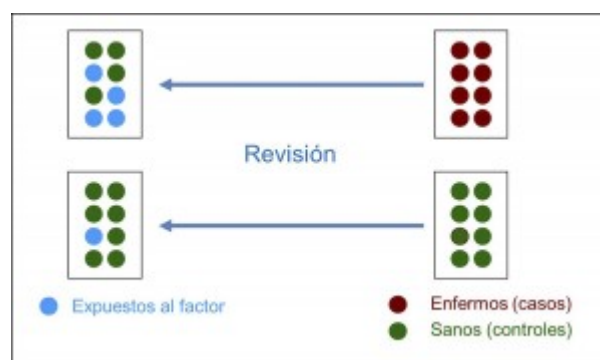


# De la gallina al huevo

Seguro que alguna persona que rebosaba ingenio a raudales os ha preguntado en alguna ocasión, con mirada de suficiencia, ¿qué fue antes, el huevo o la gallina? Pues bien, la próxima vez que os encontréis con alguien así podéis responder con otra pregunta: ¿es qué tienen algo que ver el huevo y la gallina? Porque primero habrá que saber, no solo si para tener gallinas hay primero que tener huevos (con perdón), sino también qué probabilidad hay de acabar teniéndolas, con huevos o sin ellos (alguna mente retorcida dirá que la pregunta se podría plantear al revés, pero es que yo soy de los que piensan que lo primero que hay que tener, sin ánimo de ofender, son huevos).

Este planteamiento nos llevaría al diseño de un **estudio de casos y controles**, que es un estudio **observacional** y **analítico** en la que el muestreo se hace en base a presentar una determinada enfermedad o efecto (los casos) y se compara ese grupo con otro grupo que no lo presenta (los controles), con el objetivo de determinar si existe diferencia en la frecuencia de exposición a un determinado factor de riesgo entre los dos grupos. Estos estudios son de **direccionalidad retrógrada** y de **temporalidad mixta**, por lo que la mayor parte son de naturaleza retrospectiva aunque, al igual que ocurría con los estudios de cohortes, pueden ser también prospectivos (quizás la clave más útil para distinguir entre los dos sea el muestreo de cada uno, en base a la exposición en los estudios de cohortes y en base al efecto en los de casos y controles).

En la figura que os adjunto podéis ver el diseño típico de un estudio de casos y controles. En estos estudios se parte de una población determinada de la que se extrae una muestra de casos que habitualmente suele englobar todos los casos diagnosticados y disponibles, y se comparan con un grupo control formado por una muestra balanceada de sujetos sanos que proceden de la misma población que los sanos. Sin embargo, cada vez es más frecuente encontrarse con variaciones del diseño básico que combinan características de los estudios de cohortes y de casos y controles, comparando los casos que van apareciendo en una cohorte estable a lo largo del tiempo con controles de una muestra parcial extraída de esa misma cohorte.



El más conocido de este tipo de diseños mixto es el de **casos y controles anidado en una cohorte**. En estos casos partimos de una cohorte ya conocida en la que vamos identificando los casos que se van produciendo. Cada vez

que aparece un caso, se empareja con uno o varios controles sacados también de la cohorte inicial. Si lo pensamos brevemente, es posible que un sujeto que sea seleccionado inicialmente como control se convierta en caso a lo largo del tiempo (desarrolle la enfermedad en estudio). Aunque pueda parecer que esto pueda sesgar los resultados, esto no debe ser así, ya que se trata de medir el efecto de la exposición en el momento de realizar el análisis. Este diseño puede hacerse con cohortes más pequeñas, por lo que puede ser más sencillo y económico. Además, es especialmente útil en cohortes muy dinámicas con muchas entradas y salidas a lo largo del tiempo, sobre todo si la incidencia de la enfermedad en estudio es baja.

Otra variante del diseño básico son los [estudios de cohorte y casos](#). En este, inicialmente tenemos una cohorte muy grande de la que seleccionaremos una subcohorte más pequeña. Los casos serán los enfermos que se vayan produciendo en cualquiera de las dos cohortes, mientras que los controles serán los sujetos de la subcohorte más pequeña (y más manejable). Estos estudios tienen un método de análisis un poco más complicado que los diseños básicos, ya que tienen que compensar el hecho de que los casos se ven sobrerrepresentados al poder provenir de las dos cohortes. La gran ventaja de este diseño es que permite estudiar varias enfermedades a la vez, comparando las diversas cohortes de enfermos con la subcohorte elegida como control.

Por fin, una última variación que vamos a tratar es la de los polisémicos [estudios de caso-cohorte](#), también llamados de casos y controles cruzados, también llamados de casos autocontrolados. En este diseño pareado, cada individuo sirve como su propio control, comparándose la exposición durante el periodo de tiempo más cercano a la aparición de la enfermedad (periodo caso) con la exposición durante el periodo de tiempo previo (periodo control). Este planteamiento de estudio es útil cuando la exposición es corta, con un tiempo de actuación previsible y produce una enfermedad de corta duración en el tiempo. Son muy utilizados, por ejemplo, para estudiar los efectos adversos de las vacunas.

Al igual que en los estudios de cohortes, los estudios de casos y controles permiten el cálculo de toda una serie de medidas de asociación e impacto. Claro que aquí nos encontramos con una diferencia fundamental con los estudios de cohortes. En estos partíamos de una cohorte sin enfermos en la que iban apareciendo los enfermos a lo largo del seguimiento, lo que nos permitía calcular el riesgo de enfermar a lo largo del tiempo (incidencia). Así, el cociente entre incidencias de expuestos y no expuestos nos proporcionaba el riesgo relativo, la principal medida de asociación.

Sin embargo, como puede deducirse del diseño de los estudios de casos y controles, en estos casos no podemos hacer una estimación directa ni de la incidencia ni de la prevalencia de la enfermedad, ya que la proporción de

expuestos y enfermos viene determinada por los criterios de selección del investigador y no por la incidencia en la población (se seleccionan de entrada un número fijo de casos y de controles, pero no podemos calcular el riesgo de ser caso en la población). Así, ante la imposibilidad de calcular el riesgo relativo recurriremos al cálculo de la odds ratio (OR), tal como véis en la segunda figura.

	Casos	Controles
Expuestos	a	b
No expuestos	c	d

Odds (casos) =  $a / c$   
 Odds (controles) =  $b / d$   
 OR =  $(a / c) / (b / d) = (a \times d) / (b \times c)$

La OR tiene una lectura similar a la del riesgo relativo, pudiendo valer desde cero hasta infinito. Una OR=1 quiere decir que no hay asociación entre exposición y efecto. Una OR<1 quiere decir que la exposición es un factor de protección frente al efecto. Por último, una OR>1 indica que la exposición es un factor de riesgo, tanto mayor cuanto mayor sea el valor de la OR.

De todos modos, y solo para los que gusten de meterse en complicaciones, os diré que es posible calcular las tasas de incidencia a partir de los resultados de un estudio de casos y controles. Si la incidencia de la enfermedad en estudio es baja (por debajo del 10%), pueden equipararse OR y riesgo relativo, así que podemos estimar la incidencia de forma aproximada. Si la incidencia de la enfermedad es mayor, la OR tiende a sobreestimar el riesgo relativo, así que no podemos equipararlos. De todas formas, en estos casos, si conocemos previamente la incidencia de la enfermedad en la población (obtenida de otro tipo de estudios), podemos calcular la incidencia utilizando las siguientes fórmulas:

$$I_0 = I_t / (OR \times P_e) + P_0$$

$$I_e = I_0 \times OR,$$

donde  $I_t$  es la incidencia total,  $I_e$  la incidencia en expuestos,  $I_0$  la incidencia en no expuestos,  $P_e$  la proporción de expuestos y  $P_0$  la proporción de no expuestos.

Aunque la OR permite estimar la fuerza de la asociación entre la exposición y el efecto, no informa sobre el efecto potencial que tendría eliminar la exposición sobre la salud de la población. Para ello, tendremos que recurrir a las medidas de riesgo atribuible (tal como hacíamos con los estudios de cohortes), que pueden ser absolutas o relativas.

Las medidas absolutas de riesgo atribuible son dos. La primera es el

**riesgo atribuible en expuestos (RAE)**, que es la diferencia entre la incidencia en expuestos y no expuestos y representa la cantidad de incidencia que puede ser atribuida al factor de riesgo en los expuestos. La segunda es el **riesgo atribuible poblacional (RAP)**, que representa la cantidad de incidencia que puede ser atribuida al factor de riesgo en la población general.

Por su parte, las medidas relativas de riesgo atribuible (también conocidas como proporciones o fracciones atribuibles o etiológicas) son también dos. La primera, la **fracción atribuible en expuestos (FAE)**, que representa la diferencia de riesgo relativo a la incidencia en el grupo de expuestos al factor. La segunda, la **fracción atribuible poblacional (FAP)**, que representa la diferencia de riesgo relativo a la incidencia en la población general.

- **Diferencias absolutas de riesgo**

- **Riesgo atribuible en expuestos (RAE)**

- Puede calcularse igual a  $lo \times (RR-1)$ , usando la OR como aproximación al RR (solo válido para valores de incidencia bajos). Es una mala aproximación.
    - Asumiendo que los controles son representativos de la población general y que la incidencia de la enfermedad es baja:

$$RAE = lo(OR-1)/Pcte(OR-1)+1.$$

lo = incidencia de los no expuestos en la población general, que generalmente no se conoce. Se puede aproximar por la Incidencia en la población general, normalmente de estudios previos.

Pcte = proporción de controles expuestos.

- **Riesgo atribuible poblacional (RAP)**

$$RAP = RAE \times Pe$$

- **Diferencias relativas de riesgo**

- **Fracciones atribuibles.**

- Fracción atribuible en los expuestos (FAE)

$$FAE = OR-1/OR.$$

- Fracción atribuible en la población (FAP)

$$FAP = FAE \times Pce$$

Pce = proporción casos expuestos

- **Fracciones prevenibles.**

- Fracción prevenible en los expuestos (FPE)

$$FPE = 1-OR$$

- Fracción prevenible en la población (FPP)

$$FPP = FPE \times Pce$$

Pce = proporción casos expuestos

En la tabla adjunta os muestro las fórmulas para el cálculo de estos parámetros, que es algo más complejo que en el caso de los estudios de cohortes.

El problema de estas medidas de impacto es que pueden ser, en ocasiones, difíciles de interpretar por parte del clínico. Por ese motivo, e inspirados en el cálculo del número necesario a tratar (NNT) de los ensayos clínicos, se han ideado una serie de medidas denominadas números de impacto, que nos dan una idea más directa del efecto del factor de exposición sobre la enfermedad en estudio. Estos números de impacto son el **número de impacto en expuestos (NIE)**, el **número de impacto en casos (NIC)** y el **número de impacto de los casos expuestos (NICE)**.

El NIE sería el equivalente al NNT y se calcularía como el inverso de la reducción absoluta de riesgos o de la diferencia de riesgos entre expuestos

y no expuestos. El NNT es el número de personas que deben ser tratadas para prevenir un caso en comparación con el grupo control. El NIE representa el número medio de personas que tienen que exponerse al factor de riesgo para que se produzca un nuevo caso de enfermedad en comparación con las personas no expuestas. Por ejemplo, un NIE de 10 significa que de cada 10 expuestos se producirá un caso de enfermedad atribuible al factor de riesgo estudiado.

El NIC es el inverso de la FAP, así que define el número medio de personas enfermas entre las que un caso es debido al factor de riesgo. Un NIC de 10 quiere decir que por cada 10 enfermos de la población, uno es atribuible al factor de riesgo en estudio.

Por fin, el NICE es el inverso de la FAE. Es el número medio de enfermos entre los que un caso es atribuible al factor de riesgo.

En resumen, estas tres medidas miden el impacto de la exposición entre todos los expuestos (NIE), entre todos los enfermos (NIC) y entre todos los enfermos que han estado expuestos (NICE). Será de utilidad que intentemos calcularlos si los autores del estudio no lo hacen, ya que nos darán una idea del impacto real de la exposición sobre el efecto.

Como colofón a los tres anteriores, podríamos estimar el efecto de la exposición en toda la población calculando el [número de impacto en la población](#) (NIP), para lo cual no tenemos más que hacer el inverso del RAP. Así, un NIP de 3000 quiere decir que por cada 3000 sujetos de la población se producirá un caso de enfermedad debida a la exposición.

Además de la valoración de las medidas de asociación e impacto, cuando valoremos un estudio de casos y controles tendremos que prestar especial atención a la presencia de sesgos, ya que son los estudios observacionales que tienen más riesgo de presentarlos.

Los estudios de casos y controles son relativamente sencillos de hacer, generalmente tienen menor coste que otros estudios observacionales (como los estudios de cohortes), permiten estudiar varios factores de exposición al mismo tiempo y saber cómo interactúan entre ellos y son ideales para enfermedades o factores de exposición de frecuencia muy baja. El problema de este tipo de diseño es que hay que ser sumamente cuidadoso para seleccionar los casos y los controles, ya que, como ya hemos dicho, es muy fácil caer en una lista de sesgos que, a día de hoy, no tiene aún un final conocido.

En general, los criterios de selección deberían ser los mismos para casos y controles, pero, como para ser caso hay que estar diagnosticado de la enfermedad y estar disponible para el estudio, es muy probable que los casos no sean totalmente representativos de la población. Por ejemplo, si

los criterios de diagnóstico son poco sensibles y específicos habrá muchos falsos positivos y negativos, con lo que el efecto de la exposición al factor se diluirá.

Otro posible problema depende de que elijamos casos incidentes (de nuevo diagnóstico) o prevalentes. Los estudios basados en prevalencia favorecen la selección de supervivientes (hasta ahora no se conoce ningún caso de un muerto que haya accedido a participar en ningún estudio) y, si la supervivencia está relacionada con la exposición, el riesgo detectado será menor que con casos incidentes. Este efecto es todavía más evidente cuando el factor de exposición es de buen pronóstico, situación en la que los estudios prevalentes producen una mayor sobreestimación de la asociación. Un ejemplo para comprender mejor estos aspectos: supongamos que el riesgo de infarto es mayor cuanto más se fuma. Si solo incluimos casos prevalentes excluiríamos a los muertos por infarto más grave, que presumiblemente deberían ser los que más fumasen, con lo cual el efecto del tabaco podría infraestimarse.

Pero si lo de los casos parece complicado, no es nada comparado con una buena selección de los controles. Lo ideal es que los controles hayan tenido la misma probabilidad de exposición que los casos o, dicho de otra forma más elegante, deben ser representativos de la población de la que proceden los casos. Además, esto hay que compatibilizarlo con la exclusión de aquéllos que tengan alguna enfermedad que se relacione de forma positiva o negativa con el factor de exposición. Por ejemplo, si nos sobra tiempo y queremos ver la asociación entre pasajeros de avión que tienen una tromboflebitis y la ingesta previa de aspirina, tendremos que excluir de los controles a los que tengan cualquier otra enfermedad que se trate con aspirina, aunque no la hubiesen tomado antes de emprender el viaje.

También hay que ser crítico con algunos hábitos de selección de controles. Por ejemplo, los pacientes que van al hospital por otro motivo distinto al estudiado están muy a mano, suelen ser muy colaboradores y, al ser enfermos, seguramente recordarán mejor las exposiciones pasadas a factores de riesgo. Pero el problema es ese, que son enfermos, por lo que pueden tener hábitos de exposición a factores de riesgo diferentes a los de la población general.

Otro recurso es reclutar a vecinos, amigos, familiares, etc. Éstos suelen ser muy comparables y colaboradores, pero tenemos el riesgo de que haya emparejamiento de hábitos de exposición que nos alteren los resultados del estudio. Todos estos problemas se evitan tomando los controles de la población general, pero esto es más costoso en esfuerzo y dinero, suelen ser menos colaboradores y, sobre todo, mucho más olvidadizos (los sanos recuerdan menos la exposición a factores de riesgo pasados), con lo que la calidad de la información que obtengamos de casos y controles puede ser muy

diferente.

Solo una reflexión más para terminar con este tema tan ameno. Los estudios de casos y controles comparten una característica con el resto de los estudios observacionales: detectan la asociación entre la exposición y el efecto, pero no nos permiten establecer con seguridad relaciones de causalidad, para lo cual necesitamos otro tipo de estudios como los ensayos clínicos aleatorizados. Pero esa es otra historia...

---