

# Medidas de frecuencia y de asociación en epidemiología clínica

MANUEL ENRIQUE FUENTES FERRER Y NÁYADE DEL PRADO GONZÁLEZ

Unidad de Metodología de Investigación y Epidemiología Clínica. Servicio de Medicina Preventiva. Hospital Clínico San Carlos. Instituto de Investigación Sanitaria. Hospital Clínico San Carlos (IdISSC). Madrid. España.

mfuentesferrer@gmail.com; nayade.hcsc@gmail.com

## Introducción

La epidemiología es la disciplina científica que estudia la frecuencia y distribución de fenómenos relacionados con la salud y sus determinantes en poblaciones específicas, y su aplicación al control de problemas de salud<sup>1</sup>. Dentro de sus aplicaciones básicas, se encuentra la medición de la frecuencia de la enfermedad u otros fenómenos de salud en la población y cuantificar la relación existente entre diferentes factores de riesgo y la enfermedad. Es necesario que el profesional de la medicina conozca y maneje los índices epidemiológicos existentes para el cálculo de la frecuencia

de fenómenos de salud y de magnitud de asociación entre variables que se utilizan en los diferentes estudios epidemiológicos que se publican en la literatura científica. De esta manera, podrá valorar la relevancia clínica de dichos índices obtenidos en estudios descriptivos y analíticos.

## Medidas de frecuencia de la enfermedad

Un aspecto importante de la epidemiología es la medición de la frecuencia de las enfermedades en el campo de la salud. Los índices matemáticos básicos que se utilizan para medir dichos fenómenos se dividen básicamente en tres tipos: proporción, razón y tasa. La proporción es un cociente donde el numerador está incluido en el denominador y expresa la probabilidad de ocurrencia un suceso. Es un índice adimensional y con un rango entre 0 y 1. La *razón* es todo índice obtenido al dividir 2 cantidades donde el numerador no está incluido en el denominador. Puede tener o no dimensiones. Un tipo especial de razón es la *odds*, que se calcula dividiendo una proporción entre su complementaria ( $odds = \text{proporción} / [1 - \text{proporción}]$ ) y expresa el número de veces que ha ocurrido un suceso por cada vez que no ha ocurrido. Por último, la tasa es un cociente que lleva incorporado en el denominador la variable de tiempo y expresa el cambio de la variable del numerador por unidad de cambio de la variable del denominador (medida de rapidez de cambio) de un fenómeno de salud. Es dimensional (inverso del tiempo o tiempo<sup>-1</sup>) y con un rango de 0 a infinito<sup>2</sup>.

Las medidas de frecuencia representan la ocurrencia de un fenómeno de salud (enfermedad, trastorno o muerte) en poblaciones y, por lo tanto, son fundamentales para las investigaciones descriptivas y analíticas<sup>3</sup>. Describen un evento de enfermedad o salud en relación con el tamaño de una población a riesgo. Dicha descripción se puede realizar desde una perspectiva *transversal* a través de la obtención de datos en un momento temporal (índices estáticos) o desde una perspectiva de *seguimiento* mediante el registro de datos

### Puntos clave

- Las medidas de frecuencia de la enfermedad miden la ocurrencia de un fenómeno de salud en la población.
- La prevalencia y la incidencia son 2 medidas referenciadas ampliamente en la literatura científica que responden a dos preguntas diferentes: a) ¿qué porcentaje de un grupo de individuos experimenta un proceso en un momento determinado de tiempo?, y b) ¿a qué ritmo aparecen nuevos casos en un grupo de individuos a medida que transcurre el tiempo?
- La incidencia acumulada refleja el riesgo de desarrollar una enfermedad en un periodo. Y la tasa de incidencia refleja la velocidad de aparición de un determinado evento en la población.
- Las medidas de asociación o efecto cuantifican la magnitud de la relación entre un factor de exposición (variable independiente) y una enfermedad (variable dependiente).
- El cálculo de las medidas de asociación varía en función del tipo de estudio epidemiológico.

**Tabla 1.** Medidas de frecuencia y efecto en función del tipo de estudio epidemiológico

	Transversal o de prevalencia	Cohortes observacional y experimental		Casos y controles	
<b>Medida de frecuencia</b>	Prevalencia	Incidencia acumulada	riesgo	Densidad o tasa de incidencia	Prevalencia de exposición
<b>Medidas de efecto o de asociación</b>					
<b>Absoluta</b>	Diferencias de prevalencias $P_1 - P_0$	Diferencia absoluta de riesgos (riesgo atribuible) $I_1 - I_0$		Diferencia absoluta de tasas $DI_1 - DI_0$	
<b>Relativa</b>	Razón de prevalencias $\frac{P_1}{P_0}$	Riesgo relativo $\frac{I_1}{I_0}$		Razón de tasas $\frac{DI_1}{DI_0}$	Odds ratio $\frac{\text{Odds ex p casos}}{\text{Odds ex p controles}}$

DI<sub>1</sub>: densidad o tasa de incidencia en expuestos; DI<sub>0</sub>: densidad o tasa de incidencia en no expuestos; exp: exposición; I<sub>1</sub>: incidencia en expuestos; I<sub>0</sub>: incidencia en no expuestos; P<sub>1</sub>: prevalencia en expuestos; P<sub>0</sub>: prevalencia en no expuestos.

en un periodo fijo o variable (índices dinámicos). La tabla 1 muestra las medidas de frecuencia de la enfermedad en función del tipo de estudio epidemiológico.

## Prevalencia

La prevalencia (P) es la proporción del grupo de individuos que presentan un proceso clínico o resultado en un momento determinado del tiempo (t). La prevalencia se determina mediante el muestreo representativo de una población definida en un momento determinado que contiene individuos con y sin el problema bajo estudio<sup>4</sup>. Representa la medida de frecuencia de la enfermedad que se calcula en los estudios transversales o de prevalencia. Expresa la probabilidad de que un individuo en una población tenga el fenómeno de interés en el tiempo t. Existen 2 tipos de prevalencia. La *prevalencia de punto o puntual* se determina en el momento del muestreo para cada individuo, aunque no necesariamente en el mismo momento para todos los individuos de la población definida. La *prevalencia de periodo* hace referencia a los casos presentes en cualquier momento durante un periodo específico.

$$\text{prevalencia de punto} = \frac{\text{Nº de casos presentes de enfermedad}}{\text{Total de la población estudiada}}$$

$$\text{prevalencia de periodo} = \frac{\text{Nº de casos nuevos} + \text{Nº de casos presentes de enfermedad}}{\text{Total de la población estudiada}}$$

## Incidencia

Las medidas de incidencia reflejan la dinámica de ocurrencia de un fenómeno en una población determinada<sup>5</sup>. Existen 2 medidas de incidencia, la incidencia acumulada (IA) y la tasa o densidad de incidencia.

La IA es una proporción que refleja el porcentaje de sujetos de una población, susceptible o en riesgo de desarrollar un

determinado fenómeno de salud, que desarrollan el evento a lo largo de un determinado periodo. Por lo tanto, la incidencia hace referencia a los nuevos casos de una enfermedad que se producen en una población inicialmente libre de enfermedad en un periodo. La incidencia se calcula identificando un grupo de individuos en riesgo de desarrollar una enfermedad (es decir, sujetos sin la enfermedad), examinándoles periódicamente durante un intervalo para identificar los nuevos casos de la enfermedad<sup>4</sup>.

$$\text{incidencia acumulada} = \frac{\text{Nº de casos nuevos de enfermedad}}{\text{Población susceptible de desarrollar la enfermedad}}$$

La incidencia es la medida de frecuencia de la enfermedad de elección en los estudios observacionales de cohortes y en los estudios experimentales para la valoración de factores etiológicos, factores pronósticos y la eficacia de intervenciones terapéuticas. La IA refleja el riesgo de desarrollar un determinado fenómeno de salud en un periodo de tiempo en una población.

Un segundo enfoque para estimar la incidencia es el cálculo de la tasa de incidencia que refleja el potencial instantáneo de cambio al estado de enfermedad (casos nuevos) por unidad de tiempo, relativos al tamaño de la población. Se calcula dividiendo el número de casos nuevos ocurridos durante un periodo de seguimiento entre la suma de todos los periodos de observación de los sujetos bajo estudio.

$$\text{tasa de incidencia } (t_0, t_1) = \frac{\text{Nº casos nuevos}}{\sum_{j=1}^N \Delta t_j}$$

donde t<sub>0</sub> indica el momento de inicio del periodo de seguimiento y t<sub>1</sub> el momento del final de seguimiento, por lo que el sumatorio hace referencia a la suma por paciente de todos los periodos de seguimiento de los sujetos en riesgo de desarrollar la enfermedad. Aquellos individuos pertenecientes a la población que se convierten en nuevos casos

contribuirán en el denominador de la tasa con el rango de tiempo desde su inicio de seguimiento hasta el desarrollo del evento (tiempo en riesgo). Por lo tanto, la tasa de incidencia refleja la velocidad de aparición de un determinado evento en la población.

## Medidas de asociación o efecto en los estudios epidemiológicos

En la mayoría de los estudios epidemiológicos el investigador está interesado en medir el grado de asociación entre uno o varios factores independientes (exposición) y la aparición de una enfermedad (variable dependiente). Para ello se calculan las medidas de asociación que cuantifican dicha relación.

Desde el punto de vista de la estadística, se puede evaluar, a través de un contraste de hipótesis, cómo de probable es que dicho efecto que se observa o uno mayor se haya producido por azar, suponiendo que el efecto no existiera. Esa probabilidad es la que se conoce como significación estadística y se representa mediante el conocido valor *p*. El problema es que el nivel de significación (la *p*) no informa sobre la magnitud del efecto y, por tanto, de si resulta clínicamente relevante o no, simplemente de si este resulta estadísticamente significativo, pudiendo obtener efectos pequeños (que no son clínicamente relevantes) estadísticamente significativos si el tamaño muestral es grande, y efectos grandes no estadísticamente significativos si tenemos tamaños muestrales pequeños<sup>6</sup>.

Las diferentes medidas de asociación que se pueden calcular varían en función del tipo de estudio epidemiológico. A grandes rasgos, se pueden dividir en medidas de efecto relativas (basadas en cocientes) y medidas de efecto absolutas (basadas en diferencias). Para el cálculo de las medidas de asociación los datos suelen presentarse en tablas de contingencia (2 x 2) (tabla 2).

## Medidas de efecto relativas

### Riesgo relativo

Es la medida de elección en los estudios observacionales de cohortes y en los estudios experimentales. Se calcula dividiendo la incidencia de la enfermedad en expuestos ( $I_1$ ) entre la incidencia de enfermedad en no expuestos ( $I_0$ ) (tabla 2).

$$RR = \frac{\text{incidencia en expuestos}}{\text{incidencia en no expuestos}} = \frac{\frac{a}{a+b}}{\frac{c}{c+b}}$$

Expresa el número de veces que es más probable que una enfermedad se desarrolle en el grupo de expuestos en relación con el grupo de no expuestos. Es un cociente que solo puede tomar valores positivos. Si la exposición es un factor de riesgo de esa enfermedad, entonces su valor será superior a 1 (riesgo

relativo [RR] > 1), mientras que si la exposición es un factor de protección de la enfermedad su valor será inferior a 1 (RR < 1). Si el RR toma un valor neutro (RR = 1), indica que no existe asociación entre la exposición y la enfermedad, ya que la incidencia de expuestos es la misma que los no expuestos. Cuando la medida de frecuencia de incidencia es la tasa o densidad de incidencia, el RR se calcula de manera similar dividiendo la tasa de incidencia de expuestos entre la tasa de no expuestos; en este caso, se denomina razón de tasas (RT).

### Razón de prevalencias

La razón de prevalencias (RP) es la medida de asociación que se puede calcular en los estudios transversales o de prevalencia. Se calcula dividiendo la prevalencia de enfermedad en expuestos ( $P_1$ ) entre la prevalencia de enfermedad en no expuestos ( $P_0$ ) (tabla 2).

$$RP = \frac{\text{prevalencia en expuestos}}{\text{prevalencia en no expuestos}} = \frac{\frac{a}{a+b}}{\frac{c}{c+b}}$$

La interpretación cuantitativa numérica es muy similar al RR o la razón de tasa. Su interpretación cualitativa difiere de la del RR porque en un estudio transversal no se puede realizar ninguna afirmación sobre el riesgo de enfermar, sino solo sobre el riesgo de padecer la enfermedad (prevalencia) en el momento del estudio.

### Odds ratio

En los estudios de cohorte el RR se estima directamente, ya que se conoce la incidencia de la enfermedad en los individuos expuestos y no expuestos a un factor. Por el contrario, en los estudios de casos y controles no se puede calcular la incidencia, porque la población de estudio se selecciona a partir de individuos que ya han desarrollado la enfermedad<sup>2</sup>. La *odds ratio* [OR] no es más que la razón entre la *odds* de exposición observada en los casos expuestos y la *odds* de exposición en el grupo control. Se interpreta en una escala multiplicativa como el número de veces que es mayor la *odds* de exposición en los casos que en los controles (tabla 2).

$$OR = \frac{\text{odds de exposición en los casos}}{\text{odds de exposición en los controles}} = \frac{\frac{\frac{a}{a+b}}{\left(1 - \frac{a}{a+b}\right)}}{\frac{\frac{b}{b+d}}{\left(1 - \frac{b}{b+d}\right)}} = \frac{a * d}{b * c}$$

La OR es una medida que se puede calcular en estudios de cohortes (OR de enfermedad) y en estudios transversales (OR de prevalencia). En el caso de estudios de cohortes, su interpretación es similar a la del RR, aunque la OR es solo una buena aproximación al RR en determinadas circunstancias. Si el riesgo de la enfermedad en la población de estudio es bajo (< 10%), ambas medidas son equivalentes. A medida

**Tabla 2.** Presentación de datos en una tabla de 2 x 2 para el cálculo de las medidas de asociación

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

Las casillas a, b, c y d representan diferentes combinaciones entre la enfermedad y la exposición:

a: número de personas expuestas que tienen la enfermedad; b: número de personas expuestas que no tienen la enfermedad; c: número de personas no expuestas que tienen la enfermedad; d: número de personas no expuestas que no tienen la enfermedad.

que aumenta la incidencia de la enfermedad, el cálculo del OR sobrestima la magnitud del efecto en relación con el RR<sup>7</sup>.

## Medidas de efecto absolutas

### Riesgo atribuible o diferencia de riesgo en los expuestos

A la pregunta del riesgo que tienen los expuestos, ¿cuánto se debe a la exposición?, se responde mediante el cálculo del riesgo atribuible (RA). El RA sigue siendo un riesgo (una incidencia acumulada) y se interpreta como el riesgo que se podría evitar si el grupo de expuestos no hubiera estado expuesto. Su interpretación cuantitativa y numérica es la siguiente: valor > que 0 indica que el factor es de riesgo, valor = 0 indica un efecto nulo, y valor < 0 indica un factor protector<sup>8</sup>.

$$\text{Riesgo atribuible} = \text{Incidencia en expuestos} - \text{Incidencia en no expuestos}$$

La tabla 1 muestra las principales medidas de efecto en función del tipo de diseño epidemiológico.

## Bibliografía



● Importante ●● Muy importante

1. Last JM. A Dictionary of Epidemiology. 4th ed. International Epidemiological Association. Oxford University Press; 2001.

2. ●● Armigón Pallás JM, Jiménez Villa J. Métodos de investigación clínica y epidemiológica. 3.ª ed. Madrid: Elsevier S.A.; 2000.
3. Kleinbaum DG, Kupper LL, Morgentern H. Epidemiology research. Principles and quantitative methods. New York: Van Nostrand Reinhold Company; 1982.
4. ● Fletcher RH, Fletcher SW, Wagner EH. Epidemiología Clínica. Aspectos fundamentales. 2.ª ed. Barcelona: Masson-Williams & Wilkins España, S.A.; 1988.
5. Tapia Granados JA. Incidencia; concepto, terminología y análisis dimensional. Med Clin (Barc). 1994;103:140-2.
6. ● Martínez-González MA, De Irala J, Faulin Fajardo FJ. Bioestadística Amigable. 1.ª ed. Madrid: Díaz de Santos S.A.; 2001.
7. Szklo M, Nieto J. Epidemiología intermedia. Conceptos y aplicaciones. 1.ª ed. Madrid: Díaz de Santos S.A.; 2003.
8. Mirón Canelo JA, Alonso Sardón M. Medidas de frecuencia, asociación e impacto en investigación aplicada. Med Segur Trab. 2008;54:93-102.

## Bibliografía recomendada

Armigón Pallás JM, Jiménez Villa J. Métodos de investigación clínica y epidemiológica. 3.ª ed. Madrid: Elsevier S.A.; 2000.

*Este libro en castellano desarrolla los fundamentos del método científico para aquellos profesionales que se inician en el campo de la investigación en Ciencias de la Salud. Es un libro sencillo y didáctico, en la medida de lo posible, pero a la vez riguroso. Los contenidos del libro tratan todos los aspectos principales del proceso de investigación: pregunta de investigación, tipos de estudios epidemiológicos, elaboración de protocolos de investigación e interpretación de resultados.*

Martínez-González MA, De Irala J, Faulin Fajardo FJ. Bioestadística Amigable. 1.ª ed. Madrid: Díaz de Santos S.A.; 2001.

*Este manual trata de desarrollar de manera asequible una materia que habitualmente resulta ardua para el profesional sanitario que está en fase de formación. Desarrolla los aspectos de estadística descriptiva e inferencial (intervalos de confianza y contraste de hipótesis) mediante ejemplos sencillos y aplicados. Además de desarrollar los cálculos estadísticos, describe los comandos de programas de análisis estadístico habituales en Ciencias de la Salud, como el SPSS o el Excel.*