



Departamento de Estadística
Universidad Carlos III de Madrid

BIOESTADISTICA (55 - 10536)

Medidas de frecuencia, asociación, e impacto

1. INTRODUCCION

Antes de comenzar con el tema propiamente, es conveniente recordar algunos conceptos que nos introducirán en los distintos tipos de mediciones utilizados en Epidemiología. Se han dado múltiples definiciones de Epidemiología, entre ellas una muy resumida es: "*Epidemiología es el estudio de la salud en las poblaciones humanas*". Sin embargo, nosotros utilizaremos la recogida en el Diccionario de Epidemiología de la Asociación Internacional de Epidemiología:

Estudio de la distribución y de los determinantes de los estados o acontecimientos relacionados con la salud de determinadas poblaciones, y la aplicación de este estudio al control de los problemas sanitarios.

En esta definición se reflejan los niveles de actuación de la investigación epidemiológica, que son:

- a) **Nivel Descriptivo:** Interesa describir, en forma cuantitativa y cualitativa, la distribución de un problema de salud-enfermedad con relación a variables de descripción básicas, como son las características de las personas, del lugar y la evolución temporal de la frecuencia de este problema.
- b) **Nivel de Conocimiento Etiológico:** El objetivo es inferir la participación causal o etiológica de las variables condicionantes en el problema de salud-enfermedad en estudio.
- c) **Nivel de Intervención:** Evaluación del impacto probable de las medidas tomadas en relación al problema de salud-enfermedad.

Cada uno de estos niveles de actuación necesita de formas de medición, que serán respectivamente: medidas de frecuencia, de asociación y de impacto potencial.

2. MEDIDAS DE FRECUENCIA

Medidas que caracterizan la aparición de una enfermedad o de otro atributo o hecho en una población, para un tiempo dado y una zona geográfica establecida. Existen dos maneras de medir la ocurrencia de un problema de salud: la incidencia y la prevalencia.

Incidencia (I): Número de casos de enfermedad que comienzan o de personas que caen enfermas durante un período dado en una población determinada. De modo más general, número de hechos nuevos, por ejemplo, nuevos casos de una enfermedad en una población definida dentro de un plazo determinado.

Prevalencia (P): Número de casos de una enfermedad determinada o de otra situación en una población y en un momento dados.

Ambas medidas pueden expresarse en números absolutos, sin embargo, tanto en la cuantificación del riesgo como en la carga de enfermedad, el uso de medidas relativas (proporciones o tasas) es más útil; de forma que en ocasiones encontraremos incidencia y prevalencia refiriéndose a tasa de incidencia y tasa de prevalencia, respectivamente.

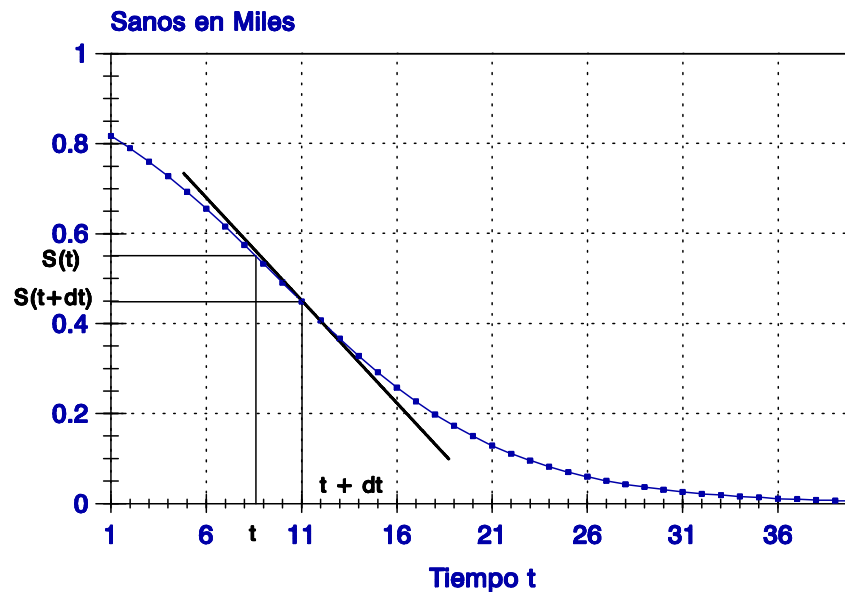
2.1. MEDIDAS DE INCIDENCIA DE USO COMUN

a) Densidad de Incidencia (DI): Medida teórica del número de nuevos casos que ocurren por unidad de tiempo-persona, por ejemplo, años-personas en riesgo. Se trata de una medida de la aparición de la enfermedad en un punto en el tiempo t , definida matemáticamente como el límite cuando Δt tiende a 0 de la siguiente expresión:

$$\frac{\text{Probabilidad de que una persona sana en el momento } t, \text{ desarrolle la enfermedad antes de } t+\Delta t}{\Delta t}$$

En la figura siguiente se muestra una distribución teórica de la cantidad de personas sanas en una cohorte $S(t)$, en este caso la densidad de incidencia está dada por:

$$\lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{S(t)\Delta t} = \frac{-S'(t)}{S(t)}$$



El valor medio de esta cantidad durante el intervalo t a $t + \Delta t$ puede estimarse por:

$$\frac{\text{Nuevos casos observados desde } t \text{ a } t + \Delta t}{\text{Número de unidades de tiempo-persona observación desde } t \text{ a } t + \Delta t}$$

Debido a la dificultad de expresar el tamaño de la población en estudio como una función matemática de t , se utiliza el valor medio de **DI**.

Asimismo, como sinónimos de la densidad de incidencia podemos encontrar: tasa de peligro, fuerza de morbilidad, tasa de incidencia instantánea y tasa de incidencia tiempo-persona.

b) Tasa de incidencia (TI): Tasa de nuevos casos en una población. El numerador es el número de nuevos hechos que se producen en un período determinado; el denominador es la población en riesgo de presentar el hecho de interés durante dicho período, y a veces se expresa en forma de tiempo persona. La tasa de incidencia que se usa más a menudo en salud pública se calcula por la fórmula:

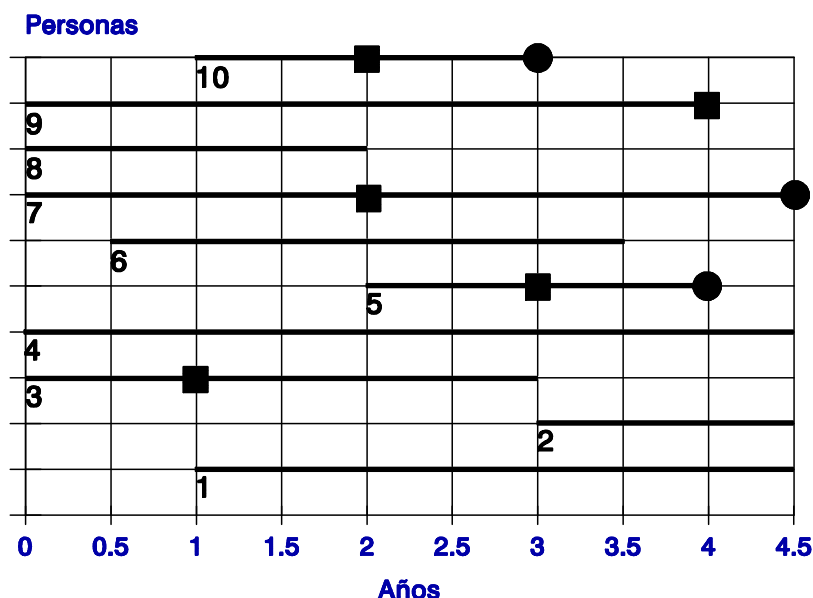
$$\frac{\text{Número de nuevos acontecimientos en un período determinado}}{\text{Número de personas expuestas al riesgo durante ese período}}$$

Como vemos la **TI** no es más que el valor medio de la **DI** en un período determinado, por esta razón en la mayoría de estudios se dice que calculan la **DI** cuando en realidad están calculando la **TI**.

El denominador de esta tasa se puede obtener según el tipo de diseño:

- i) Si el diseño permite conocer la duración del tiempo de observación de cada uno de los individuos de la población, entonces el denominador será la suma de los tiempos individuales de observación.

EJEMPLO: En la figura siguiente se representa una población de diez personas seguidas durante 4.5 años y que al inicio del estudio estaban libres de enfermedad, ■ indica el punto en que la enfermedad es diagnosticada, ● indica el punto en que la persona muere y - representa la observación de una persona.



En este caso conocemos el tiempo de observación de las personas y el momento de contraer la enfermedad, por tanto el denominador se calcula por la siguiente suma: Persona-años en riesgo = $3.5 + 1.5 + 1 + 4.5 + 1 + 3 + 2 + 2 + 4 + 1 = 23.5$, y la densidad de incidencia de la enfermedad está dada por: $DI = 5/23.5 \approx 0.213 \text{ año}^{-1}$.

- ii) Si se tiene una población estable en el tiempo (no cambia en número de individuos ni en la distribución por edades), entonces el denominador será el producto del número de personas al inicio del período por la longitud de este.

EJEMPLO: En los estudios de seguimientos de trabajadores de un tipo específico de industria y la aparición de una enfermedad particular, en particular, si tomamos 1000 trabajadores de una mina de plomo y en un período de seguimiento de 4 años aparecen 10 casos de saturnismo, obtenemos $DI = 10/1000 \times 4 = 0.0025 \text{ año}^{-1}$.

- iii) Si se tiene una población dinámica, entonces el denominador será el tamaño medio de la población, a menudo la población a la mitad del período multiplicado por la longitud del período de observación.

EJEMPLO: Se desea calcular la DI en un período de 3 años de una población que al inicio del estudio tenía 10000 habitantes y al final 15000 habitantes y durante ese período hubo 100 casos de la enfermedad. El tiempo-persona en riesgo es igual a: $3 \times (10000 + 15000)/2 = 37500$, y $DI = 100/37500 \approx 0.00267 \text{ año}^{-1}$.

- c) Incidencia Acumulada (IA):** Número o proporción de personas en un grupo, en las que comienza un hecho relacionado con la salud durante un intervalo de tiempo determinado. El denominador está constituido por el número de individuos sin la enfermedad al comienzo del período de observación. Esta es la forma en que habitualmente se expresa la incidencia de una enfermedad y es equivalente al riesgo promedio de contraer la enfermedad de cada individuo del grupo en estudio.

EJEMPLO: Se realiza un encuesta serológica en 10000 mujeres con edades entre 15 y 30 años. Se encuentra 10 mujeres positivas al antígeno de superficie al virus de Hepatitis B (AgsHB). Estas mujeres se siguen por un período de 10 años, en el que aparecen 20 nuevos casos positivos a AgsHB. $IA = 20/(10000-10) \approx 0.002$ en 10 años.

2.2. MEDIDAS DE PREVALENCIA DE USO COMUN

- a) Tasa de Prevalencia (TP):** Número total de individuos que presentan un evento o enfermedad en un momento determinado, dividido por la población en riesgo de tener el evento en dicho momento. Se puede expresar como el cociente entre el número de casos nuevos y antiguos y la población en estudio (prevalencia puntual **PP**).

$$PP = \frac{\text{Número de casos nuevos y antiguos}}{\text{Población en estudio}}$$

- b) Prevalencia lápsica (PL):** Se utiliza en estudios de prevalencia, cuando durante el proceso de determinación de los casos en la población, se originan casos nuevos de la enfermedad. Si se incluyen estos en el numerador estamos midiendo la prevalencia lápsica.

$$PL = \frac{\text{Número de casos prevalentes} + \text{Número de casos nuevos}}{\text{Población en estudio}}$$

EJEMPLO: Supongamos que realizamos un estudio de transversal, o de encuesta, para estimar la prevalencia de resfriado común en una escuela secundaria con 600 alumnos, que fueron distribuidos en 6 grupos de 100 alumnos cada uno. El estudio tiene una duración de 6 días, cada día se examinarán los alumnos de un grupo, aunque en los días consecutivos se reportarán por los profesores, padres y los propios alumnos de los grupos ya examinados la aparición de síntomas compatibles con la enfermedad. En la siguiente tabla se presentan los resultados del estudio:

	1	2	3	4	5	6	Total
1 Grupo	10	1	2	1	0	1	15
2 Grupo	-	6	1	2	1	0	10
3 Grupo	-	-	11	1	0	3	15
4 Grupo	-	-	-	8	1	1	10
5 Grupo	-	-	-	-	10	5	15
6 Grupo	-	-	-	-	-	10	10
Totales	10	7	14	12	12	20	75

En las columnas se representan los casos de resfriado común detectados y reportados, entonces la diagonal representa los casos detectados en el estudio, y el resto los reportados. Si calculamos la prevalencia puntual sólo incluimos los casos detectados, o sea, $10+6+11+8+10+10 = 55$, $PP = 55/600 \approx 0.092$.

Si calculamos la prevalencia lápsica, o sea considerando los casos que se reportaron en los días siguientes al análisis, tendremos $PL = 75/600 = 0.125$.

3. MEDIDAS DE ASOCIACION

Medidas que expresan la fuerza de la asociación entre variables. Se entiende por asociación a la *dependencia estadística* entre dos sucesos, o variables. Consideraremos dos tipos de medidas de asociación: el riesgo relativo (**RR**) y el riesgo atribuible (**RA**).

Riesgo Relativo (RR): Se obtiene del cociente entre el riesgo de enfermedad o muerte de un grupo que denominaremos expuestos, y el riesgo en los no expuestos, este modo de utilizar el término es sinónimo de razón de riesgos. Alternativamente, se obtiene de la razón de la **IA** en expuestos dividida por la **IA** en no expuestos. El término **RR** se ha utilizado como sinónimo de razón de ventaja, y de razón de **DI**; esto se debe a que en enfermedades "raras" (por ejemplo la mayoría de las neoplasias malignas) estas cantidades son aproximadamente iguales.

Así, definiremos las siguientes dos medidas de asociación:

i) Riesgo Relativo:
$$RR = \frac{\text{Tasa de incidencia acumulada en individuos expuestos}}{\text{Tasa de incidencia acumulada en indiv. no expuestos}}$$

ii) Razón de Densidad de Incidencia:
$$RDI = \frac{\text{Densidad de incidencia en individuos expuestos}}{\text{Densidad de incidencia en no indiv. expuestos}}$$

Por tanto en un estudio de seguimiento en dependencia de la medida de frecuencia que se utilice se tomará el **RR** o **RDI**.

EJEMPLO: Supongamos que en la figura anterior, donde se representa una población de diez personas seguidas durante 4.5 años, ahora los individuos representados en la se dividen en fumadores los números impares 1-3-5-7-9 y en no fumadores los números pares 2-4-6-8-10, la RDI en este caso;

Tenemos que la densidad de incidencia en fumadores se calcula por:

$$\text{Personas-años}_{\text{expuestos}} = 3.5 + 1 + 1 + 2 + 4 = 11.5$$

$$DI_{\text{expuestos}} = 4/11.5 \approx 0.348 \text{ años}^{-1}$$

Tenemos que la densidad de incidencia en no fumadores se calcula por:

$$\text{Personas-años}_{\text{no expuestos}} = 1.5 + 4.5 + 3 + 2 + 1 = 12$$

$$DI_{\text{no expuestos}} = 1/12 \approx 0.083 \text{ años}^{-1}$$

$$RDI = 0.348/0.083 \approx 4.193$$

Si consideramos en este ejemplo el cálculo de la incidencia acumulada, obtendremos:

$$IA_{\text{expuestos}} = 4/5 = 0.8$$

$$IA_{\text{no expuestos}} = 1/5 = 0.2$$

$$RR = 0.8/0.2 = 4$$

En estudios de cohortes o de seguimiento, donde se conocen las poblaciones expuestas y no expuestas, el cálculo de **RDI** o **RR** se realiza como sigue:

	Casos	No Casos	Total	Tiempo en observación
Expuestos	a	b	n_1	l_1
No Expuestos	c	d	n_0	l_0
Total	m_1	m_0	n	l

$$DI_{\text{expuestos}} = a/l_1$$

$$DI_{\text{no expuestos}} = c/l_0$$

$$RDI = \frac{a/l_1}{c/l_0}$$

$$IA_{\text{expuestos}} = a/n_1$$

$$IA_{\text{no expuestos}} = c/n_0$$

$$RR = \frac{a/n_1}{c/n_0}$$

En estudios de casos-controles, donde no se conocen a priori las poblaciones de expuestos y no expuestos y por tanto no es posible calcular las tasas de incidencia, se utiliza otra medida que denominamos Razón de Odds.

Razón de Odds (OR): Cociente de las probabilidades en favor de la exposición entre los casos a/c , divididas por las existentes entre los controles b/d , lo que queda reducido a $\frac{ad}{bc}$.

	Casos	Controles	Total
Expuestos	a	b	n_1
No Expuestos	c	d	n_0
Total	m_1	m_0	n

Si representamos con una E la exposición al factor en estudio y por D la enfermedad, **OR** se expresa por: $RD = \frac{\Pr(E|D)/\Pr(\bar{E}|D)}{\Pr(E|\bar{D})/\Pr(\bar{E}|\bar{D})}$.

Notemos a su vez que $\Pr(E|D) = \frac{a}{m_1}$, $\Pr(\bar{E}|D) = \frac{c}{m_1}$, $\Pr(E|\bar{D}) = \frac{b}{m_0}$, y $\Pr(\bar{E}|\bar{D}) = \frac{d}{m_0}$, de donde obtenemos nuevamente **OR** = $\frac{ad}{bc}$.

Se demuestra que en estudios de casos-controles de enfermedades "raras" ($\Pr(D) < 0.1$) el **OR** es una estimación aceptable del **RR**. Veamos esta demostración. Utilizando el teorema de Bayes obtenemos:

$$\Pr(D|E) = \frac{\Pr(D)\Pr(E|D)}{\Pr(D)\Pr(E|D) + \Pr(\bar{D})\Pr(E|\bar{D})}$$

y

$$\Pr(D|\bar{E}) = \frac{\Pr(D)\Pr(\bar{E}|D)}{\Pr(D)\Pr(\bar{E}|D) + \Pr(\bar{D})\Pr(\bar{E}|\bar{D})}$$

Sabemos que $RR = \frac{\Pr(D|E)}{\Pr(D|\bar{E})}$, sustituyendo las expresiones anteriores tenemos:

$$RR = \frac{\Pr(E|D)}{\Pr(\bar{E}|D)} \left[\frac{\Pr(D)\Pr(\bar{E}|D) + \Pr(\bar{D})\Pr(\bar{E}|\bar{D})}{\Pr(D)\Pr(E|D) + \Pr(\bar{D})\Pr(E|\bar{D})} \right]$$

Luego, si $\Pr(D) \approx 0$, podemos simplificar esta expresión:

$$RR \approx \frac{\Pr(E|D)/\Pr(\bar{E}|D)}{\Pr(E|\bar{D})/\Pr(\bar{E}|\bar{D})} = OR$$

Este resultado justifica en alguna manera el uso en la literatura del término RR en trabajos donde se calcula la OR .

EJEMPLO: Para estudiar la asociación positiva en un estudio de casos-controles de consumo de cigarrillos y cáncer de bucofaringe, se obtiene la siguiente distribución:

	Casos	Controles	Total
Expuestos	352	228	580
No Expuestos	48	122	170
Total	400	350	750

Así, la $OR = 352 \times 122 / 228 \times 48 \approx 3.92$, esto indica que el riesgo de desarrollar cáncer de bucofaringe en fumadores es 3.92 veces más elevado que el de los no fumadores.

Riesgo Atribuible (RA): Tasa de una enfermedad u otro evento de salud en expuestos, que pueda atribuirse a dicha exposición. Esta medida es la diferencia entre la tasa de una enfermedad entre los individuos expuestos y la misma tasa en los no expuestos; se supone que las demás variables diferentes de la que se investiga ejercen el mismo efecto sobre ambos grupos. Se conoce también por diferencias de tasas.

De manera similar al riesgo relativo se definen, según la medida de frecuencia que se utiliza, las siguientes medidas:

- i) $RA = \text{Incidencia Acumulada en expuestos} - \text{Incidencia Acumulada en no expuestos}$
- ii) $DDI = \text{Densidad de Incidencia en expuestos} - \text{Densidad de Incidencia en no expuestos}$

EJEMPLO: De nuevo considere el ejemplo de la figura anterior, donde se representa una población de diez personas seguidas durante 4.5 años, donde obtendríamos: $RA = 0.6$ y $DDI = 0.265$, respectivamente.

4. MEDIDAS DE IMPACTO

Estas medidas reflejan el efecto esperado al cambiar uno o más factores de riesgo o realizar una acción de carácter preventivo en una población particular.

Fracción Atribuible (Etiológica) en Expuestos (*FAe*): Es la proporción en que se reduciría entre los expuestos la tasa de incidencia del evento de salud si se eliminara la exposición. También recibe el nombre de riesgo atribuible porcentual en los expuestos (*RA* %). Es una medida que se obtiene a partir del *RA* y se puede calcular con las siguientes fórmulas:

$$FAe = \frac{TI \text{ en expuestos} - TI \text{ en no expuestos}}{TI \text{ en expuestos}}$$

y

$$FAe = \frac{RR - 1}{RR}$$

Fracción Atribuible (Etiológica) en Población (*FAP*): Es la proporción en que se reduciría en la población la tasa de incidencia del evento de salud si se eliminara la exposición. También recibe el nombre de Riesgo Atribuible de Levin o poblacional porcentual (*RAP* %). Se puede calcular con las siguientes fórmulas:

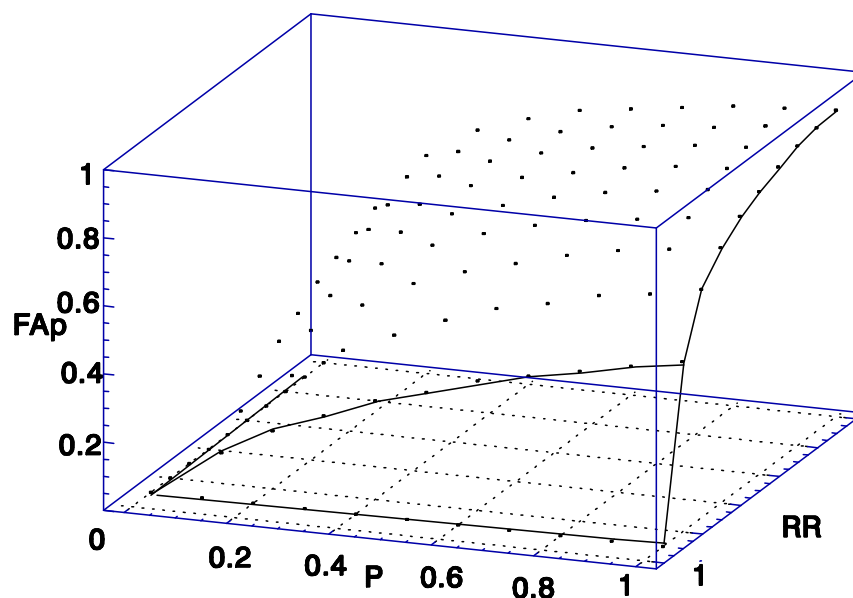
$$FAP = \frac{TI \text{ global} - TI \text{ en no expuestos}}{TI \text{ global}}$$

y

$$FAP = \frac{P(RR-1)}{1+P(RR-1)}$$

Donde *P* es la prevalencia del factor en estudio y *RR* es la razón entre la *TI* global y *TI* en no expuestos. En estudios de casos-controles en que los casos y controles sean representativos del universo se puede sustituir por *RD*.

La expresión anterior nos dice que el impacto en la población depende de la prevalencia del factor y de su asociación con la enfermedad. En la figura, vemos como el *FAP* es más pronunciado cuando mayor es *P*, y cuando mayor es *RR*.



Debemos enfatizar que estas medidas deben interpretarse en conjunto, pues si bien, por ejemplo, la exposición a sustancias radiactivas está fuertemente asociada a la aparición de cáncer, el impacto poblacional de una medida tendiente a evitar la exposición a esos materiales, será pobre debido a que la exposición en población general es muy pequeña.

Fracción Evitada en Población (FEP): Cuando estamos en una situación en que se cree que la exposición a un determinado factor protege frente a una enfermedad, la fracción evitada es la proporción de la carga hipotética total de la enfermedad en la población que se ha evitado con la exposición a ese factor. Se conoce también por fracción de prevención poblacional porcentual (**FPP** %). Se puede calcular con la siguiente fórmula:

$$FEP = \frac{TI \text{ en no expuestos} - TI \text{ global}}{TI \text{ en no expuestos}}$$

Esta medida nos muestra la proporción de nuevos casos potenciales que se han prevenido por la exposición a un factor determinado, lo que permite evaluar la importancia a nivel colectivo de los factores asociados con un **RR** menor que 1. Los resultados obtenidos tienen que interpretarse con cuidado, dado que parte o la totalidad del aparente efecto protector pueden deberse a otros factores asociados a éste.