

Tomado del Libro: Métodos y Técnicas para la Investigación en Atención Primaria de Salud. Parte II. Tema 7.

Autores: Dr. Héctor D. Bayarre Veá. Profesor ENSAP.

Dra. Maritza Oliva Pérez. Prof. Facultad "Enrique Cabrera"

Tema 7 - Estudios observacionales analíticos.

Objetivos

Que al finalizar el tema seas capaz de:

- Diseñar estudios analíticos de prevalencia, cohortes y caso-control; para su aplicación en el marco de la comunidad que atiendes.
- Interpretar medidas de asociación específicas de cada diseño estudiado, a partir del cálculo de estos indicadores; para su aplicación en la investigación de la Atención Primaria de Salud.
- Interpretar intervalos confianza y pruebas de hipótesis pertinentes para los diseños estudiados, a partir de su selección y cálculo, que permita su utilización en investigaciones de la Atención Primaria de Salud.

Introducción

Una vez que conoces el comportamiento de un evento de salud, digamos cierta enfermedad, en la comunidad que atiendes; seguramente te interesará conocer cuál es la relación que existe entre la frecuencia de ésta y diversos factores potencialmente causantes de la misma. Desde luego que tal inquietud no puede ser resuelta con un estudio descriptivo.

Como verás en las líneas que siguen, esta pretensión la resuelve la Epidemiología mediante diseños experimentales y observacionales analíticos. Estos últimos, los más utilizados en el ámbito sanitario, constituyen objeto de estudio del tema que nos compete.

Aquí encontrarás la información necesaria para que te inicies en el diseño y análisis de estos estudios, a la vez que te orientará hacia dónde dirigir tu mirada en caso que desees profundizar en la temática.

7.1 Estudio observacional analítico. Objetivos.

En este epígrafe no pretendemos establecer los objetivos de un tipo concreto de diseño epidemiológico analítico. Aquí encontrarás objetivos generales propios de cualquier investigación de este nivel.

Así, los objetivos más comúnmente planteados son:

- Identificar factores de riesgo (o pronóstico) de una enfermedad o condición determinada.
- Estimar efectos crónicos en la salud. Ej. Discapacidad, supervivencia.
- Generar nuevas hipótesis etiológicas.
- Sugerir mecanismos de causalidad.
- Sugerir potencial para prevenir enfermedades o consecuencias.

7.2 Clasificación de los estudios observacionales analíticos.

La investigación observacional analítica puede clasificarse como:

1. Transversales: estudios de prevalencia.
2. Longitudinales: estudios de cohortes y estudios de caso-control¹.

Los primeros, como ya conoces, se realizan en un momento dado; en tanto que los segundos exigen de un seguimiento, que en concordancia con las relaciones que investiga (exposición a determinado(s) factor(es) y la enfermedad), pueden ser prospectivos, retrospectivos o ambispectivos.

Como puedes ver, esta clasificación es la misma, en términos genéricos, que la estudiada en el tema anterior. Ahora bien, no te desanimes, en lo adelante conocerás elementos de estos diseños y (o) técnicas para el análisis que distinguen éstas de las estudiadas en el tema de referencia.

7.3 Estudios transversales analíticos.

Los estudios de prevalencia transitan desde el nivel descriptivo al analítico. De ahí que clasifiquen en ambos. Por ello consideramos oportuno recordarte, que lo que define su clasificación son los objetivos. Así, aquellos estudios cuyos objetivos no intenten contrastar hipótesis etiológicas serán descriptivos y en caso contrario analíticos.

Si bien, en los estudios de prevalencia descriptiva, se trata de estimar la prevalencia tanto global como específica, por característica o atributo de interés; en los estudios de prevalencia analíticos los objetivos van dirigidos a buscar la posible participación de estos factores en la producción del evento, aportando evidencias a las hipótesis o interrogantes que en torno al evento existen hasta ese momento.

Aquí los grupos de comparación surgen al determinar en una población o en una muestra aleatoria, los individuos que están enfermos; en estos y en los no enfermos se estudiará la presencia de los factores de riesgo (hipotéticamente causales).

Como puedes darte cuenta, el esquema general presentado en la figura 6.3.1 es válido en este caso. Las diferencias, en virtud de los objetivos, se producen fundamentalmente, en el análisis estadístico a emplear.

Las razones expuestas nos llevan a que centremos nuestra atención en este punto del diseño metodológico expuesto en el tema 6.

7.3.1 Análisis estadístico.

Aquí puede comenzarse por la estimación o cálculo de las medidas de frecuencia descritas en los temas 5 y 6. Sin embargo, ellas no son características de este nivel.

Las medidas de asociación son las que distinguen este nivel. Para estudiarlas, nos apoyaremos en la tabla siguiente², la cual no es nueva para ti pues ya la viste en el tema 5, pero por lo frecuente de su uso te la reproducimos a continuación:

¹ En castellano casos – testigos.

² Una tabla de este tipo suele denominarse como tabla de contingencia de 2*2.

Tabla 7.3.1- Tabla de contingencia clásica de 2x2.

Factor de exposición n	Enfermedad		Total
	Presente	Ausente	
Presente	a	b	n ₁
Ausente	c	d	n ₂
Total	m ₁	m ₂	n

Del tema 5 ya conoces el significado de cada símbolo.

Para el cálculo de las medidas de asociación se parte de medidas descriptivas. Así, se pueden calcular las prevalencias de enfermedades (tanto en expuestos como en no expuestos) y las prevalencias de exposición (en enfermos y no enfermos) y las oportunidades.

Medidas de frecuencia:

A partir de ellas se pueden calcular las medidas de frecuencia relativas entre el factor y la enfermedad. Dentro de ellas las más frecuentes son:

$$\frac{P(E+)}{1-P(E+)} = \frac{a/n_1}{b/n_1} = \frac{a}{b}$$

P(E+) = Probabilidad de estar enfermo en los expuestos

1-P(E+) = Probabilidad de no enfermar en ese grupo

Esta medida ha sido denominada por algunos como “**Oportunidad de enfermedad en expuestos**”. La misma es la razón entre la probabilidad de presentar la enfermedad en los expuestos y la probabilidad de no presentarla en este mismo grupo.

También se puede calcular la **oportunidad de enfermedad en no expuestos**.

$$\frac{P(E-)}{1-P(E-)} = \frac{c/n_2}{d/n_2} = \frac{c}{d}$$

P(E-) = Probabilidad de enfermar en no expuestos

1-P(E-) = Probabilidad de no enfermar dentro del grupo.

De la misma forma, se puede realizar el cálculo de las oportunidades de la exposición tanto en enfermos como en no enfermos.

Oportunidad de exposición en enfermos:

$$\frac{P(+E)}{1-P(+E)} = \frac{a/m_1}{c/m_1} = \frac{a}{c}$$

P(+E) = Probabilidad de exposición en los enfermos

1-P(+E) = Probabilidad de no exposición en ese grupo

Oportunidad de exposición en los no enfermos:

$$\frac{P(+\bar{E})}{1-P(+\bar{E})} = \frac{b/m_2}{d/m_2} = \frac{b}{d}$$

$P(+\bar{E})$ = Probabilidad de exposición en los no enfermos

$1-P(+\bar{E})$ = Probabilidad de no exposición en ese grupo

Las oportunidades calculadas y las prevalencias, son medidas relacionadas y equivalentes. Así, a mayor prevalencia mayor oportunidad, y a cada valor de prevalencia le corresponde un único valor de oportunidad. Note que, dado que la prevalencia es una probabilidad, su valor estará entre cero y uno, en tanto que la oportunidad puede tomar valores entre cero e infinito.

La precisión de estas medidas vendrá dada por la construcción de intervalos de confianza.

El **intervalo de confianza al 95%** para la **P** (prevalencia estimada), si se trata de un muestreo simple aleatorio vendrá dada por:

$$P \pm 1.96 \sqrt{\frac{PQ}{n}} \quad Q = 1-P$$

Ello será válido, como conoces, solo si: $n \cdot P$ y $n \cdot Q \geq 5$.

Medidas de asociación.

Mensurar la asociación entre la exposición a cierto(s) factor(es) y la enfermedad es el elemento distintivo del análisis en este tipo de investigación. De hecho, se han descrito varias medidas, de ellas la **prevalencia relativa** y la **oportunidad relativa** son las más comúnmente utilizadas.

Se pueden calcular dos prevalencias relativas, ello depende de la lectura que se le de a la tabla 7.3.1; si se analiza de forma horizontal se calcula la **prevalencia relativa de enfermedad (PR)**. Aquí se divide la prevalencia de enfermedad en expuestos por la prevalencia de enfermedad en no expuestos:

$$PR = \frac{P(E+)}{P(E-)} = \frac{a/n_1}{c/n_2}$$

Indica cuantas veces es más frecuente la enfermedad en individuos con el factor en relación con los que no lo tienen, dicho de otro modo, cuantas veces es mayor el riesgo de enfermar en los sujetos expuestos a cierto factor, respecto a los no expuestos a este.

El segundo tipo de prevalencia relativa, surge de la lectura vertical de la tabla 7.3.1, la **prevalencia relativa de exposición (PR_(x))**, la cual se obtiene de dividir las prevalencias de exposición en enfermos y no enfermos.

$$PR_{(x)} = \frac{P(+E)}{P(+\bar{E})} = \frac{a/m_1}{b/m_2}$$

Indica cuantas veces es más frecuente la exposición en enfermos con relación a los no enfermos; dicho de otro modo, cuantas veces es mayor el riesgo de exposición en los enfermos respecto a los no enfermos.

La interpretación de la PR es similar a la del riesgo relativo (RR). En el caso que el factor de riesgo y la enfermedad no estén asociados la PR sería igual a 1, es decir, la prevalencia es igual en expuestos que en no expuestos. Si están asociados la prevalencia es mayor en expuestos que en no expuestos y en tanto mayor sea la PR más fuerte será la asociación.

La verificación de la existencia de la asociación significativa (no debida al azar) suele realizarse mediante dos procedimientos. La prueba χ^2 y la construcción de intervalos de confianza.

Pruebas de independencia (χ^2).

Usos: cuando tenemos una muestra aleatoria y las **n** unidades de análisis clasifican respecto a dos variables (exposición a cierto factor y presencia de una enfermedad).

Hipótesis:

H₀: Existe, a nivel poblacional, asociación entre la exposición y la enfermedad.

H₁: No existe asociación, a nivel poblacional, entre la exposición y la enfermedad.

Estadígrafo de prueba:

$$\chi^2 = \frac{n (|ad-bc| - \frac{n}{2})^2}{n_1 n_2 m_1 m_2} \sim \chi^2_{1 \text{ gl}}$$

Regla de decisión:

Rechazar la hipótesis nula si χ^2 observado $>$ χ^2 teórico [1 gl; (1- α)]

La zona de aceptación y rechazo para $\alpha = 0.05$, se representa en la siguiente figura. Fíjate que el valor de χ^2 teórico en este caso es igual a 3.84, lo cual puedes comprobarlo en la tabla C del Apéndice.

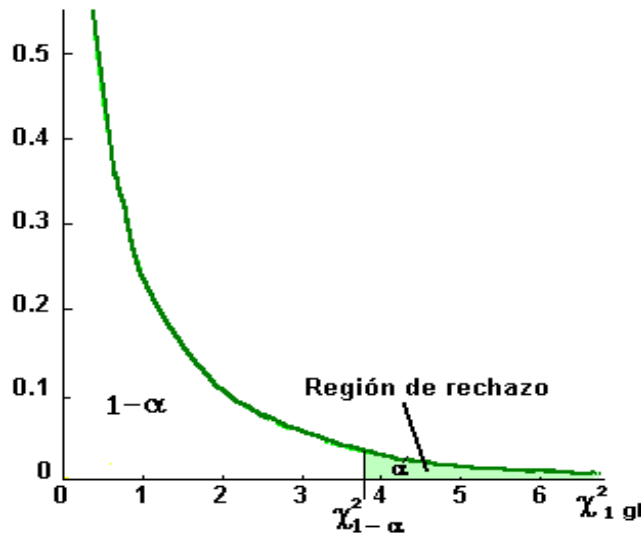


Figura 7.3.1- Distribución χ^2 con 1 gl. Zona de rechazo de H₀ para $\alpha = 0.05$

Intervalo de confianza al 95%.

$$PR \pm 1.96 \sqrt{\text{Var}(PR)}$$

Dos problemas se presentan al construir el intervalo de confianza de la PR por la forma descrita: el primero es que este presupone que el parámetro estimado sigue una

distribución normal y el segundo es que precisa conocer el valor de la varianza de la PR. Esto se resuelve mediante la aplicación del logaritmo neperiano (ln):

$$\text{Ln PR} \pm 1.96 \sqrt{\text{Var}(\text{Ln PR})} \quad \text{Donde:} \quad \text{Var}(\text{Ln PR}) = \frac{b/a}{n_1} + \frac{d/a}{n_2}$$

Luego de obtenido los dos valores, debes calcular el antilogaritmo (exponencial), y así quedarán computados los límites del intervalo de confianza.

Para la interpretación, si el intervalo no contiene al 1, entonces la asociación es estadísticamente significativa.

Oportunidad relativa (odds ratio).

Si bien, teóricamente se pueden calcular dos tipos de odds ratio (de enfermedad y de exposición), en la práctica estas se calculan igual:

$$\text{OR} = \frac{ad}{cb}$$

Indica cuantas veces la exposición incrementa la oportunidad de presentar la enfermedad.

La estimación puntual debe complementarse con su intervalo de confianza. Aquí presentamos el propuesto por Woolf.

$$\text{OR}_I = \text{OR} e^{-1.96\sqrt{\text{Var}(\text{Ln OR})}} \quad \text{OR}_S = \text{OR} e^{+1.96\sqrt{\text{Var}(\text{Ln OR})}}$$

Donde OR_I y OR_S son los límites inferior y superior del intervalo de confianza y $\text{Var}(\text{Ln OR})$ es la varianza del logaritmo neperiano del OR, estimada por:

$$\text{Var}(\text{Ln OR}) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

El escoger la prevalencia relativa o el odds ratio es intrascendente en este tipo de diseño.

Ahora bien, hasta ahora hemos descrito métodos univariados. Ellos tienen la desventaja que en la mayoría de las situaciones hay factores (confusores) que no fueron controlados en el diseño y que deberían serlo en el análisis, para ello suelen emplearse (además del ajuste de tasas) otros métodos como el de Mantel Haenszel y la regresión logística, pero ellos serán explicados en un momento posterior.

Ilustremos lo explicado con el ejemplo siguiente:

Ejemplo 7.3.1

En una comunidad se ha detectado una tendencia ascendente del bajo peso al nacer (BPN). Por tal motivo se decide realizar un estudio que permita identificar cuáles de un conjunto de presuntos factores de riesgo del BPN, están asociados al evento referido. Se escoge una muestra simple aleatoria de recién nacidos y se evalúan los factores seleccionados y la condición de BPN. La ganancia de peso durante el embarazo (GPDE) fue uno de esos factores. Los resultados relacionados con éste se exponen en la tabla siguiente:

Tabla 7.3.2 – Recién nacidos según ganancia de peso de la madre durante el embarazo y presencia de BPN

GPDE (Kg)	BPN		Total
	Si	No	
< 8	50	10	60
≥ 8	4	136	140
Total	54	146	250

- a) calcule las prevalencias relativas y el odds ratio.
 b) ¿Puede afirmarse que existe asociación significativa entre la ganancia de peso durante el embarazo y el bajo peso al nacer?

Solución.

$$a) \text{ PR} = \frac{50/60}{4/140} = \frac{0.83}{0.03} = \underline{\underline{27.7}} \quad \text{PR}_{(x)} = \frac{50/54}{10/146} = \frac{0.92}{0.07} = \underline{\underline{13.14}} \quad \text{OR} =$$

$$\frac{50 * 136}{4 * 10} = \frac{6800}{40} = \underline{\underline{170}}$$

b) Con los resultados anteriores podemos decir que “parece” existir asociación entre la GPDE y el BPN, pues, analizando la PR, es casi 28 veces más frecuente el BPN en los recién nacidos de madres con ganancia de peso inferior a los 8 Kg durante el embarazo, respecto a los de las que ganaron 8 Kg o más. Puedes realizar el análisis de las otras medidas con resultados similares. Ahora bien, ¿Por qué “parece”? Bueno, este es el resultado del estudio de una muestra aleatoria, por tanto no puedo saber si se debe al azar o no hasta tanto realice una prueba de significación estadística o construya un intervalo de confianza para los parámetros estimados.

Construyamos el IC al 95% para cada una de las medidas anteriores, no te alarmes, ya sé que son fórmulas muy complejas para ti y no pretendemos que hagas los cálculos, nos auxiliaremos del programa EPIINFO, así tenemos:

IC al 95% para el OR **(45.92 a 701.49)**, como era de esperar, con un OR tan elevado, no contiene el 1.

No obstante, recuerda que no debemos ser categóricos, pues siempre hay un margen de error que no puede ser despreciado, por lo tanto podemos concluir que existe asociación estadísticamente significativa entre la GPDE y el BPN, lo cual no quiere decir que sea de importancia clínica o epidemiológica.

7.3.2 - Ventajas y desventajas.

Entre las ventajas de los estudios transversales tenemos:

- Son útiles para el estudio de enfermedades de larga duración y frecuentes en la comunidad (crónicas no transmisibles).
- Sus resultados son fácilmente generalizables al total poblacional ya que en su mayoría se basan en la selección de una muestra representativa de la población.
- Es menos costoso y requiere menos tiempo que otros diseños analíticos.
- Es muy útil en la planificación sanitaria y en la descripción de las características de la población.

Como desventaja fundamental podemos citar la siguiente:

- Su diseño no permite distinguir entre un factor de riesgo de una enfermedad y un factor pronóstico de la misma, lo que afecta el proceso inferencial.

7.4 Estudios de cohortes.

Un estudio de cohortes puede ser definido como sigue a continuación:



Es un método de estudio epidemiológico en el que se pueden identificar grupos de una población determinada, que están, han estado o pueden estar en el futuro, expuestos o no, o expuestos en diferente grado, a un factor que supuestamente influye sobre la probabilidad de aparición de una enfermedad dada u otro evento de salud; los que son seguidos durante un intervalo de tiempo para estimar las tasas de incidencia de la enfermedad en los diferentes grupos y establecer comparaciones.

Se trata básicamente de un estudio longitudinal prospectivo (hacia adelante, de la causa al efecto), en el que 2 grupos o más de sujetos (cohortes) sanos o libres de la enfermedad o del efecto objeto de estudio, uno expuesto al supuesto factor de riesgo y el otro no expuesto o en menor grado, se van a observar y seguir en el tiempo con el objetivo de valorar y cuantificar la aparición de la enfermedad, para establecer comparaciones entre ambos y ver así la relación existente entre la enfermedad y el factor.

El término cohorte, se refiere a que los grupos de estudio se comportan como tal al compartir la característica común de estar expuestos o no al factor.
Las características principales de este diseño son:

- Los individuos que serán estudiados se encuentran libres de la enfermedad o del efecto que se desea estudiar en el momento de su inclusión en el estudio.
- Siempre que sea posible se debe cuantificar el grado de exposición al factor.
- Siempre se debe poder cuantificar de manera precisa y estandarizada los casos de la enfermedad que aparezcan en ambos grupos durante el seguimiento.

El tiempo de seguimiento será variable en dependencia de los objetivos, los recursos disponibles y las características del proceso estudiado.
Los principales puntos metodológicos a seguir en el diseño de un estudio de cohortes son:

1. Planteamiento de una hipótesis en forma precisa y operacional: Ésta se refiere básicamente a la suposición inicial de que el factor se encuentra asociado a la enfermedad.
2. Definición de la variable independiente (factor de riesgo) y dependiente (enfermedad): durante el estudio se debe obtener información tanto del factor como de la enfermedad. El factor a estudiar debe ser definido claramente, estableciéndose de forma precisa cómo se medirá la exposición y en caso de ser posible cuantificar el grado de la misma. Las características de la exposición deben ser evaluadas en términos de su intensidad, duración, regularidad y variabilidad, puede que con el transcurso del tiempo se requiera de una frecuente reevaluación de las exposiciones, si el estatus de las mismas cambia. La medición objetiva o los marcadores biológicos de la exposición son preferibles a las mediciones subjetivas.

La enfermedad o el efecto de interés debe ser definido también de manera clara y precisa, pues es muy importante que ambos grupos estén libres de la misma al inicio del estudio. Esto es difícil debido a la presencia de formas subclínicas e inaparentes de enfermedad, una forma de evitar esto es excluir todos los casos de la enfermedad que surjan precozmente en el curso del seguimiento, con la suposición que la enfermedad comenzó antes del inicio del estudio.

3. Identificación de las fuentes y criterios de elección de las cohortes que serán estudiadas: Este punto es de vital importancia en este tipo de estudio, abordaremos por separado las fuentes de selección de cada grupo.

Selección de la cohorte expuesta: Esta depende de una serie de factores de orden científico y de factibilidad. Si se trata de exposiciones frecuentes es posible seleccionar a los sujetos de la población general, por ejemplo el hábito de fumar, el consumo de café o grasas en la dieta, etc.; en cambio exposiciones poco frecuentes hacen necesario seleccionar grupos específicos particularmente expuestos, por ejemplo profesionales de Rayos x, mineros, etc. También puede ser grupos definidos geográficamente. No debes olvidar la factibilidad, la población debe ser accesible y estar motivada a participar.

Selección de la cohorte no expuesta: Depende también de elementos de orden científico y de factibilidad, pero, depende en gran medida de las características y la fuente de obtención de los expuestos. El principio fundamental de la selección de este grupo, es que debe garantizar una favorable comparación con la cohorte expuesta en otros aspectos como edad, sexo, etc. En caso de exposiciones frecuentes puede ser tomada una muestra de la población general y ser dividida de acuerdo a la exposición. En ocasiones se compara la frecuencia de enfermar del grupo expuesto con la frecuencia de enfermar en la población general, esto puede hacerse cuando la probabilidad de exposición en la población general es pequeña.

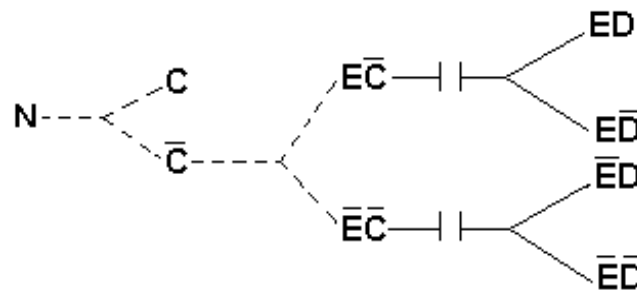
De forma general, se deben tener en cuenta los siguientes aspectos para la selección de las cohortes de estudio:

- Las personas no expuestas deben ser seleccionados de la misma población (o comparable) que la cohorte de expuestos.
- Ambos grupos deben estar libres de la enfermedad o del efecto de interés al iniciarse el estudio.

- Las características iniciales de ambos grupos no deben diferenciarse sistemáticamente a excepción de la exposición de interés.
- Se debe disponer de la misma información en cantidad y calidad sobre la exposición y la enfermedad en ambos grupos.
- Ambos grupos deben estar igualmente disponibles para el seguimiento.

En cualquier caso las posibilidades dependen de los objetivos, costos, accesibilidad, disponibilidad de informes, etc.

El diseño general de los estudios de cohortes se esquematiza en la siguiente figura:



Leyenda:

N = Población

C = Caso prevalente y \bar{C} su complemento, es decir, no caso.

E = Individuo con el factor y \bar{E} sin el factor.

D = Individuo enfermo y \bar{D} no enfermo.

|| = Seguimiento

Figura 7.4.1- Esquema del diseño de un estudio de cohortes.

4 - Seguimiento.

El objetivo fundamental de esta fase es la detección del efecto pero también permite valorar las fluctuaciones en la exposición y las pérdidas de sujetos participantes.

En estos estudios resulta muy importante que la vigilancia de la enfermedad sea de la misma forma en ambos grupos, la frecuencia de los exámenes, lo detallado de éstos y la duración del seguimiento debe ser similar en ambos y dependen del tipo de exposición y de las características del efecto o enfermedad bajo vigilancia.

Se debe seguir un procedimiento diagnóstico estandarizado que no tenga en cuenta el estatus de la exposición. La exactitud y la confiabilidad del diagnóstico no deben ser diferentes entre los dos grupos para asegurar una correcta comparación entre ellos.

7.4.1 – Análisis estadístico.

El análisis estadístico en este tipo de estudios se caracteriza por el uso de medidas de asociación aunque, por supuesto, se emplean las medidas de frecuencia.

El primer paso es la medición de la frecuencia de la enfermedad en los grupos, la cual se realiza mediante las tasas de incidencia de la enfermedad. La información se resume en la clásica tabla de contingencia 2x2 (tabla 7.3.1). a partir de ella podemos calcular la incidencia acumulada tanto en expuestos como en no expuestos:

$$IA_{\text{exp}} = \frac{a}{n_1} \qquad IA_{\text{no exp}} = \frac{c}{n_2}$$

Existen fórmulas que tienen en cuenta las pérdidas, pero no las veremos para evitar complejizar los cálculos, de todas formas si se realiza una buena selección de las cohortes el porcentaje de pérdidas es bajo y no es necesario tenerlo en cuenta.

También puede calcularse la densidad de incidencia, pero no es necesario calcular ambas medidas, se usará una o la otra en dependencia de los objetivos y de la duración del período de incubación o de riesgo de la enfermedad.

Una vez calculadas las incidencias en ambos grupos se procede a medir el grado de asociación existente mediante el cálculo del riesgo relativo, RR, o la Razón de Densidad de Incidencia en caso de que sean calculadas las DI; para posteriormente comprobar la significación de la asociación.

$$RR = \frac{IA_{\text{exp}}}{IA_{\text{no exp}}} \quad \text{o} \quad RDI = \frac{DI_{\text{exp}}}{DI_{\text{no exp}}}$$

Como recordarás, este indicador señala la probabilidad de contraer una enfermedad en el grupo expuesto con relación a los no expuestos. Su interpretación ya la vimos en el tema 5.

La significación de la asociación se lleva a cabo mediante una prueba de hipótesis, de forma similar a la ya vista en los estudios transversales, el estadígrafo de prueba es el mismo χ^2 que conociste. Además del cálculo de un intervalo de confianza para el RR, al igual que vimos para la PR, pero la fórmula para el cálculo de los límites de dicho intervalo son complejas y no las abordaremos, de todas formas en la práctica se emplea un programa estadístico para hacerlo, por ejemplo el EPIINFO.

También puede ser calculada la proporción de casos que puede ser explicada por la exposición, a través del cálculo del riesgo atribuible, RA. Éste ofrece información sobre el efecto absoluto de la exposición y se define como la diferencia entre las tasas de incidencia de los expuestos y no expuestos:

$$RA = IA_{\text{exp}} - IA_{\text{no exp}} \quad \text{Puede ser expresado por 1000 o 100000 habitantes.}$$

El riesgo relativo y el riesgo atribuible ofrecen diferentes tipos de información. El primero es una medida de la fuerza de la asociación entre la exposición y la enfermedad, ofrece información que puede ser utilizada para evaluar la probabilidad de una relación causal, mientras que el segundo, es una medida del impacto de una exposición para la salud pública, suponiendo que la asociación sea de causa y efecto.

Desde el punto de vista de su importancia para la salud pública y para la implementación de políticas de prevención, el riesgo atribuible suele ser empleado ya que expresa el número de casos que pudieran evitarse en la población con la eliminación de la exposición.

7.4.2 El sesgo en los estudios de cohortes.

Los sesgos que con más frecuencia afectan este tipo de estudio son los sesgos de clasificación o de información, las pérdidas durante el seguimiento, sesgos de confusión y en menor medida los de selección.

Sesgos de clasificación: la clasificación errónea de los sujetos respecto a la exposición y al estado de la enfermedad, suele ser una de las principales fuentes de error de estos estudios. La magnitud de este error está determinada por la sensibilidad (probabilidad de que un individuo expuesto sea clasificado como tal) y especificidad (probabilidad de que un individuo no expuesto sea clasificado como no expuesto). Estos conceptos los verás con más detalle en próximos cursos.

Ya vimos en el tema 5 que la clasificación errónea puede ser no diferencial (cuando ocurre en igual magnitud en los dos grupos) y diferencial (cuando la exactitud de la información es diferente entre los grupos de estudio) la cual puede dar lugar a una subestimación o sobreestimación de la asociación.

Pérdidas durante el seguimiento: esta es una de las fuentes más frecuentes de error junto con los errores de clasificación. Varios individuos en los grupos de expuestos y no expuestos pueden perderse durante el seguimiento, si el porcentaje de pérdidas es alto (30 o 40%) esto plantea interrogantes sobre la validez del estudio, sin embargo, aún en los casos que las pérdidas sean pocas, si la probabilidad de abandonar el estudio está relacionada con la exposición o el estado de la enfermedad afectará la validez de los hallazgos.

Sesgos de selección: estos se pueden presentar en diferentes situaciones, por ejemplo, cuando el grupo estudiado no refleja la misma composición por edad y sexo que la población de referencia, el grado en que la población accede a participar no representa el verdadero espectro de la exposición en dicha población, también cuando se incluyen sujetos con enfermedad latente o cuando la distribución de otras variables ajenas a la exposición que pueden estar relacionadas con la incidencia, no están igualmente representada en las cohortes, etc.

Sesgos de confusión: es producido por la existencia de factores de confusión, es decir, una variable que actúa como factor de riesgo para la enfermedad, está asociada al factor de riesgo en estudio y no es una variable intermedia en la cadena causal de la enfermedad sometida a estudio; que no muestran la misma distribución en las cohortes.

Esto puede corregirse empleando un método adecuado de selección de los grupos, por ejemplo, pareando por el factor de confusión o seleccionando los sujetos que se encuentren en una categoría específica del factor de confusión (restricción), o en el análisis siempre y cuando se registró la información referente a dicho factor confusor (estratificación y análisis multivariado).

La probabilidad de cometer errores se incrementa cuando se establecen conclusiones sobre la base de hallazgos o asociaciones que no formaron parte de los objetivos de la investigación.

7.4.3 Ventajas y desventajas.

Veamos primero las ventajas.

- Proporciona una descripción completa de las experiencias posteriores a la exposición, con lo cual se demuestra más claramente la secuencia temporal entre la exposición y la aparición de la enfermedad.
- Por su diseño posibilita el estudio de múltiples efectos potenciales de una única exposición.

- Se obtienen estimaciones directas de la incidencia de la enfermedad y del cálculo de riesgos.
- Ofrece menos posibilidad de cometer sesgos de selección en relación con otros diseños analíticos.
- Resultan muy convenientes para evaluar el efecto de exposiciones poco frecuentes.
- Permite valorar la relación entre el factor de riesgo y la mortalidad.

Entre sus desventajas tenemos las siguientes:

- Son costosos.
- Requieren mayor consumo de tiempo que otros diseños.
- Los criterios para el diagnóstico de la enfermedad pueden cambiar a lo largo del período de seguimiento lo que puede dar al traste con la investigación.
- No se pueden aplicar para enfermedades poco frecuentes.
- Por su larga duración está expuesto a que el avance de otras investigaciones paralelas hagan irrelevantes los resultados.
- Puede que el propio estudio influya en cambios de hábitos y costumbres de los sujetos estudiados y por tanto modificarse la exposición.
- Pueden producirse pérdidas durante el seguimiento, tanto de sujetos como de investigadores.

7.5 Estudios de caso-control.

Los estudios de caso-control son muy usados debido a su menor coste en relación con otros diseños, pueden ser tanto descriptivos como analíticos, siendo la mayoría de ellos analíticos. Pueden definirse como sigue a continuación:



Estudios que comienzan con la identificación de personas con la enfermedad u otro evento de interés y un grupo adecuado de personas de control sin la enfermedad o característica de interés, donde se compara la frecuencia de presentación de uno o más factores de riesgo en ambos grupos.

Este tipo de estudio tiene como objetivo principal establecer la asociación entre posibles factores de riesgo y la ocurrencia de un evento en particular, son frecuentemente empleados para detectar las fuentes de infección u otros factores causales durante las epidemias, para la evaluación de vacunas y para la identificación de factores que hacen algunos grupos poblacionales más susceptibles a sufrir un evento de salud.

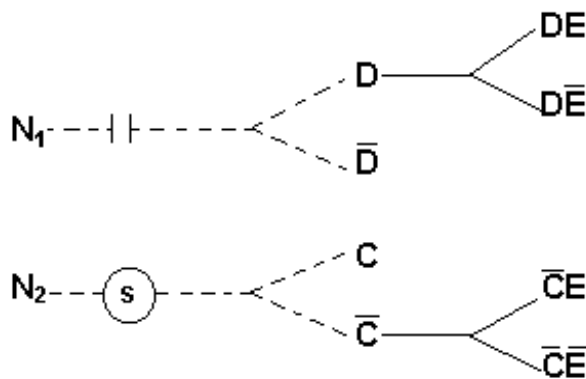
Se basa generalmente en la selección de dos grupos de sujetos, uno con la enfermedad o evento de interés, llamado casos y el otro sin dicho evento, los controles. Ambos grupos son comparados con respecto a la frecuencia de presentación de un determinado factor o factores con el objetivo de esclarecer su papel en la etiología de la enfermedad en estudio.

Características principales.

- Generalmente son retrospectivos, hacia atrás, se parte del efecto y se busca en el pasado la posible causa.

- Los casos son personas que tienen el problema de salud bien definido al comienzo del estudio.
- Los controles son personas que no han desarrollado el problema de salud al comienzo del estudio y son seleccionados para establecer la comparación.
- Ambos grupos son investigados para comparar la frecuencia de exposición a los factores de riesgo potenciales.
- Los casos y los controles deben estar muy bien definidos, lo cual constituye la clave de estos estudios. Es necesario especificar las características de los casos, los criterios para la selección de un control, la forma de selección y los criterios de exclusión.

El diseño básico se esquematiza en la siguiente figura.



Leyenda

N_1 = población de referencia para los controles

N_2 = población de referencia para los casos

C = Caso prevalente y \bar{C} , no caso.

E = Individuo con el factor y \bar{E} , sin el factor.

D = individuo enfermo y \bar{D} , no enfermo.

S = Muestra aleatoria

Figura 7.5.1 Esquema gráfico de un estudio de caso-control.

Como indica la figura, el investigador generalmente selecciona por separado los grupos de poblaciones de casos y controles disponibles. En algunas situaciones es posible seleccionar ambos grupos de la misma población de referencia, por ejemplo cuando la enfermedad y la exposición son comunes en la población, lo cual cambiaría completamente el esquema anterior.

La confección del diseño presupone el seguimiento de los siguientes puntos metodológicos:

1. Planteamiento de una hipótesis en forma precisa y operacional.
2. Definición de la variable dependiente: la enfermedad o acontecimiento de interés.
3. Definición de las variables independientes: factores de riesgo a estudiar.

4. **Fuente y criterio de selección de los casos.** Es muy importante definir los criterios diagnósticos necesarios para incluir un caso, es preciso contar con medios diagnósticos estandarizados y con alta sensibilidad y especificidad para evitar el sesgo de clasificación. Se deben precisar los criterios de exclusión de posibles casos y los criterios para clasificar la etapa o estadio de la enfermedad.

Una vez definidos los aspectos anteriores, es necesario determinar la fuente de obtención de los casos. Puede ser el total o una muestra de las personas atendidas por la enfermedad, en una o varias instituciones sanitarias durante un período dado; o el total o una muestra de las personas que sufren la enfermedad, en una comunidad durante un lapso de tiempo. La primera suele ser la más usada, por ser más accesible, económica y permite mayor cooperación por parte de los enfermos, sin embargo, es posible que los resultados obtenidos no puedan ser generalizados a la población general, por lo que deben ser definidas las limitaciones de la fuente empleada en cada caso. Se recomienda usar solamente los casos incidentes de la enfermedad.

5. **Fuente y criterio de selección de los controles:** la definición y selección de los controles es la cuestión más difícil. Se requiere que los controles permitan la comparación de la frecuencia de exposición observada. Estos deben ser seleccionados de forma tal, que sean similares a los casos en todos los aspectos con excepción de la enfermedad, esto no es posible por razones prácticas, sin embargo deben acercarse lo más posible a este ideal.

Fuentes más frecuentes de selección de los controles:

- Muestra de pacientes atendidos en la institución de la que fueron seleccionados los casos que están libres de la enfermedad objeto de estudio. Tiene como ventaja que son accesibles, fáciles de identificar, muestran disposición a colaborar, pero estos al estar enfermos por otras causas son diferentes a la población general, además los criterios para ingresar en las diferentes instituciones no es el mismo para todos los enfermos lo cual trae consigo un sesgo de selección.
- Muestra de familiares: los familiares de los casos pueden ser tomados como controles, estos son accesibles, colaboran pero se corre el riesgo de que los factores analizados sean compartidos por los miembros de la familia.
- Vecinos: los vecinos de los casos son una opción para seleccionar los controles, estos ofrecen la ventaja de compartir un medio similar y a la vez tener diferentes antecedentes genéticos respecto a los casos. Como desventaja se puede decir que es muy trabajoso obtener la información sobre ellos.
- Muestra de la población general: es generalmente más costoso y consume más tiempo.
- Estadísticas nacionales o regionales: consiste en comparar los casos con los datos registrados en las estadísticas de la región.

Sea cual sea la fuente empleada, es importante evaluar las limitaciones de la misma al dar las conclusiones del estudio. Por otra parte, se pueden seguir diferentes estrategias al seleccionar los controles para que sean comparables:

1. No imponer ningún criterio en la selección, pero deberán registrarse las variables confusoras para controlarlas en el análisis.
2. Adopción de criterios de restricción al formar el grupo control. Ésta puede ser total o parcial.
 - a) La restricción total consiste en lograr similitud total entre ambos grupos con excepción de la enfermedad y los factores de riesgo. Esto incrementa la validez del estudio pero lo encarece y se hace casi imposible conseguir los controles.
 - b) La restricción parcial consiste en controlar algunas de las variables confusoras conocidas, lo cual es más fácil de llevar a la práctica. Puede realizarse mediante un muestreo estratificado por esa variable confusora o mediante el apareamiento.

7.5.1 Análisis estadístico.

La base del análisis estadístico es la comparación entre los casos y los controles con relación a la frecuencia de exposición a los factores de riesgo. Para ello se necesita del cálculo tanto de medidas de frecuencia como de asociación.

Como ya sabes, en estos estudios no es posible estimar las tasas de incidencia por lo que no se puede calcular el riesgo relativo, entonces el análisis se basará en otras medidas conocidas. Veamos, primero es necesario presentar la información en la tabla clásica de 2x2 (tabla 7.3.1). Ahora podemos calcular la proporción de expuestos en ambos grupos: Proporción de expuestos en los casos (P_1) y Proporción de expuestos en los controles (P_2):

$$P_1 = \frac{a}{m_1} \qquad P_2 = \frac{b}{m_2}$$

El objetivo es comprobar si existen diferencias entre éstas, la pregunta concreta es si P_1 es mayor que P_2 . Para responder esa interrogante se calcula la razón de productos cruzados (RPC) también conocida como Odds ratio (OR):

$$OR = \frac{ad}{bc} \qquad \text{El significado ya lo vimos en el tema 5, es muy similar al}$$

RR.

Aquí también es válido lo dicho para el RR, en lo referente a la significación estadística de la asociación: puede realizarse una prueba de hipótesis (χ^2) o construir un intervalo de confianza para el OR.

7.5.2 Los sesgos y los factores de confusión

La probabilidad de sesgos es mayor en este diseño respecto a otros, debido fundamentalmente a que son retrospectivos, se estudia la exposición partiendo del efecto. A pesar de esto, suelen ser muy usados por ser en muchas ocasiones el diseño más factible.

Entonces, ante una asociación esta puede ser real o como resultado de distorsiones que se producen en el proceso de recopilación, análisis o interpretación de los resultados (sesgos) o de la superposición de uno o varios factores confusores.

Los sesgos en este tipo de estudio pueden ser incluidos en tres grupos fundamentales: de selección, de clasificación y por factores de confusión.

Sesgos de selección: se produce cuando hay una desigualdad en la inclusión de los casos y (o) los controles fundamentada en el estatus de exposición y (o) de enfermedad. Entre los más conocidos tenemos los siguientes:

- Paradoja de Berkson. Se produce cuando los casos y los controles son hospitalarios y estos difieren de forma sistemática de las poblaciones a las que se pretende generalizar los resultados, debido a que la probabilidad de hospitalización es diferente para los casos y los controles.
- Falacia de Neyman. Se produce cuando se estudian casos prevalentes de la enfermedad, siendo la exposición un factor pronóstico de ésta.
- Sesgo de detección. Se debe al hecho de que la exposición conduce a una mayor vigilancia facilitando el proceso de detección de la enfermedad.
- Sesgo de los no respondentes. Se produce por la existencia de diferencias entre los sujetos participantes y los que se negaron a participar de los designados inicialmente para colaborar en el estudio.
- Sesgo por inclusión – exclusión. Puede surgir cuando se incorporan o eliminan sistemáticamente otras enfermedades que están relacionadas con la exposición objeto de estudio.

Sesgos de clasificación: estos sesgos se producen cuando hay una valoración inexacta del status de exposición y (o) de enfermedad. Este error en la clasificación puede ser diferencial o no diferencial, como vimos en el tema 5.

- Error de clasificación no diferencial. En este caso los errores de clasificación de una variable (por ej. La exposición) no dependen del nivel de clasificación de la otra (la enfermedad), por lo tanto los errores se comportarán de manera similar en ambos grupos.
- Error de clasificación diferencial. En este caso los errores de clasificación de una variable dependen del nivel de la otra. Entre las situaciones más comunes que pueden llevar a este tipo de sesgo tenemos las siguientes:
 - a) Sesgo de memoria: cuando los datos de la exposición dependen de la memoria del sujeto y ésta está influida por el status de la enfermedad. Es frecuente que los casos estén más motivados a recordar la exposición que los controles, o por el contrario, que la memoria esté debilitada por la enfermedad.
 - b) Sesgo protopático: puede producirse cuando las manifestaciones precoces de la enfermedad condicionan cambios en la exposición de los casos.
 - c) Sesgos de la entrevista: se produce cuando las características de la entrevista difieren entre los casos y los controles. Es el resultado de diferencias sistemáticas al solicitar, registrar o interpretar la información procedente de los grupos participantes en el estudio. Este reviste particular importancia en este tipo de estudio, ya que al conocer el investigador el estado patológico del sujeto al inicio del estudio, es posible que ponga mayor interés en encontrar la exposición en el grupo de los casos.

Sesgos de confusión: este sesgo se produce por la distorsión que puede ocurrir por la existencia de un factor de confusión, de forma que la posible asociación entre la exposición y la enfermedad esté alterada en mayor o menor medida.

El papel de los factores de confusión puede ser controlado, como ya dijimos, tanto en el diseño como en el análisis.

7.5.3 Ventajas y desventajas.

Entre las múltiples ventajas que ofrece el uso de este diseño se pueden citar las siguientes:

- Son relativamente fáciles de diseñar y de llevarse a cabo, lo que los hace poco costosos.
- Son muy adecuados para el estudio de enfermedades con largos períodos de latencia.
- Permiten estudiar enfermedades raras.
- Requieren, comparativamente con otros diseños, menos sujetos.
- Suelen ser eficientes para pesquisar un amplio espectro de factores de riesgo.
- No suponen, generalmente, riesgos para los participantes.
- Son fácilmente reproducibles.

Entre las desventajas tenemos que:

- Existe una dificultad grande para conseguir un grupo adecuado de comparación.
- Son ineficientes para evaluar exposiciones poco frecuentes.
- Mayor posibilidad de sesgos, en ocasiones, difícilmente medibles.
- No ofrecen tasas de incidencia y las asociaciones encontradas no demuestran asociaciones causales.
- El control de los factores de confusión suele ser incompleto.
- La validación de la información referente a la exposición es difícil y a veces imposible.

Resumen

En este tema estudiaste que:

1. Los estudios observacionales analíticos constituyen un elemento esencial en la solución de problemas de investigación donde el objetivo principal es la búsqueda de asociaciones causales entre fenómenos o eventos.
2. Entre los objetivos generales de estos estudios se encuentra la identificación de factores de riesgo de enfermedades, la estimación de efectos crónicos a la salud, contrastar hipótesis causales y generar otras, además de aportar elementos muy útiles en la prevención de daños a la salud. Estos estudios pueden clasificarse en transversales: de prevalencia, y longitudinales: caso control y cohortes.
3. La realización de un estudio observacional analítico conlleva un riguroso proceso de planificación, que exige un profundo conocimiento, por parte de los investigadores, tanto del problema a estudiar como de metodología de la investigación científica.
4. Los estudios de prevalencia analíticos nos permiten, además de describir y estimar las características del evento estudiado, establecer la posible participación de

determinados factores o variables de interés en la producción de dicho evento. El análisis estadístico de estos estudios, se caracteriza por el uso de medidas de frecuencia y de asociación, entre las primeras se destacan las llamadas **oportunidades de enfermedad y de exposición**, que pueden ser calculadas tanto para los expuestos, como los no expuestos, y para los enfermos y no enfermos, respectivamente. Entre las medidas de asociación tenemos la **prevalencia relativa**, cuya interpretación es similar al riesgo relativo.

5. La desventaja fundamental de los estudios de prevalencia analíticos radica en que tanto el presunto factor de riesgo como la enfermedad (el efecto), se miden en el mismo momento, es decir, no se tiene en cuenta la secuencia temporal de los sucesos.
6. Los estudios de cohortes se caracterizan por el seguimiento de dos o más grupos de personas aparentemente sanas o libres del efecto de interés (enfermedad u otro daño a la salud), con diferentes grados de exposición a un presunto factor de riesgo (incluida la no exposición), que serán evaluados periódicamente durante un lapso de tiempo determinado, con el objetivo de medir la frecuencia de aparición de dicho efecto (tasas de incidencia) y establecer comparaciones entre los grupos examinados.
7. El objetivo general de los estudios de cohortes es establecer la presencia y magnitud de asociaciones causales, lo cual se realiza fundamentalmente mediante el cálculo del **riesgo relativo**, el cual expresa la probabilidad de contraer la enfermedad en el grupo expuesto con relación a los no expuestos. Además puede calcularse el **riesgo atribuible**, que si bien no mide la magnitud de la asociación, si nos brinda información acerca del efecto absoluto de la exposición sobre la frecuencia de aparición de la enfermedad.
8. Entre las ventajas de los estudios de cohortes tenemos que demuestran más claramente la secuencia temporal entre la exposición y la enfermedad, es posible calcular las tasas de incidencia de la enfermedad y minimizan las probabilidades de cometer sesgos de selección, entre otras. Como desventaja fundamental está el hecho de que pueden prolongarse en el tiempo con la consecuente pérdida de sujetos y (o) de investigadores, lo que encarece su realización y disminuye la validez de sus resultados, e incluso, puede suceder que ocurran otros descubrimientos durante el período de seguimiento que provoquen que los resultados obtenidos sean totalmente irrelevantes.
9. Los estudios de caso control se caracterizan por la comparación de la frecuencia de presentación de uno o más factores de riesgo en dos grupos de individuos: uno con la enfermedad o el efecto objeto de investigación, los casos, y otro sin dicha enfermedad, los controles. Su objetivo general es establecer relaciones causales entre los eventos estudiados. Suelen ser más fáciles de realizar y menos costosos que los de cohortes, pero no permiten calcular las tasas de incidencia y por tanto el riesgo relativo y son particularmente propensos al sesgo de selección.
10. Cualquiera sea el diseño empleado, es de particular importancia la selección de la población de referencia y de los sujetos a estudiar. Deben definirse claramente las fuentes de selección de los casos, controles, cohortes o muestra en general, según sea el caso, pues cada fuente tiene sus ventajas y limitaciones, además esta información es muy útil en el proceso inferencial. Por otro lado, deben quedar bien establecidas las definiciones de caso, no caso; la exposición, los grados de la misma; los criterios de selección y exclusión; pues de ello depende grandemente la probabilidad de sesgos.

Ejercitación

- De los objetivos enunciados a continuación seleccione los que se corresponden con los estudios observacionales analíticos.
 - Probar hipótesis etiológicas
 - Identificar personas o grupos con alto riesgo
 - Identificar factores de riesgo de una enfermedad
 - Identificar factores pronósticos de una enfermedad
 - Evaluar el alcance de los objetivos de la salud pública
 - Sugerir mecanismos de causalidad
 - Generar nuevas hipótesis etiológicas
- De las medidas empleadas en la investigación epidemiológica diga ¿cuáles son las de mayor uso en los estudios observacionales analíticos?
- A continuación te brindamos una lista de medidas de asociación que pueden ser calculadas en los estudios observacionales analíticos (**Columna A**) y los tipos de estudios en que pueden ser empleadas (**Columna B**). Relacione convenientemente ambas columnas teniendo en cuenta que es posible más de una medida por tipo de estudio. Anota en el guión de la columna A el número correspondiente de la columna B.

Columna A	Columna B
<input type="checkbox"/> Oportunidad relativa	1 - Estudios transversales
<input type="checkbox"/> Razón de productos cruzados	2 - Estudios de Caso - Control
<input type="checkbox"/> Prevalencia relativa de exposición	3 - Estudios de cohortes
<input type="checkbox"/> Riesgo relativo	
<input type="checkbox"/> Prevalencia relativa de enfermedad	
<input type="checkbox"/> Riesgo atribuible	

- Un grupo de investigadores del CITED³ estimó una prevalencia de discapacidad física del 35% en cierta área de salud de la capital. Ahora, desean conocer los factores de riesgo de esta afectación a la salud.
 - ¿Qué tipo de investigación usted les sugeriría?. Fundamente su respuesta
- Un grupo de investigadores considera a la hiperfiltración renal como un posible factor de riesgo de la insuficiencia renal crónica (IRC). Para estudiar esta relación diseñaron un estudio de casos y controles cuyos resultados se resumen en la siguiente tabla.

Filtración renal	IRC		Total
	Si	No	
Hiperfiltración	255	50	305
Normofiltración	45	250	295
Total	300	300	600

³ Son las siglas del Centro Internacional de la Tercera Edad que se encuentra en el hospital "Calixto García"

a) Calcule la(s) medida(s) que considere conveniente(s). Interpretela(s)

Autoevaluación

1. Mencione tres medidas de asociación empleadas en los estudios observacionales analíticos. Diga en qué tipo de estudios se pueden calcular.
2. Un grupo de investigadores está interesado en estimar la prevalencia de cierta enfermedad, tanto global como específica según un conjunto de variables seleccionadas.
 - a) ¿Qué tipo de investigación usted les sugeriría?
 - b) Mencione tres medidas que podrían utilizarse en tal empeño. Diga su interpretación teórica.
3. Luego de solucionado el problema anterior, los investigadores de referencia desean identificar los factores de riesgo de la enfermedad estudiada.
 - a) ¿Qué estudios podrían realizarse con tal finalidad?
 - b) Exponga esquemáticamente el diseño de uno de ellos.
4. Mencione dos sesgos que pueden presentarse en los estudios de casos y controles. Explíquelos brevemente.
5. El consumo excesivo de sal en la dieta es considerado como un factor de riesgo de la HTA. En un estudio de cohortes realizado para evaluar esta relación se obtuvo un riesgo relativo igual a 1.89. Interprete dicho resultado.

Bibliografía

1. Jenicek M D, Cléroux R. **Epidemiología. Principios, técnicas, aplicaciones.** Barcelona: Salvat Editores, S.A; 1987.
2. Streiner GR, Norman DL. **Bioestadística.** Madrid: Harcourt Brace. España; 1998
3. Rothman, K J. **Epidemiología moderna.** Madrid: Díaz de Santos, 1987.
4. Kleinbaum D G, Kupper L L, Morgenstern H. **Epidemiologic research, principles and quantitative methods.** New York: Van Nostrand Reinhold Co; 1982.
5. Argimon Pallas J M, Jiménez Villa J. **Métodos de investigación. Aplicados a la atención primaria de salud.** Barcelona: Ediciones Dayma; 1993
6. Beaglehole R, Bonita R, Kjellström. **Epidemiología básica.** Washington DC: OPS; 1996.
7. Evans R, Albornoz R. **Principios de Epidemiología Moderna.** Universidad central de Venezuela; 1994.
8. Colimon K M. **Fundamentos de la Epidemiología.** Madrid: Díaz de Santos, S.A; 1990.
9. Ahlbom A, Norell S. **Fundamentos de Epidemiología.** Madrid: siglo veintiuno de España editores, S.A; 1995.
10. Armitage P, Berry G. **Statistical Methods in Medical Research.** 3rd ed. Oxford: Blackwell Scientific Publications, 1994
11. Altman DG. **Practical statistics for medical research.** London: Chapman and Hall. 1992
12. Daniel WW. **Bioestadística. Base para el análisis de las ciencias de la salud.** 3ra edición. México. D. F: Limusa; 1997
13. Alonso Fernández A M. **Curso de Epidemiología.** Escuela de Estadística. Mérida, Venezuela; 1994.

14. Fleiss J L. **Statistical methods for rates and proportions**. 2nd Edition. New York: John Wiley & sons; 1981.
15. Shultz, James M. **Epicize. Essentials and exercises in Epidemiology**. 4th Edition. Weston, Florida; 1998.