

Estudios analíticos observacionales (cohortes y casos-contróles): investigando asociaciones causales



¹A. Gallego Iborra, ²G. Moreno Muñoz, G. del Castillo Aguas³

¹Pediatra. CS Trinidad. Málaga. España.

²Pediatra. CS Puerto de la Torre. Málaga. España.

³Pediatra. CS Carihuela. Torremolinos, Málaga. España.

Miembros del Grupo de Investigación de la AEPap

RESUMEN

Cuando la pregunta de investigación haga referencia a la causa, cuando se intente evaluar la relación causa-efecto, deberemos usar estudios analíticos. Los estudios de cohortes y los de casos-contróles son estudios analíticos observacionales longitudinales. En las **cohortes o estudios de seguimiento** partiremos de dos grupos de individuos sanos: un grupo estará expuesto a un factor de riesgo y el otro no, y seguiremos a las dos cohortes a lo largo del tiempo, midiendo la aparición de la enfermedad en los dos grupos. Por el contrario, en un estudio de **casos-contróles** se identifica a un grupo de personas con una enfermedad (casos) y se compara con un grupo apropiado que no tenga la enfermedad (contróles); si en el grupo de enfermos hubiera mayor exposición a un determinado factor que en el grupo control, podremos decir que hay una asociación entre el factor y el efecto. Generalmente, las asociaciones causales pueden determinarse de forma más probable con los estudios de cohortes, pero en ocasiones deberemos recurrir a los de casos-contróles. Ambos tipos de estudio tienen sus limitaciones, sus posibles sesgos y sus propias medidas de asociación.

PUNTOS CLAVE

- Cohortes y casos-contróles son estudios analíticos observacionales longitudinales.
- Evalúan la relación causa-efecto. Se usarán cuando la pregunta haga referencia a la etiología de un evento.
- Los estudios de cohortes parten de individuos sanos que están expuestos o no a un posible factor de riesgo.
- Los estudios de cohortes permiten medir la incidencia en expuestos y en no expuestos, y a partir de ahí calcular el riesgo relativo, así como estudiar la historia natural de la enfermedad.
- Los estudios de cohortes son costosos, largos en el tiempo, las pérdidas son frecuentes y no son útiles para estudiar enfermedades raras o con largos periodos de latencia.
- Los estudios de casos-contróles parten de un grupo de individuos enfermos (casos), y se busca un grupo de individuos sanos (contróles) para comparar la posible presencia de determinados factores de riesgo. Los contróles deben ser comparables a los casos.
- En los estudios de casos-contróles no se puede estimar la incidencia, estimándose la magnitud del riesgo por medio de la *odds ratio*.
- Son más rápidos y baratos que los estudios de cohortes, permiten estudiar enfermedades poco frecuentes y pueden evaluar múltiples factores de riesgo para una misma enfermedad.
- Los problemas más importantes en los estudios de casos-contróles es encontrar un grupo control adecuado y los sesgos de memoria.

¿CUÁNDO Y CÓMO PODEMOS DISEÑAR UN ESTUDIO DE COHORTES?

El uso de un diseño de cohortes nos proporciona información sobre causalidad y sobre factores de riesgo. Lo usaremos, por tanto, cuando la pregunta de investigación haga referencia a: “¿a qué se debe...?” “¿el factor “x” favorece la aparición de tal fenómeno?”.

Estos diseños se clasifican según los criterios taxonómicos clásicos como estudios observacionales, analíticos y longitudinales, normalmente (pero no siempre) prospectivos.

Y consisten en el **seguimiento** de una o más **cohortes** de individuos **sanos** que presentan diferentes grados de **exposición** a un **factor de riesgo** en quienes se mide la aparición de la **enfermedad o condición de estudio**.

Antes de plantear un estudio de este tipo, debemos recoger información sobre eventuales factores de riesgo para el suceso de estudio, elegir el más relevante o factible y, sobre la base de este, delimitar las cohortes. Una de ella tendrá el factor de riesgo y la otra no. Por definición, en ninguno de los dos grupos debe haber individuos con la enfermedad en estudio o resultado.

Por ejemplo: “¿Se asocia el bajo peso al nacimiento con trastornos en el habla?”.

Se elegirá una cohorte con bajo peso al nacimiento (variable independiente o factor de exposición) y otra con peso normal. Se les sigue durante un tiempo, pongamos tres años, y al final se compara la existencia de trastornos del habla (variable dependiente o resultado) en uno y otro grupo.

Lo deseable sería que ambos grupos, expuestos y no expuestos, fueran comparables, es decir tuvieran características similares; lo ideal sería (aunque resulta difícil) que la única diferencia entre los dos grupos fuera el factor de exposición en estudio (en nuestro ejemplo, el peso al nacimiento). De cualquier forma, caracterizaremos los dos grupos describiendo en ellos la distribución de variables que en teoría podrían influir en el factor de resultado (por ejemplo: perímetro craneal al nacimiento, adquisición de hitos motores, aparición de balbuceo, nivel formación de los padres, etc.).

Recapitulando, **puntos importantes** en un estudio de cohortes:

- Los sujetos están libres del evento de interés al iniciarse el estudio.
- Se clasifican en función de su exposición o no al/los factor/es de estudio.

- Se compara la frecuencia de aparición del evento entre las cohortes.

Tipos de estudios de cohortes

En función del tiempo y del momento de observación, se describen tres tipos:

- **Prospectivos**: diseño clásico, ya descrito. El estudio comienza a la vez que la exposición.
- **Retrospectivo**: la exposición y el efecto ya han ocurrido cuando se inicia el estudio. La población expuesta y el momento de inicio de la exposición deben ser claramente definibles a pesar del paso del tiempo.
- **Ambispectivos**: el factor de exposición se inició antes del inicio de estudio, pero todavía no ha aparecido la enfermedad o resultado.

En función del número de cohortes, se describen:

- **Cohorte única**: existe una única cohorte de expuestos y se utiliza, como referencia, información registrada procedente de la población general. Se conoce también como “antes y después”. Por ejemplo, después de un accidente nuclear, medir la aparición de casos nuevos (incidencia) de una determinada patología, comparándola con la incidencia previa al evento.
- **Dos cohortes**: corresponde al formato clásico ya descrito, en el cual se seleccionan dos grupos de individuos sanos que solo difieren en que uno presenta el factor de exposición (cohorte expuesta) y el otro no (cohorte no expuesta).
- **Cohortes múltiples**: hacemos varias cohortes de individuos sanos en función del grado de exposición a un determinado factor de riesgo; la cohorte control será la de menor o nula exposición y con esta se comparará cada una de las otras. Sirve para relacionar el grado de exposición con las tasas de incidencia (relación dosis-respuesta).
- **Casos-controlados anidados**: al finalizar un estudio de cohortes tendremos un grupo de individuos en los que ha aparecido la enfermedad, individuos que ya están analizados y estudiados por el investigador. Constituyen un grupo de casos (anidado en una cohorte), y le buscamos un grupo control seleccionado aleatoriamente entre los que no desarrollaron la enfermedad. Se trata de un diseño híbrido que contribuye al control de sesgos, particularmente los de medición (memoria), frecuentes en los casos-controlados (**Figura 1**).

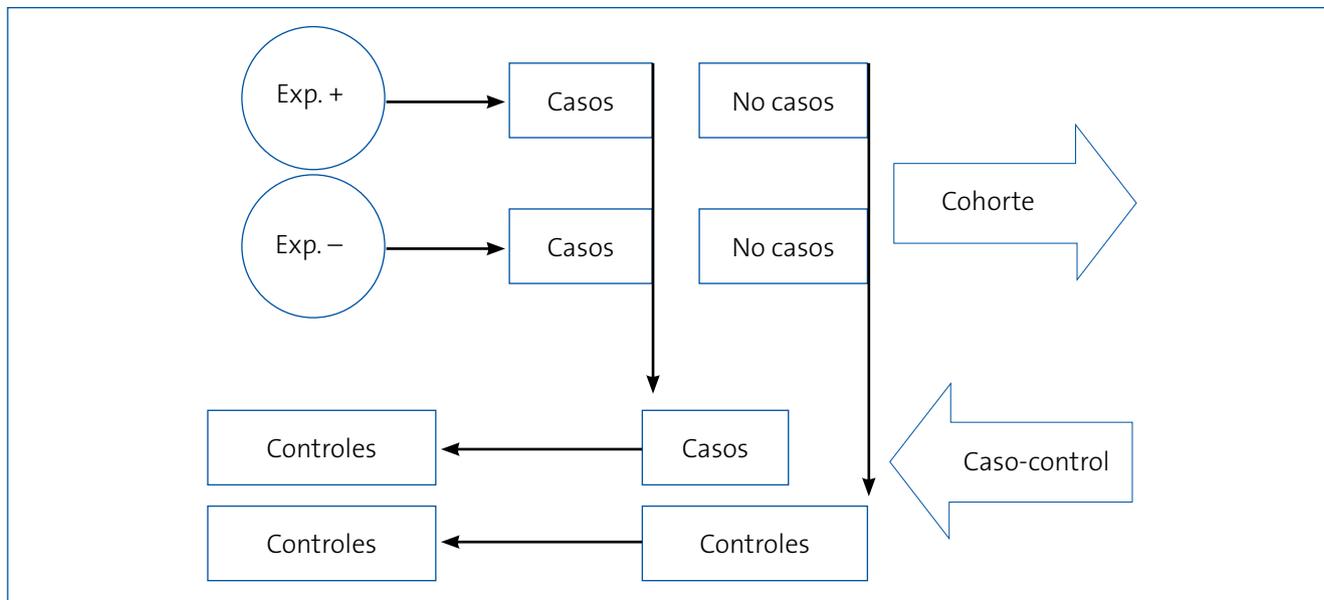


Figura 1. Esquema de un caso-control anidado en una cohorte

Datos obtenibles de un estudio de cohortes

1. Plantear hipótesis de causalidad y de riesgo: son, entre los estudios observacionales, los que proporcionan información más rigurosa sobre la causalidad y el riesgo. Sigue el mismo sentido de la lógica del pensamiento clínico (la exposición precede al desenlace).

La cuantificación de esta asociación la podemos calcular construyendo una razón entre la incidencia del fenómeno en los expuestos a la variable (I_e) y la incidencia del fenómeno en los no expuestos (I_o). Esta razón entre incidencias se conoce como riesgo relativo (Tabla 1).

2. Medir la incidencia de una enfermedad: la tasa de incidencia es el total de casos nuevos que se presentan en un grupo de individuos en un periodo determinado de tiempo. Como al principio del estudio todos los sujetos están sanos, podremos calcu-

lar la incidencia en expuestos, en no expuestos, y en ambos grupos en conjunto.

3. Estudiar la historia natural de la enfermedad.
4. Estudiar más de una consecuencia para la salud derivada de una exposición (un factor de riesgo puede tener varias consecuencias).

Desventajas de los estudios de cohortes (Tabla 2)

- No son eficientes para el estudio de enfermedades poco frecuentes.
- No sirven para el estudio de enfermedades con un largo periodo de latencia.
- Requieren un número muy elevado de participantes.
- Tienen un coste elevado.
- Son frecuentes las pérdidas.

Tabla 1. Cálculo de RR

	Enfermos	Sanos	
Expuestos	a	b	a + b
No expuestos	c	d	c + d
Total	a + c	b + d	a + b + c + d

$$\text{Riesgo relativo} = \frac{\text{Incidencia en expuestos}}{\text{Incidencia en no expuestos}} = \frac{I_e}{I_o} = \frac{a/(a+b)}{c/(c+d)}$$

Tabla 2. Ventajas y limitaciones de estudios de cohortes

Ventajas	Limitaciones
<ul style="list-style-type: none"> • Estiman incidencia • Menor posibilidad de sesgos en la medición de la exposición 	<ul style="list-style-type: none"> • Coste elevado • Dificultad en la ejecución • No son útiles en enfermedades raras • Requieren generalmente un tamaño muestral elevado • El paso del tiempo puede introducir cambios en los métodos y criterios diagnósticos • Posibilidad de pérdida en el seguimiento

¿Cuáles son los sesgos más frecuentes?

- **Sesgo de desgaste:** pérdidas de sujetos a lo largo del tiempo. Habrá que estudiar si es posible el motivo de la pérdida, si tienen relación con el factor de estudio, y si las pérdidas han sido similares en los dos grupos.
- **Sesgo de detección:** la variable resultado ha de estar perfectamente definida, pero además debe ser medida y medible en los dos grupos por igual.
- **Sesgo de selección por muestreo inadecuado** (voluntarios, población laboral/población general – efecto del trabajador sano–).
- **Sesgo de mala clasificación:**
 - Mala clasificación de la exposición: puede considerar exposiciones superiores o inferiores (hábitos socialmente aceptables o inaceptables) en alguna de las cohortes, esto llevaría a diluir el efecto del factor de riesgo. Tiene importancia si no se encuentra asociación significativa.
 - Mala clasificación del efecto/enfermedad: si la vigilancia diagnóstica en el grupo de expuestos es mayor que en el no expuesto se puede sobreestimar el efecto del factor de riesgo al detectar más enfermedad en los expuestos.

¿CUÁNDO Y CÓMO DEBEMOS DISEÑAR UN ESTUDIO DE CASOS-CONTROLES?

Cuando la enfermedad que queremos estudiar es poco frecuente, tenemos poco tiempo y pocos recursos o bien no tenemos una idea clara de los factores etiológicos, deberemos optar por diseñar un estudio de casos-controles (Figura 2).

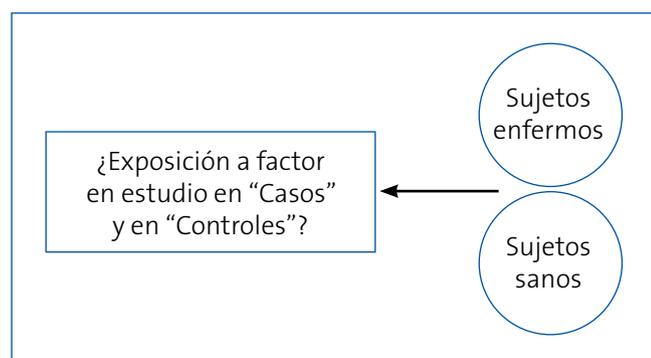


Figura 2. Estructura básica de un diseño de casos-controles

Selección de casos

Para la identificación de los casos, es imprescindible un criterio diagnóstico muy claro y preciso que no plante dudas en cuanto a la inclusión o exclusión del sujeto en el estudio. Los casos se pueden obtener de hospitales, centros de salud, o de cualquier institución donde pueda haber un registro, incluso de la población general.

Se pueden utilizar dos tipos de casos:

- **Casos incidentes:** se incluyen los nuevos casos que se van diagnosticando a partir del comienzo del estudio (**carácter prospectivo**).
- **Casos prevalentes:** se utilizan los casos que ya existían cuando comenzó el estudio (**carácter retrospectivo**).

En enfermedades cuyo desenlace puede ser la muerte, es preferible utilizar los casos incidentes, ya que en el momento del diagnóstico están presentes las formas graves y menos graves de la enfermedad, evitando así el **sesgo de supervivencia o falacia de Neyman**. Ejemplo de este sesgo es un estudio de casos-controles sobre leucemia aguda linfoblástica en el que los casos prevalentes presentaban una mayor presencia de HLA-A2 que los niños del grupo control, sugiriendo que la presencia de este Ag. era un factor de riesgo. Sin embargo, el seguimiento de los enfermos reveló que la presencia del HLA-A2 era un factor de buen pronóstico responsable de una mayor supervivencia de los pacientes y esta era la razón por la cual los casos prevalentes lo presentaban con tanta frecuencia. Este sesgo no se produce cuando los casos prevalentes representan a todos los casos existentes, lo que ocurre en enfermedades que no tienen un desenlace mortal. Por ejemplo, estudio de casos-controles sobre niveles de mercurio en orina en niños con trastornos del espectro autista.

Selección de controles

Probablemente este es el punto más difícil del diseño. Los controles deben provenir de la misma población de la que proceden los casos. Deben ser comparables a los casos, en el sentido de haber tenido la misma probabilidad de haber estado expuestos al factor de riesgo.

El grupo control estará compuesto por “sanos” (o al menos que no sufran la enfermedad estudiada), dentro del cual hay expuestos y no expuestos a la posible causa.

Una regla útil para decidir si un control es adecuado es que, si hubiese desarrollado el efecto, este debería haber aparecido en el grupo de los casos. Por ejemplo, un niño

que utiliza el sistema sanitario público no sería, en principio, un buen control para un estudio realizado con casos diagnosticados en una clínica privada.

Se puede escoger más de un sujeto control por cada caso, lo que redundaría en beneficios estadísticos. Sin embargo, la complejidad del estudio, su duración y su coste se ven incrementados por esta decisión.

Indicaciones y ventajas de los estudios de casos-controles

- Cuando la enfermedad es rara. Si la incidencia de enfermedad es de 1/1000, llevar a cabo un estudio de cohortes implicaría hacer un seguimiento de 10 000 niños durante un año para poder observar diez casos de la enfermedad, seguimiento que además de costoso es muy difícil de realizar.
- Cuando el intervalo entre la exposición y la aparición de la enfermedad es prolongado.
- Posibilidad de estudiar simultáneamente el papel de varios factores de riesgo. Cuando no se tiene una hipótesis concreta, es preferible hacer un estudio de casos-controles para buscar asociaciones significativas que posteriormente pueden ser verificadas por un estudio de cohortes.
- En líneas generales son más baratos, se pueden realizar en menos tiempo y son más fáciles de ejecutar que los estudios de cohortes.

Limitaciones y posibles sesgos

- No poder estimar directamente la incidencia (casos nuevos) de la enfermedad entre los expuestos y los no expuestos, ya que para ello es necesario realizar el seguimiento de un grupo de personas que previamente no presenten el efecto que queremos medir (estudios de cohortes). Obviamente, solo podemos estudiar el papel que los factores de riesgo tienen en la enfermedad. Es un error frecuente en estudios de casos-controles confundir la frecuencia de exposición con la incidencia de enfermedad.
- Sesgos frecuentes de selección y de información:
 1. Sesgo de información: puesto que se trata de averiguar datos que ocurrieron en el pasado, el aspecto relacionado con la memoria o recuerdos del entrevistado adquiere una importancia capital. Para evitarlos, se recomienda elaborar un cuestionario lo más detallado posible, en el que se incluyan todos los puntos que puedan ayudar

a mejorar la memoria y aplicarlo de forma exacta tanto en los casos como en los controles.

- Sesgo del observador: que el investigador conozca la hipótesis de estudio puede llevar a que se realice un interrogatorio más detallado en los casos introduciendo un sesgo que lleva a detectar falsamente más exposición en estos que en los controles.
 - Sesgo del observado: los casos tienen mayor tendencia que los controles a dar detalles acerca de sus posibles fuentes de exposición. Si esto ocurre, se introduce un sesgo en el sentido de producir un aparente exceso de exposición en los casos
2. Sesgo de selección: si la exposición de interés determina la selección de los casos y los controles (Tabla 3).

Tipos de estudios de casos-controles

Estudios de casos-controles emparejados

Los controles se escogen de forma tal que a cada caso corresponde un control igual en ciertas características previamente determinadas. Este proceso se denomina apareamiento o elaboración de pares y permite reducir el tamaño de la muestra, ya que aumenta la precisión del estudio.

Las características utilizadas con más frecuencia para hacer el apareamiento son edad y sexo, y también variables tales como el nivel socioeconómico y lugar de residencia. Al restringir los controles, de forma que tengan una característica dada igual a los casos, evidentemente se anula la posibilidad de estudiar si esa característica es diferente en los casos y en los controles. Por ejemplo, si se decide tomar un control del mismo nivel socioeconómico que los casos, es imposible saber si el nivel socioeconómico

Tabla 3. Ventajas y limitaciones de los estudios de casos-controles

Ventajas	Limitaciones
<ul style="list-style-type: none"> • Relativamente menos costosos que los estudios de seguimiento • Corta duración • Aplicaciones para el estudio de enfermedades raras • Permiten el análisis de varios factores de riesgo para una determinada enfermedad 	<ul style="list-style-type: none"> • No estiman directamente la incidencia • Facilidad de introducir sesgos de selección y/o información • La secuencia temporal entre exposición y enfermedad no siempre es fácil de establecer

mico afecta a la enfermedad estudiada. Por tanto, se debe evitar abusar del apareamiento, incluyendo más variables de las necesarias.

Estudios de casos-controles anidados

También llamados estudios caso-cohorte. Se realiza dentro de un estudio de cohortes (prospectivo o retrospectivo) y es una estrategia para reducir costes (Figura 1).

Estudios de casos-controles alternantes (case-crossover)

También llamados caso-caso. Cada caso sirve de su propio control. Este tipo de diseño es una manera científica de cuestionar y responder a la pregunta “¿estuvo haciendo el sujeto algo inusual justo antes del inicio de la enfermedad o evento de estudio?”. Se compara la exposición inmediatamente antes del desenlace, con exposiciones en otros momentos anteriores. Son útiles en eventos rápidos e imprevisibles. Por ejemplo, influencia de la falta de sueño y estrés previos al desarrollo de una crisis epiléptica.

Estimación del riesgo en los estudios de casos-controles

La medida más utilizada para cuantificar la asociación entre la exposición y la presencia de enfermedad es la *odds ratio* (OR). Cuantifica la magnitud del riesgo. La OR es una división entre dos *odds*:

$$\frac{\text{Odds de tener la enfermedad según se esté expuesto o no expuesto}}{\text{Odds de no tener la enfermedad según se esté expuesto o no expuesto}} = \text{OR}$$

El concepto de *odds* corresponde a la razón (cociente cuyo numerador no está incluido en el denominador) entre la probabilidad de que un evento ocurra y la probabilidad de que no ocurra. Para su cálculo tenemos que recurrir a la tabla 2 x 2 (Tabla 4).

Tabla 4. Cálculo de la *odds ratio*

	Casos	Controles
Expuestos	a	b
No expuestos	c	d
Odds ratio (razón de productos cruzados, oportunidad relativa)		
$\text{Odds ratio} = \frac{a \times d}{b \times c}$		

Odds de tener la enfermedad (*odds* 1) según se esté o no expuesto al factor corresponde a:

$$\text{Odds 1} = \frac{a/(a+c) \text{ enfermos expuestos en relación con el total de enfermos}}{c/(a+c) \text{ enfermos no expuestos en relación con el total de enfermos}}$$

Odds de no tener la enfermedad (*odds* 2) según se esté o no expuesto al factor corresponde a:

$$\text{Odds 2} = \frac{b/(b+d) \text{ sanos expuestos en relación con el total de enfermos}}{d/(b+d) \text{ sanos no expuestos en relación con el total de enfermos}}$$

$$\text{OR 2} = \frac{\frac{a/(a+c)}{c/(a+c)}}{\frac{b/(b+d)}{d/(b+d)}} \quad \text{OR} = \frac{a \cdot d}{b \cdot c}$$

Si el valor de la OR es 1 significa que no hay asociación entre las variables analizadas (la exposición positiva o negativa no hace diferencia respecto al riesgo de enfermar).

Un valor mayor de 1 indica una mayor frecuencia de exposición entre los enfermos (casos). Por tanto, el factor actúa asociándose con un mayor riesgo de enfermar.

Finalmente, si el valor del OR es inferior a 1, esto indica mayor frecuencia de exposición entre los sujetos sanos (controles). En este caso, la presencia del factor reduce el riesgo de enfermar (actuando como un factor de protección).

LECTURAS RECOMENDADAS

- Perea-Milla López E. Estudios de casos y controles. En: Burgos Rodríguez R (ed.). Metodología de Investigación y estructura científica en clínica. Serie de Monografías de la Escuela Andaluza de Salud Pública. 1998. p. 95-124.
- Perea-Milla López E. Estudios de seguimiento. En: Burgos Rodríguez R (ed.). Metodología de Investigación y estructura científica en clínica. Serie de Monografías de la Escuela Andaluza de Salud Pública. 1998. p. 77-93.
- Pértegas Díaz S, Pita Fernández S. Estimación del riesgo en los estudios de casos y controles [en línea]. Disponible en: http://www.fisterra.com/mbe/investiga/muestra_casos/casos_controles.asp#estudios

Son repastos en profundidad de las limitaciones, ventajas, tipos, y análisis de los datos obtenidos.

BIBLIOGRAFÍA RECOMENDADA

- Aedo S, Pavlov S, Clavero F. Riesgo relativo y odds ratio ¿Qué son y cómo se interpretan? Rev Obstet Ginecol. 2010;5:51-54.
- Bradstreet J, Geier DA, Kartzinel JJ, Adams JB, Geier MR. Estudio caso-control sobre la carga de mercurio en niños con desórdenes

del espectro autista [en línea]. Disponible en: <http://www.autismoava.org/metales-pesados/articulos/estudio-caso-control-sobre-la-carga-de-mercurio-en-ninos-con-desordenes-de>

- De Irala J, Martínez-González MA, Seguí-Gómez M, Jiménez Monleón JJ. Diseños de investigación en epidemiología. En: De Irala J, Martínez-González MA, Seguí-Gómez M (eds.). *Epidemiología aplicada*. Barcelona: Ariel; 2008. p. 179-272.
- Goldberg M. Epidemiología analítica o etiológica, o de finalidad explicativa. En: Goldberg M. (ed.). *La epidemiología sin esfuerzo*. Madrid: Ediciones Díaz de Santos, S.A.; 1994; p. 107-48.
- Guerrero R, González CL, Medina E. Estudios de casos y controles o retrospectivos. En: Guerrero R, González CL, Medina E (eds.). *Epidemiología*. , México D.F.: Addison-Wesley Iberoamericana, S.A.; 1986. p. 101-11.
- Lazcano-Ponce E, Salazar-Martínez E, Hernández-Avila M. Case-control epidemiological studies: theoretical bases, variants and applications. *Salud Pública Mex*. 2001;43:135-50.
- Merino T. Estudios de casos y controles. En: Rada G. *Epidemiología Analítica*. Pontificia Universidad Católica de Chile, 2007 [en línea]. Disponible en: <http://escuela.med.puc.cl/recursos/receptem/epianal7.htm>
- Perea-Milla López E. Estudios de casos y controles. En: Burgos Rodríguez R (ed.). *Metodología de investigación y escritura científica en clínica*. Granada: Escuela Andaluza de Salud pública; 1998. p. 95-128.
- Pértegas Díaz S, Pita Fernández S. Estimación del riesgo en los estudios de casos y controles [en línea]. Disponible en: http://www.fisterra.com/mbe/investiga/muestra_casos/casos_controles.asp#estudios
- Pita Fernández S. Tipos de estudios clínico-epidemiológicos [en línea]. Disponible en: http://www.fisterra.com/mbe/investiga/6tipos_estudios/6tipos_estudios.asp