



Errores metodológicos. Sesgos

Methodological errors. Biases

10.20960/angiologia.00665

08/02/2024

Errores metodológicos. Sesgos

Methodological errors. Biases

Antonio Martín Conejero¹, Víctor Quirós González²

¹Servicio de Angiología, Cirugía Vascular y Endovascular. Hospital Clínico San Carlos. Madrid. ²Medicina Preventiva. Oficina de la Estrategia 2020-2024 “Transforma 12”. Hospital Universitario 12 de Octubre. Madrid

RESUMEN

Uno de los aspectos más importantes en la investigación es el hecho de ser capaces de comunicar adecuadamente los resultados que obtenemos en nuestros estudios. Sin embargo, comunicar sobre un trabajo que contiene errores solo sirve para distorsionar la realidad científica, por lo que es esencial conocer qué situaciones pueden afectar negativamente a la calidad de nuestros estudios. Presentamos en este artículo, a modo de resumen, algunos de ellos.

Palabras clave: Sesgos. Factor de confusión.

ABSTRACT

One of the most important aspects of research is being able to adequately communicate the results we obtain in our studies. However, communicating about studies that contains errors only serves to distort the scientific reality, so it is essential to know which situations can negatively affect the quality of our studies. We present in this article, by way of summary, some of them.

KeyWords: Bias. Confusion factor.

INTRODUCCIÓN

En este capítulo veremos algunas de las formas en las que los errores metodológicos pueden alterar nuestro trabajo de investigación.

Veamos un par de situaciones a las que se enfrenta un investigador:

- Imaginemos que queremos hacer un estudio para comparar si el tabaco es un factor de riesgo de los AAA y para ello usamos como grupo control de comparación a pacientes de la planta de urología, que fuman tanto o más que nuestros pacientes. ¿Tendrá eso consecuencias en la asociación tabaco-AAA que estamos estudiando?
- Y si cuando decidimos estudiar la permeabilidad del *bypass* y del *stent* en la patología oclusiva de la arteria femoral superficial observamos que el porcentaje de diabéticos es mucho mayor en el grupo *stent*, ¿influirá ello en la permeabilidad de los grupos de comparación?

Los errores sistemáticos, o también llamados *sesgos*, afectan a la validez interna del estudio. Es decir, influyen negativamente en la capacidad de medir lo que realmente queremos medir en nuestro estudio. Nos impiden, por ejemplo, saber si un factor de riesgo se asocia a una enfermedad o si un tratamiento es diferente de otro.

Pero no todos los errores que podemos cometer son problemas de validez. Algunos son problemas de precisión relacionados con el azar y, por tanto, con el tamaño muestral. Imaginemos esta situación clásica, que seguro que conocerá (Fig. 1): tenemos unas dianas y nos piden que lancemos unos dardos con el siguiente resultado:

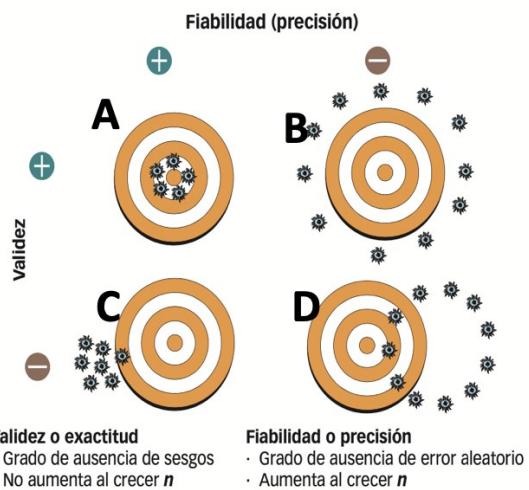


Figura 1. Diferencia entre error aleatorio y error sistemático. n : tamaño muestral.

En la diana A no cometemos ningún tipo de error. Damos en la diana, habremos hecho un estudio correctísimo. En la diana D nos equivocamos completamente. No damos ni una bien, quizá lo nuestro no es la investigación.

Las dianas B y C nos permiten diferenciar lo que es un error aleatorio de lo que es un error sistemático. En la situación B el investigador conoce el centro de la diana; es decir, el objetivo que persigue. El problema es que es incapaz de llegar a él. Sería como si pidiésemos que disparase a una persona que no sabe. Dispararía a cualquier parte, pero si le pedimos que dispare muchas veces, es posible que finalmente acierte. La diana C es diferente. Sería como si al mejor tirador del mundo le diésemos una escopeta trucada. Dispararía de maravilla a donde no debe. Siempre en el mismo sentido. Por mucho que disparase, siempre erraría en el mismo sentido.

La situación B es la que representa un error aleatorio; es decir, un error debido al azar. Si analizamos la relación entre tabaco e IAM, un error aleatorio implicaría decir que no encontramos asociación entre ambos. No “damos” en la diana. Es lo que ocurre cuando tenemos una p no significativa u obtenemos un intervalo de confianza muy

ancho: en ambas situaciones dicho error podrá atenuarse aumentando la muestra.

La situación C afecta a la validez. Si analizamos la relación entre tabaco e IAM, el sesgo podría dar lugar a encontrar una asociación *protectora* entre ambos parámetros, que no corregiremos, aunque aumentemos el tamaño de la muestra.

Existen multitud de sesgos que podemos cometer al hacer un estudio epidemiológico. En este texto veremos algunos de los que más frecuentemente se cometan en nuestra rutina investigadora, sobre todo cuando hacemos estudios de casos y controles.

SESGO DE SELECCIÓN

Se produce cuando se asignan sujetos al grupo control que difieren significativamente, en alguna característica clave, del grupo problema. Este tipo de sesgos puede controlarse mediante un proceso de aleatorización en la constitución de los distintos grupos de estudio. Elegimos mal, por diferentes motivos, a los sujetos que formarán parte de nuestro estudio.

Se exponen a continuación (Tabla I) algunos de los sesgos de selección que se comenten más frecuentemente en función del tipo de estudio realizado:

Tabla I. Sesgo de selección y tipo de estudio

Diseños prospectivos	Diseños retrospectivos
Sesgo del trabajador sano	Falacia de Neyman
Seguimiento incompleto de participantes	Sesgo de Berkson
Sesgo de autoinclusión	Sesgo de sospecha diagnóstica

Sesgo de autoselección o del voluntario

La participación o autoderivación del individuo al estudio compromete su validez. Imaginemos que un paciente oncológico terminal nos pide participar en un ensayo con un nuevo tratamiento prometedor. ¿Qué ocurriría si incluimos en nuestro estudio a pacientes con tan mal pronóstico? La respuesta terapéutica no sería buena, pero no porque el fármaco no sea eficaz, si no que muy probablemente sería por la situación tan avanzada de la enfermedad de nuestros pacientes. Es decir, el sesgo del voluntario implica que la “autoinclusión” de estos pacientes condicione una respuesta terapéutica diferente a la del paciente “habitual”.

Sesgo diagnóstico o sesgo de Berkson

Es el que se produce cuando se elige una muestra hospitalaria como control y el factor de riesgo que está estudiándose se asocia a una mayor probabilidad de hospitalización (Fig. 2). Veamos un ejemplo. Supongamos que queremos analizar la asociación entre el consumo de alcohol y el desarrollo de cirrosis hepática. Para ellos seleccionamos los casos, nuestros pacientes, de la planta de digestivo, preguntándoles sobre el antecedente de exposición al alcohol. Lógicamente la mayoría nos dirá que, en efecto, abusaron del alcohol. El problema es si seleccionamos los controles (sanos sin cirrosis) del entorno hospitalario, por ejemplo, de la planta de psiquiatría. A los pacientes allí ingresados les preguntamos sobre el antecedente de consumo de alcohol, de forma que es posible que también muchos tengan el antecedente de esta exposición. Muchos casos expuestos (lo habitual) y muchos controles expuestos (el problema) van a condicionar que se minimice la asociación entre alcohol y cirrosis, dada la igualdad entre casos y controles con la exposición al alcohol.

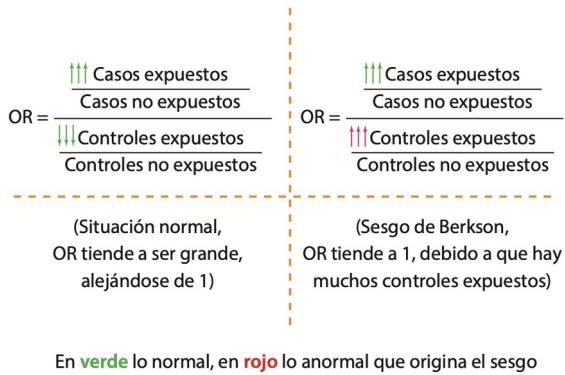


Figura 2. Influencia del sesgo de Berkson sobre la OR

Sesgo del obrero sano

La salida del trabajador enfermo del mercado laboral compromete la validez del estudio. Imaginemos que queremos estudiar si hay más riesgo de mesotelioma en trabajadores expuestos a aislantes como el asbesto. Podría darse la paradoja de que, si vamos a buscar enfermos a una fábrica en la que se emplea asbesto, no hubiera enfermos. Erróneamente concluiríamos que el asbesto no se asocia al mesotelioma, cuando realmente lo que ocurre es que no hay pacientes con mesotelioma trabajando en el momento de elegir los enfermos porque se encuentran de baja laboral.

Falacia de Neyman

Se produce en los estudios de casos-controles (Fig. 3) al seleccionar casos prevalentes (ya existentes) en vez de casos incidentes (casos nuevos). Esto conlleva que en los casos sea menos frecuente la exposición a aquellos FR que disminuyen la supervivencia; esto es, los casos representarían a individuos con rasgos de mayor resistencia a la enfermedad o a formas menos graves de esta. Por ejemplo, quiere averiguar la asociación entre el consumo de setas y el padecimiento de una hepatitis fulminante. Si seleccionamos del hospital pacientes con hepatitis fulminante con antecedente de consumo de setas, veremos que paradójicamente hay pocos. Esto se debe a la rápida letalidad de esta enfermedad. De nuevo encontraremos que se reduce la asociación entre el factor de riesgo y

la enfermedad, en este caso porque encontramos pocos casos expuestos (debido a la letalidad del cuadro clínico) y pocos controles expuestos (lo habitual).

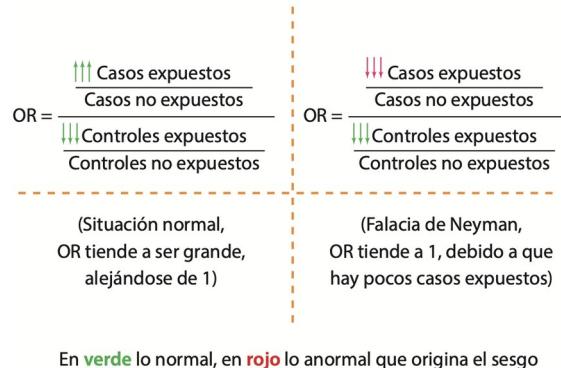


Figura 3. Influencia de la falacia de Neyman sobre la OR.

Pérdida de seguimiento

Es lógico pensar que este sesgo es habitual de los diseños prospectivos, en los que en uno de los grupos se pierden sujetos. El problema aparece cuando las pérdidas o los abandonos pueden influir en la variable final.

Supongamos que diseñamos un estudio para evaluar la asociación entre el tabaco y el desarrollo de cáncer de esófago. Para ello planteamos un estudio de cohortes en el que los expuestos son los fumadores y los no expuestos, los no fumadores. Les seguimos a lo largo del tiempo esperando la aparición de los casos incidentes. Lo que condiciona este sesgo es que el seguimiento es mucho más estrecho en el grupo de fumadores, de forma que no se nos escaparán los casos incidentes en ese grupo. Sin embargo, en los no fumadores una mínima disfagia puede pasarse por alto, con lo que se pierden casos (casos no seleccionando) en el grupo no expuesto.

SESGO DE INFORMACIÓN, CLASIFICACIÓN O DE MEMORIA

En el punto anterior explicábamos lo que sucedía cuando elegíamos de manera incorrecta o incompleta a los sujetos de estudio (sesgos de selección). Ahora vamos a ponernos en una situación diferente:

hemos seleccionado a nuestra muestra de manera perfecta, pero cometemos un error sistemático en el proceso de recogida de la información. Por tanto, definimos los sesgos de información o de clasificación (que también podemos encontrar en la literatura como *sesgos de observación*) como la incorrecta clasificación de los participantes de un estudio con respecto a las variables recogidas por el investigador.

Existen múltiples razones para que esta clasificación de los sujetos se desarrolle de manera errónea. Tres son las principales:

- Que exista un **fallo en el procedimiento de medida**. Puede ser un problema de calibración del instrumento, que la persona recuerde de manera errónea su exposición/enfermedad o que empleemos para la recogida de datos las historias clínicas (especialmente cuando no se encuentran informatizadas).
- Cuando **en lugar de la variable de interés empleamos variables sucedáneas, denominadas en epidemiología variables proxy**. A pesar de que aportan mayor comodidad al investigador, restan validez a los datos. Una pregunta: ¿de qué nos fiaríamos más a la hora de determinar la capacidad económica de una persona? ¿De cómo defina su nivel socioeconómico (variable *proxy*) o de la revisión que hagamos de sus activos bancarios y del resto de bienes? Si queremos conocer el peso de una persona, ¿se lo preguntamos (variable *proxy*) o utilizamos una báscula?
- **Si las variables no están definidas o lo están de una manera poco específica**. Cuando realizamos estudios que relacionan la obesidad con diversas patologías, definimos como “persona con factor de riesgo” a aquella con un IMC superior a 30 kg/m^2 , pues es el umbral que establece la Organización Mundial de la Salud. ¿Cuál es el umbral entre clases sociales? ¿Y entre razas? Como

vemos, existen variables en las que es difícil determinar un límite claro.

La clasificación incorrecta puede afectar o no de igual manera a todos los grupos de estudio. A continuación, y siempre a través de ejemplos, repasaremos los principales sesgos de clasificación e información.

Clasificación incorrecta diferencial

La probabilidad de clasificación errónea de la exposición (en un estudio de casos-controles) afecta de forma desigual a sanos y enfermos o la clasificación de enfermo o no enfermo (en un estudio de cohortes) se hace de manera distinta, según se esté o no expuesto al factor de estudio. Este tipo de sesgo produce una subestimación o una sobreestimación del efecto del factor de exposición en la enfermedad.

Sesgo diagnóstico. Sesgo de sospecha de exposición

Imaginemos que comienza un estudio de casos y controles en nuestro centro, del que somos investigadores principales (IP), en el que se pretende poner en valor la relación entre el consumo de determinados fármacos y los eventos cardiovasculares. Para ello, elegiremos como casos a los pacientes que ingresaron en el servicio de cardiología de nuestro centro con determinados diagnósticos en el último año. Los controles serán sanos y se obtendrá la información por entrevista telefónica. En los pacientes con diagnóstico (y que, por tanto, están incluidos en la historia clínica electrónica del hospital) repasamos de manera perfecta todas las prescripciones farmacológicas efectuadas durante el último año. Dado que muchos de los controles no han precisado ingreso hospitalario, nos limitaremos a preguntarles qué fármacos han tomado en el pasado, de forma que el antecedente de exposición al fármaco (factor de riesgo) afectará preferentemente a los controles.

Sesgo de memoria

El hecho de padecer la enfermedad hace que se esté más motivado para recordar posibles antecedentes de exposición. Vamos a situarnos en un caso alejado de las situaciones frecuentes de investigación, pero que reflejará a la perfección lo que queremos expresar: un accidente de tráfico. Imagínate que vives a 20 kilómetros del hospital en el que trabajas y necesitas vehículo propio para llegar todas las mañanas. Un día, a las 12:00 h, te pregunto cómo ha sido tu viaje hasta allí. Dado que lo haces todos los días, has perdido tu capacidad de sorpresa, por lo que no podrías darme ni un solo detalle particular de esa mañana. Ahora pongámonos en otra situación más desagradable: esa mañana, de camino al trabajo, hemos sufrido un golpe al entrar en una rotonda que ha supuesto la rotura del espejo retrovisor izquierdo. A las 12:00 h, de nuevo, te repito la pregunta: ¿cómo ha sido el viaje? Además de contarme que has tenido un accidente, me dirás la posición exacta por la que entraste en la rotonda, de qué color era el vehículo que iba delante, si el coche que te golpeó había dado o no los intermitentes, si en ese momento llovía o el sol te deslumbraba, etc.

En definitiva, los casos recuerdan siempre mejor la exposición que los controles, por lo que, de nuevo, se sobreestimará el efecto. El ejemplo clásico en epidemiología son los estudios de casos y controles y, dentro de ellos, los que relacionan el consumo de fármacos durante el embarazo con la aparición de malformaciones congénitas en el recién nacido. Las madres de niños que han presentado estas patologías recordarán mejor qué medicamentos consumieron durante ese periodo que las madres de niños sanos. Es decir, que puede que clasifiquemos como no expuestas a las mujeres de niños sanos (controles).

Sesgo de atención o efecto Hawthorne

Entre los años 1927 y 1937, en la Western Electric's Hawthorne, una central eléctrica situada en Chicago, y en un contexto norteamericano general de búsqueda constante de la productividad, se desarrollaron diversos estudios sobre la satisfacción de sus trabajadores. Tras la introducción de diversas mejoras laborales, vieron cómo se disparó la productividad. No obstante, un avisado investigador pensó que estaban mejorando en demasiada proporción para las pocas mejoras que la empresa había puesto a su disposición. ¿Qué ocurría? ¿Estaban mejorando por la nueva estrategia de la empresa? La respuesta es que no. Estaban mejorando porque, por primera vez, se sentían vigilados y observados. Lo mismo sucede en muchos de nuestros estudios, en los que los sujetos del estudio se sienten observados modificando su exposición, por ejemplo, por temor a represalias. Imaginemos que la inclusión o no en un ensayo clínico depende de que el paciente fume o no fume. Muy probablemente el sujeto continuará fumando, aunque nos diga que no. Esto implica que en nuestro estudio le clasifiquemos como no expuesto, cuando realmente sí lo es.

Sesgo de regresión a la media

Consiste en que una variable puede ser extrema en su primera medición por motivos fisiológicos, pero puede volver a la normalidad posteriormente. Por ejemplo, efecto de la bata blanca, se clasifica al sujeto expuesto, cuando realmente no lo es. Si la medición se repite en varias ocasiones, la tensión arterial de los sujetos tenderá a la normalización. Es decir, que una única medición nos clasificará erróneamente como hipertensos a sujetos que realmente no lo son.

Clasificación incorrecta no diferencial

La probabilidad de clasificación errónea ocurre en todos los grupos de estudio de forma similar. El error de clasificación no diferencial ocasiona una infraestimación (hace que la medida de asociación tienda a 1) del efecto del factor de exposición estudiado en la

enfermedad. Imaginemos que queremos evaluar la asociación entre hipertensión e ictus. Para ello elegimos un manguito de tensión de nuestra consulta y dividimos a la muestra en hipertensos y no hipertensos en función de una diastólica mayor o menor de 90 mmHg. ¿Qué pasaría si el manguito mide a veces por error valores altos y a veces valores bajos? El resultado será que en el grupo de hipertensos se colará algún normotenso y en el de normotensos, algún hipertenso. Esto iguala ambos grupos, con lo que se infraestima la asociación entre FR y enfermedad.

FACTOR DE CONFUSIÓN

Uno de los factores más importantes debido a su presencia en la investigación es sin duda el concepto de *confusión*, definido como la distorsión del efecto estimado en la exposición sobre la enfermedad debida a la interposición de un efecto nuevo producido por un factor extraño (factor de confusión) que no quiere estudiarse y que origina una mezcla de efectos. Es decir, tenemos una variable extraña al estudio que modifica los resultados que se obtienen. Todo factor de confusión debe cumplir tres condiciones (vemos como ejemplos entre *factor de riesgo tabaco*, *factor confusor alcohol* y *enfermedad cáncer de esófago*):

- Ser un factor de riesgo para la enfermedad. El alcohol, factor confusor, se asocia a la enfermedad.
- Estar asociado a la exposición. Es más frecuente el alcohol en fumadores que en no fumadores.
- No ser un paso intermedio entre la exposición y la enfermedad. El alcohol no es un nexo entre el tabaco y el cáncer.

A pesar de suponer un error con gran trascendencia en nuestros resultados, el sesgo de confusión puede cuantificarse (a diferencia de los sesgos de selección e información). Para ello, compararemos directamente la estimación cruda o bruta del efecto (OR_b , *en el total de la muestra*) con la estimación sin la confusión (OR_e , *en el grupo*

estratificado). Llegaríamos a la magnitud de la confusión a partir de la fórmula:

$$\frac{OR_e - 1}{OR_b - 1}$$

Imaginemos que, para la asociación entre un factor de riesgo y una enfermedad, se obtiene una OR de 3. Al estratificar en dos categorías en función de una variable que podría estar distorsionando el efecto, se obtienen dos OR de 2. Aplicando la fórmula:

$$\frac{OR_e - 1}{OR_b - 1} = \frac{2 - 1}{3 - 1} = \frac{1}{2} = 0,5$$

La magnitud de la confusión es de un $(1 - 0,5) \times 100 = 50\%$ del efecto. Es decir, el efecto verdadero o real, una vez controlada la confusión, es un 50 % menor (justo la mitad) que el estimado mediante el análisis bruto, que sí se encontraba distorsionado. Consideramos que la confusión es suficientemente importante como para tener que controlarla cuando la OR cruda difiere de la ajustada (o en nuestro caso, de la OR_e) en más del 10 %, como es el caso que nos ocupa.

Para prevenir los sesgos de confusión existen distintas técnicas, aplicables tanto en la fase de diseño como en la de análisis estadístico. Esto supone una de las grandes ventajas de este error sistemático: es controlable *a posteriori*. En la fase de diseño encontramos la aleatorización (estudios experimentales), el apareamiento y la restricción. En la fase de análisis estadístico: análisis estratificado (se dividen los datos globales en dos grupos, según la presencia o no de factor de confusión, como veremos más adelante) y análisis multivariante o técnicas de regresión, que se escapan del alcance que queremos tratar en este artículo. Es importante asumir que en la metodología de la investigación la

perfección tampoco existe. Siempre existirá *confusión residual*, entendida como aquella que persiste tras intentos fallidos por controlarla.

Íntimamente relacionadas con la confusión, deben conocerse otras de las modificaciones del efecto estimado: la **interacción**. A diferencia del factor de confusión, la interacción no modifica el efecto estimado de manera errónea. Entre los principales puntos en común destaca la necesidad de valorar variables externas para alcanzar una interpretación correcta en nuestro estudio.

Veamos todo lo anteriormente expuesto con nuestro ejemplo (Fig. 4).

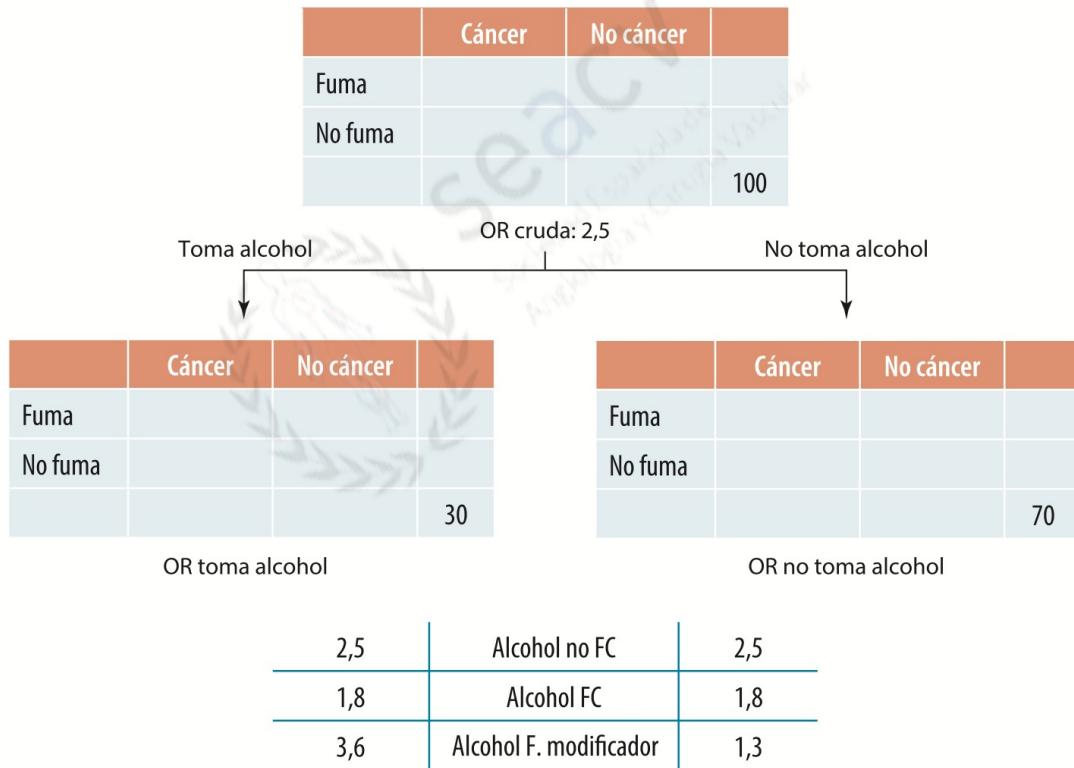


Figura 4. Análisis estratificado.

Queremos estudiar la asociación entre el cáncer de esófago y el tabaquismo. Para ello, diseñamos una tabla de 2×2 con las variables “cáncer (sí / no)” y “tabaco (fuma / no fuma)”. Tras realizar el cociente de productos cruzados, obtenemos una OR, que llamaremos OR cruda, de 2,5.

No obstante, nos inquietamos al repasar la bibliografía, pues muchos estudios comentan la influencia que el consumo de alcohol puede tener en esta relación. Puede ser un factor confusor en nuestro estudio.

Para salir de dudas decidimos estratificar según los sujetos tengan o no el supuesto factor de confusión (en este caso, el consumo de alcohol). Obtenemos dos nuevas tablas que pueden hacernos llegar a tres situaciones diferentes:

- Si la OR cruda que habíamos calculado se mantiene al estratificar (OR 2,5 en el grupo que toma alcohol y OR 2,5 en el grupo que no toma alcohol), significa que no está sucediendo nada. Dicho de un modo más metodológico: el alcohol no es un factor de confusión ni modificador del efecto en nuestro estudio.
- Si las nuevas OR, una para cada estrato, son de 1,8 (es decir, inferiores a la OR cruda que habíamos calculado al inicio), nos encontramos ante un claro factor de confusión. Tendremos que poner remedio o nuestro estudio estará seriamente comprometido. Poniendo en práctica todo lo aprendido vamos a cuantificar el problema mediante la fórmula:

$$\frac{OR_e - 1}{OR_b - 1} = \frac{1,8 - 1}{2,5 - 1} = \frac{0,8}{1,5} = 0,53$$

La magnitud de la confusión es de un $(1 - 0,53) \times 100 = 47\%$ del efecto. Es decir, el efecto verdadero o real, una vez controlada la confusión, es un 47 % menor que el estimado mediante el análisis bruto, que sí se encontraba distorsionado. Al ser superior al límite consensuado del 10 %, consideraremos que el factor de confusión está invalidando los resultados.

- Si las nuevas OR son inferiores a la OR cruda, pero con diferentes valores en ambos grupos, nos encontramos ante un factor modificador. Recapitulando, hablábamos de interacción cuando el efecto de la exposición sobre la enfermedad es

diferente a distintos niveles de una tercera variable. En nuestro ejemplo, diremos que el tabaco siempre actúa como factor de riesgo, ya que el OR es mayor de 1 en los 2 estratos. Sin embargo, afirmaremos que los alcohólicos tienen 3,6 veces más de riesgo por tabaco, mientras que los no alcohólicos tienen sólo 1,3 veces más riesgo de cáncer por tabaco. Como vemos, el alcohol modifica el riesgo de cáncer debido al tabaco.

Si nuestro lector ha sido capaz de leer todo lo anterior es que realmente está interesado en este tema y deberá acudir a libros en los que se tratan los errores metodológicos de forma más desarrollada.

En este artículo hemos querido dar solo unas pinceladas de algunas de las situaciones que más frecuentemente se plantean en la rutina de un aspirante a investigador.

