Incrementar Poder

El tamaño de muestra necesario para detectar una diferencia significativa es llamado el *poder* de un estudio.

1. Eligiendo las mediciones más *precisa y segura* de la exposición y del resultado, tienen el efecto de incrementar el poder de nuestro estudio, debido a que las varianzas de las mediciones del resultado, que entran en las pruebas estadísticas, están disminuídas.
2. Teniendo un **tamaño de muestra adecuado** de sujetos en estudio.

Principios Claves en Sesgos y Confusores

El factor que crea el sesgo, o la variable confusora, *deberá estar asociada con las variables dependiente e independiente* (v.gr. Con la exposición y con el resultado). Asociación del sesgo o confusor sólo con una de las dos variables no es suficiente para producir un resultado espurio.

**Buena Planeación del estudio
Protege contra Todas las
Formas de Error**

En el ejemplo dado:

El SESGO, llamado registro incompleto tiene que estar asociado al **tipo de alimentación** (variable independiente) y también con el registro de **gastroenteritis** (variable dependiente) para producir un resultado falso.

La VARIABLE CONFUSORA (o confusor) **mejor higiene**, tiene que estar asociado con el **tipo de alimentación** y también con **gastroenteritis** para producir un resultado espurio.

Si el sesgo o el confusor están asociados *sólo* con la variable dependiente o *sólo* con la variable independiente, no producirán sesgo ni confusión.

Esto tiene una norma útil:

Si puedes mostrar que un confusor potencial **no** está asociado con una de las variables en el estudio (exposición o resultado), confusión puede ser excluído.