

Lectura 3. El error en epidemiología

Los estudios científicos tratan de responder a una pregunta concreta, que suele estructurarse en una serie de componentes, como son la población de estudio, la intervención o exposición que se evalúa, con qué se compara y el resultado¹. La pregunta debe contestarse de la forma más objetiva y fiable posible, por lo que, en la fase de diseño del estudio, debe hacerse un esfuerzo para evitar cualquier tipo de error en la respuesta. De esta forma, cualquier diferencia entre los grupos del estudio será debida al efecto de la intervención o exposición estudiada.

Sin embargo, en la práctica puede haber factores que invaliden los resultados del estudio, ya sea por defecto de diseño o de análisis, o por las propias características de las variables de resultado medidas. Esto puede producirse por la presencia de errores en la selección, medición o en el análisis.

TIPOS DE ERRORES

Desde un punto de vista más general, distinguimos dos tipos de errores: el error aleatorio y el error sistemático².

El error aleatorio se debe al azar y puede provenir de dos fuentes distintas. La primera sería por errores inherentes al muestreo dentro de la población. Cuando obtenemos una muestra, lo hacemos con el objetivo de estimar un parámetro poblacional a través de su estudio en una muestra. Sin embargo, debido al azar de muestreo, podemos obtener una muestra que no sea representativa de la población (aunque obtengamos múltiples muestras, todas serán diferentes unas de otras). Esto es más probable que ocurra cuando los tamaños muestrales son pequeños y cuando se emplean técnicas de muestreo no probabilístico³.

La otra fuente de error aleatorio es la variabilidad en la medición, ya sea por imprecisiones del instrumento de medida o por la propia variabilidad biológica. Por ejemplo, si medimos la presión arterial varias veces a un mismo individuo, los resultados de las distintas mediciones no serán exactamente iguales, aunque podemos minimizar el error en este caso repitiendo varias mediciones y calculando una media de ellas.

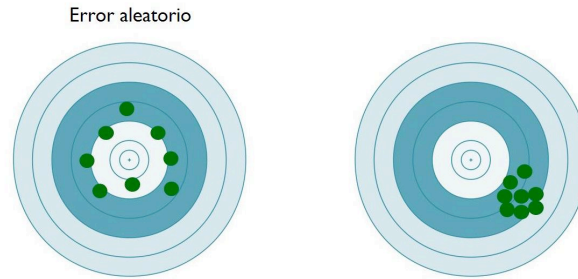
El error sistemático, también llamado sesgo, se produce por un error en el diseño o análisis del estudio, que produce una estimación incorrecta o no válida del efecto que se estudia. Imaginemos, por ejemplo, que vamos a pesar a un grupo de escolares y a los niños los pesamos sin ropa y descalzos y a las niñas las pesamos con ropa; en este caso incurriremos en un sesgo en la medición del peso.

Vemos, pues, que existen diferencias entre los errores aleatorio y sistemático. El error aleatorio se relaciona con la precisión del resultado. Una medida será tanto más precisa cuanto menor sea el componente aleatorio, por lo que el error puede disminuirse aumentando el tamaño de la muestra o mejorando la precisión de las mediciones. Este error no afecta a la validez interna ni externa.

Por su parte, el error sistemático se relaciona con la validez de los resultados, ya que la estimación solo puede ser válida si representa el verdadero valor. Como es lógico, este tipo de error es más difícil de controlar y no se corrige aumentando el tamaño muestral.

En la figura 1 vemos representado un símil para aclarar la diferencia entre error aleatorio y sesgo. En el error aleatorio, si aumentamos el número de disparos acabaremos centrando el resultado en la diana, mientras que en el error sistemático no. Existen innumerables tipos de sesgos descritos, aunque todos pueden agruparse dentro de tres grandes categorías: sesgos de selección, sesgos de información o clasificación y sesgos de análisis⁴.

FIGURA 1. DIFERENCIAS ENTRE ERROR ALEATORIO Y SESGO



SESGOS DE SELECCIÓN

Se producen cuando la muestra de la que se obtienen los resultados no es representativa de la población, pero no por efecto del azar, sino por un error en los criterios de inclusión o exclusión, o en la sistemática de reclutamiento. Estos sesgos afectan solo la validez externa, ya que el diseño y análisis de los resultados pueden ser correctos, pero las conclusiones solo pueden aplicarse a la muestra estudiada. Este tipo de sesgos puede darse en diversas situaciones:

- Selección incorrecta de los grupos de estudio: si evaluamos el efecto de un factor de exposición o intervención, los grupos comparados solo deberían diferenciarse en el factor evaluado, sin embargo, estos grupos pueden presentar diferencias en otros factores implicados en el efecto. Esta comparabilidad se consigue en los estudios experimentales con la aleatorización, por lo que los estudios observacionales son más susceptibles a este tipo de sesgos.
- Pérdidas durante el seguimiento: cuando los que se pierden son sistemáticamente diferentes de los que no en variables relacionadas con el factor de estudio o la variable respuesta. Es lo que se denomina como pérdidas informativas.
- Ausencia de respuesta: esto es propio de las encuestas transversales. Hay que tratar de recoger información para saber si los que no contestan difieren en algo de los que sí lo hacen.
- Supervivencia selectiva: se produce cuando se seleccionan casos prevalentes en lugar de incidentes. Este tipo de sesgo es típico de los estudios de casos y controles, en los que los casos más graves o más leves están infrarrepresentados por exitus o curación.
- Sesgo de autoselección por participación de voluntarios: en general, existe riesgo de que estos individuos tengan características diferentes a los no voluntarios.

SESGO DE INFORMACIÓN O CLASIFICACIÓN

Se producen cuando las mediciones de las variables de estudio se realizan de forma deficiente en todos o parte de los sujetos de estudio. Esto puede ocurrir por el uso de pruebas poco sensibles o poco específicas, por utilizar criterios diagnósticos incorrectos o variables y por imprecisión en la recogida de datos.

Cuando nos equivocamos de manera similar en los dos grupos, hablamos de un sesgo de información no diferencial. Por el contrario, si nos equivocamos preferente o exclusivamente en un grupo, el sesgo es diferencial. El sesgo no diferencial sesga los resultados a favor de la hipótesis nula (tiende a disminuir la magnitud de las diferencias entre grupos), por lo que, si podemos rechazarla (las diferencias son significativas)

TRATAMIENTO

El resultado tendrá valor. Sin embargo, el sesgo diferencial es imposible de predecir y compromete seriamente la validez del estudio.

Dentro de los sesgos de información podemos hacer referencia a dos muy típicos de los estudios de casos y controles⁵ (también de los de cohortes retrospectivas): sesgo de memoria (los que tienen un problema de salud recuerdan sus antecedentes de forma diferente a los que no lo tienen) y sesgo del entrevistador (la información se solicita o interpreta de distinta manera según el grupo al que pertenece el sujeto).

SESGOS DE ANÁLISIS

Los sesgos de análisis son los que se producen cuando, al evaluar el efecto de una variable de exposición o intervención en una variable de resultado, no se tiene en cuenta la influencia de terceras variables relacionadas con las anteriores. Una peculiaridad de los sesgos de análisis es que además de poder evitarse en la fase de diseño del estudio, pueden ser corregidos en la fase de análisis, mediante diferentes técnicas estadísticas. Podemos diferenciar dos tipos de variables implicadas en estos sesgos: las variables de confusión y las variables de interacción o de modificación del efecto.