

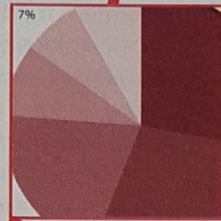
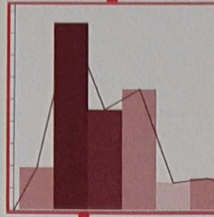
Haz clic en la imagen para abrir tu plataforma

## Estadística y epidemiología ENARM - México

### Manual CTO de Medicina y Cirugía

3.<sup>a</sup> edición





# Manual CTO de Medicina y Cirugía

3.<sup>a</sup> edición

## Estadística y epidemiología ENARM - México

### Revisor

Luis Cabeza Osorio

### Autores

Manuel Moreu Gamazo  
Antonio Martín Conejero

## NOTA

La medicina es una ciencia sometida a un cambio constante. A medida que la investigación y la experiencia clínica amplían nuestros conocimientos, son necesarios cambios en los tratamientos y la farmacoterapia. Los editores de esta obra han contrastado sus resultados con fuentes consideradas de confianza, en un esfuerzo por proporcionar información completa y general, de acuerdo con los criterios aceptados en el momento de la publicación. Sin embargo, debido a la posibilidad de que existan errores humanos o se produzcan cambios en las ciencias médicas, ni los editores ni cualquier otra fuente implicada en la preparación o la publicación de esta obra garantizan que la información contenida en la misma sea exacta y completa en todos los aspectos, ni son responsables de los errores u omisiones ni de los resultados derivados del empleo de dicha información. Por ello, se recomienda a los lectores que contrasten dicha información con otras fuentes. Por ejemplo, y en particular, se aconseja revisar el prospecto informativo que acompaña a cada medicamento que deseen administrar, para asegurarse de que la información contenida en este libro es correcta y de que no se han producido modificaciones en la dosis recomendada o en las contraindicaciones para la administración. Esta recomendación resulta de particular importancia en relación con fármacos nuevos o de uso poco frecuente. Los lectores también deben consultar a su propio laboratorio para conocer los valores normales.

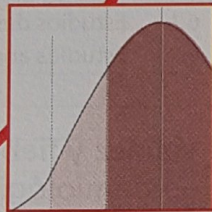
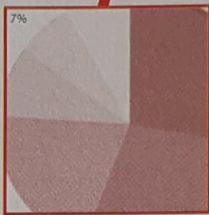
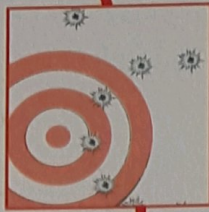
No está permitida la reproducción total o parcial de este libro, su tratamiento informático, la transmisión de ningún otro formato o por cualquier medio, ya sea electrónico, mecánico, por fotocopia, por registro y otros medios, sin el permiso previo de los titulares del *copyright*.

© CTO EDITORIAL, S.L. 2017

Diseño y maquetación: CTO Editorial

C/ Albarracín, 34; 28037 Madrid  
Tfno.: (0034) 91 782 43 30 - Fax: (0034) 91 782 43 43  
E-mail: [ctoeditorial@ctomedicina.com](mailto:ctoeditorial@ctomedicina.com)  
Página Web: [www.grupocto.es](http://www.grupocto.es)

ISBN Estadística y epidemiología: 978-84-17095-81-9  
ISBN Obra completa: 978-84-17095-72-7  
Depósito legal: M-29877-2017



# Estadística y epidemiología

## ENARM - México

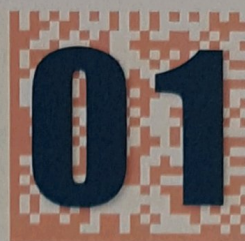
### Manual CTO

### de Medicina y Cirugía

3.<sup>a</sup> edición

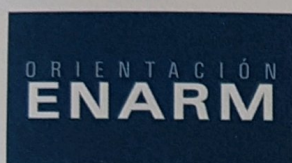
<b>01. Estudio de un test.</b>		
Parámetros de uso.....	<b>1</b>	
1.1. Estudio de un test.....	1	
1.2. Relación entre prevalencia y valores predictivos.....	2	
1.3. Aceptabilidad de un método diagnóstico.....	3	
1.4. <i>Screening</i> .....	3	
1.5. Coeficiente Kappa.....	3	
<b>02. Conceptos y uso de la epidemiología.....</b>	<b>5</b>	
2.1. Concepto de epidemiología.....	5	
2.2. Indicadores de riesgo. Causas de enfermedad.....	5	
<b>03. Medidas de frecuencia de la enfermedad.....</b>	<b>7</b>	
3.1. Formas básicas de medida.....	7	
3.2. Medidas de frecuencia de la enfermedad.....	7	
<b>04. Medidas de asociación.....</b>	<b>9</b>	
4.1. Riesgo relativo (RR).....	9	
4.2. <i>Odds ratio</i> (razón de desventaja) (OR).....	9	
4.3. Razón de prevalencia.....	10	
<b>05. Medidas de impacto o del efecto.....</b>	<b>12</b>	
5.1. Diferencia de incidencias o riesgo atribuible (RA).....	12	
5.2. Fracción atribuible en expuestos (FAE) o fracción etiológica del riesgo.....	12	
5.3. Reducción absoluta del riesgo (RAR).....	12	
5.4. Reducción relativa del riesgo (RRR).....	12	
5.5. Número de sujetos necesario para tratar.....	12	
<b>06. Tipos de estudios epidemiológicos.....</b>	<b>14</b>	
6.1. Estudios descriptivos.....	14	
6.2. Estudios analíticos.....	15	
<b>07. Validez y fiabilidad de los estudios epidemiológicos.....</b>	<b>18</b>	
7.1. Tipos de error.....	18	
7.2. Validez y fiabilidad.....	19	
7.3. Validez de un test diagnóstico.....	20	
<b>08. Ensayo clínico.....</b>	<b>21</b>	
8.1. Introducción.....	21	
8.2. Etapas en la realización de un ensayo clínico.....	21	
8.3. Tipos de ensayos clínicos.....	23	
8.4. Estudios de bioequivalencia.....	24	
8.5. Ensayo clínico de tamaño único.....	24	
8.6. Recomendaciones para la publicación de ensayos clínicos (recomendaciones CONSORT).....	25	
<b>09. Niveles de calidad de la evidencia científica.....</b>	<b>28</b>	
9.1. Metanálisis.....	28	
9.2. Factores determinantes de la calidad de la evidencia científica.....	28	
9.3. Escalas de gradación de la evidencia científica.....	28	

<b>10. Estadística. Generalidades</b> .....	<b>30</b>	<b>14. Estimación de parámetros. Estadística inferencial</b> .....	<b>39</b>
10.1. Tipos de estadística.....	30	14.1. Estimación de medias.....	39
10.2. Población y muestra.....	30	14.2. Estimación de un porcentaje.....	40
10.3. Variables.....	30	14.3. Estimación de medidas de asociación.....	40
<b>11. Estadística descriptiva</b> .....	<b>32</b>	<b>15. Contraste de hipótesis</b> .....	<b>41</b>
11.1. Representación gráfica.....	32	15.1. Hipótesis nula, hipótesis alternativa y grado de significación estadística.....	41
11.2. Síntesis de los datos.....	33	15.2. Errores $\alpha$ y $\beta$ .....	41
 		15.3. Pruebas de significación estadística.....	42
<b>12. Probabilidad</b> .....	<b>35</b>	 	
12.1. Probabilidad.....	35	<b>16. Análisis de correlación y de regresión</b> .....	<b>45</b>
12.2. Probabilidad condicionada.....	35	16.1. Análisis de correlación y de regresión.....	45
12.3. Regla de la multiplicación.....	35		
12.4. Regla de la suma.....	36	 	
 		<b>17. Tamaño muestral</b> .....	<b>47</b>
<b>13. Principales leyes de distribución de variables aleatorias</b> .....	<b>37</b>	17.1. Tamaño muestral.....	47
13.1. Principales leyes de distribución de variables aleatorias.....	37		
13.2. Distribución binomial.....	37	 	
13.3. Distribución de Poisson.....	37	<b>Bibliografía</b> .....	<b>49</b>
13.4. Distribución normal o de Gauss.....	37		



Haz clic en la imagen para abrir tu plataforma

## Estudio de un test. Parámetros de uso



*Este es un tema sencillo.  
Hay que conocer los diferentes  
conceptos, así como resolver  
las tablas.*

### 1.1. Estudio de un test

La actividad epidemiológica estudia la frecuencia de enfermedad. Sin embargo, todas sus medidas son realmente de la frecuencia de diagnósticos de enfermedad, de ahí la importancia de conocer la auténtica correspondencia entre el diagnóstico y la realidad patológica. Muy pocas pruebas diagnósticas, quizá ninguna, identifican con certeza si el paciente padece o no la enfermedad.

La eficacia de una prueba diagnóstica depende de su capacidad para detectar correctamente la presencia o la ausencia de la enfermedad que se estudia, lo que se expresa matemáticamente mediante los cuatro índices siguientes: sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo. Los dos primeros marcan lo que se denomina validez interna del test y los dos segundos, la validez externa; esto es, la que se obtiene al aplicar esa prueba en un entorno poblacional determinado.

Los citados índices se obtienen a partir del análisis de una serie de pacientes a los que se les realiza una prueba diagnóstica (prueba problema), comparándose los resultados con los de una prueba de superior rendimiento diagnóstico (prueba de referencia). Los resultados obtenidos se expresan de la siguiente manera (**Tabla 1**):

- **Sensibilidad (S):** se define como la probabilidad de que un individuo enfermo tenga un test positivo. La sensibilidad indica la proporción del total de enfermos que el test es capaz de detectar:

$$S = \frac{\text{Individuos enfermos con test (+)}}{\text{Total enfermos}} = \frac{VP}{VP + FN} = \frac{a}{a + c}$$

- **Especificidad (E):** es la probabilidad de que un individuo sano tenga un test negativo. La especificidad indica la proporción de individuos sanos confirmados como tales por el resultado negativo del test.

	Enfermo	Sano	
Positivo	VP a	FP b	Total de positivos a + b
Negativo	FN c	VN d	Total de negativos c + d
	Total de enfermos a + c	Total de sanos b + d	Total de individuos a + b + c + d

Verdaderos positivos: resultados positivos en sujetos enfermos. Verdaderos negativos: resultados negativos en sujetos sanos  
Falsos positivos: resultados positivos en sujetos sanos. Falsos negativos: resultados negativos en sujetos enfermos

**Tabla 1.** Resultado de la prueba diagnóstica evaluada

$$E = \frac{\text{Individuos sanos con test (-)}}{\text{Total sanos}} = \frac{VN}{VN + FP} = \frac{d}{b + d}$$

- **Tasa de falsos negativos (TFN):** es la probabilidad de que un individuo, estando enfermo, sea clasificado como sano.

$$TFN = \frac{\text{Individuos enfermos con test (-)}}{\text{Total enfermos}} = \frac{FN}{VP + FN} = \frac{c}{a + c}$$

- **Tasa de falsos positivos (TFP):** mide la probabilidad de que a un individuo sano se le clasifique como enfermo.

$$TFP = \frac{\text{Individuos sanos con test (+)}}{\text{Total sanos}} = \frac{FP}{VN + FP} = \frac{b}{b + d}$$

- **Razón de probabilidad positiva:** compara la probabilidad de que un paciente enfermo presente un resultado positivo frente a la probabilidad de que el resultado positivo se presente en un individuo sano.

$$RP (+) = \frac{\text{Probabilidad test (+) en enfermos}}{\text{Probabilidad test (+) en sanos}} = \frac{VP/\text{enfermos}}{FP/\text{sanos}}$$

$$RP+ = \frac{S}{1 - E}$$

- **Razón de probabilidad negativa:** relaciona la probabilidad de que un paciente enfermo presente un resultado negativo con la probabilidad de que el resultado negativo se dé en un individuo sano.

$$RP (-) = \frac{\text{Probabilidad test (-) en enfermos}}{\text{Probabilidad test (-) en sanos}} = \frac{FN/enfermos}{VN/sanos}$$

$$RP (-) = \frac{1 - S}{E}$$

- **Valor predictivo positivo:** se trata de la proporción de verdaderos positivos entre aquellos que han sido identificados como positivos en una prueba de test.

$$VPP = \frac{VP}{VP + FP} = \frac{a}{a + b}$$

- **Valor predictivo negativo:** se trata de la proporción de verdaderos negativos entre aquellos que han sido identificados como negativos en un test.

$$VPN = \frac{VN}{VN + FN} = \frac{d}{c + d}$$

- **Valor global (eficiencia) del test:** indica la proporción de resultados válidos entre el conjunto de resultados.

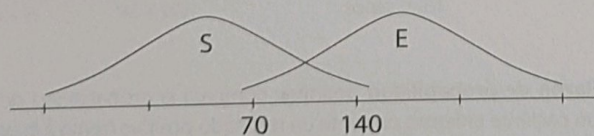
$$VGT = \frac{VP + VN}{VP + VN + FP + FN} = \frac{a + d}{\text{Total}}$$

- **Sensibilidad + tasa de falsos negativos = 100%.**
- **Especificidad + tasa de falsos positivos = 100%.**

La sensibilidad y la especificidad clasifican a los sujetos en función del estatus de enfermedad. Los valores predictivos lo hacen en función del resultado del test.



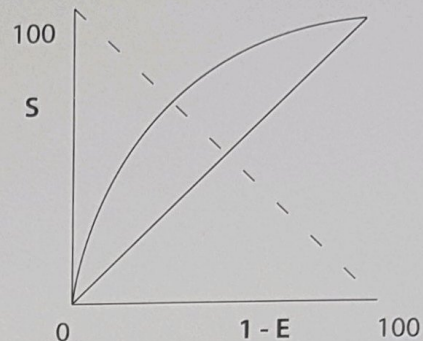
El resultado de un test puede ser continuo (por ej.: niveles de glucemia en mg/dl) y entonces hay que decidir cuál se considerará como resultado positivo del mismo, por lo que es necesario elegir un punto de corte.



El punto de corte escogido determinará la sensibilidad y la especificidad de la prueba (si elegimos 70, la prueba será muy sensible y poco específica; si se selecciona 140, será poco sensible y muy específica).

Para determinar el punto de corte se pueden utilizar las curvas de Características Operativas para el Receptor (COR). Se seleccionan varios puntos de corte y se estima la sensibilidad y la especificidad para cada uno de

ellos. Posteriormente, se representa gráficamente la sensibilidad en función de (1-Especificidad). La prueba ideal se sitúa en el ángulo superior izquierdo (S y E = 1). En general, para hacer la elección del mejor punto de corte se suele emplear lo que se denomina área bajo la curva COR (o ROC en su terminología inglesa), a mayor área bajo la curva, mejor capacidad de discriminación de esa prueba.



Una prueba sin ningún valor sigue la diagonal que va del ángulo inferior izquierdo al ángulo superior derecho (cada incremento de la sensibilidad se asocia a una pérdida de igual magnitud de la especificidad).

De esta forma, la sensibilidad y la especificidad son valores interdependientes, de manera que si aumenta la sensibilidad, disminuye la especificidad y viceversa.

Si se adoptan criterios de diagnóstico muy estrictos, decrece la sensibilidad (hay menos enfermos que cumplen estos criterios), y paralelamente se incrementa la especificidad (pocos sanos cumplen los citados criterios).

## 1.2. Relación entre prevalencia y valores predictivos

Los valores predictivos de un test son variables, dependen de la prevalencia de la enfermedad en la población. La sensibilidad y la especificidad son características propias del test y no se modifican con cambios en la prevalencia.

Si la prevalencia de la enfermedad aumenta, se incrementa el valor predictivo positivo, mientras que disminuye el valor predictivo negativo.

Esto es fácil de entender, por ejemplo, en el caso del VPP, si en una población la prevalencia de una enfermedad es elevada, tendrá mayor probabilidad de hacer test a individuos realmente enfermos, lo que disminuirá la tasa de falsos positivos y, por ello, aumentará el VPP.

Por el contrario, si la prevalencia de una enfermedad es muy pequeña, se puede tener una alta probabilidad de realizar test que den resultados falsos positivos, porque en la mayor parte el test se estará aplicando sobre individuos sanos, incluso con cifras de buena sensibilidad del test, el VPP disminuye, en ocasiones notablemente. Por eso es tan importante conocer no sólo los valores de sensibilidad y especificidad sino el ámbito en el que se aplican para entender los conceptos de VPP y VPN.

Por el contrario, si la prevalencia de la enfermedad decrece, aumenta el valor predictivo negativo y se reduce el valor predictivo positivo.



- ↑ PREVALENCIA: ↑ VPP Y ↓ VPN
- ↓ PREVALENCIA: ↓ VPP Y ↑ VPN



Recuerda

### 1.3. Aceptabilidad de un método diagnóstico

No existe un parámetro guía útil para todas las situaciones. La aceptabilidad de un test depende de la patología estudiada y de las condiciones reales en el medio y en la colectividad.

Si lo que interesa es detectar el mayor número posible de enfermos, se debe usar un test con alta sensibilidad. Así se escapan pocos enfermos, aunque con un coste de obtención de bastantes "falsos positivos".

Se elegirá un test sensible cuando:

- La enfermedad sea grave y no pueda pasar desapercibida.
- La enfermedad sea tratable.
- Los resultados falsamente positivos no supongan un trauma psicológico en los individuos examinados.

Como un test sensible da muchos falsos positivos, la rentabilidad del mismo es que su resultado sea negativo.



Recuerda

Si lo que se quiere es "asegurar" el diagnóstico, se ha de emplear un test cuya especificidad sea máxima. Se utilizará un test lo más específico posible cuando:

- La enfermedad sea importante, pero difícil de curar o incurable.
- Los resultados falsamente positivos puedan suponer un trauma psicológico para el individuo examinado.
- El tratamiento de los falsos positivos pudiera tener graves consecuencias.

### 1.4. Screening

El *screening* es una estrategia de detección precoz de la enfermedad. Lo ideal es aplicar primero un test muy sensible (que detecta todos los casos posibles de enfermedad; por tanto se obtienen muchos FP y pocos FN) y en una segunda fase, aplicar un test muy específico (con el que se confirma el diagnóstico de esos posibles enfermos; se obtienen muy pocos FP).

En la puesta en marcha de un programa de *screening* deben tenerse en cuenta los criterios dependientes de la enfermedad, del test y de la población diana que se detallan a continuación.

### Ideas clave

- La sensibilidad y la especificidad se refieren respectivamente a la capacidad del test para detectar correctamente a los individuos enfermos y a los sanos. La sensibilidad y la especificidad no dependen de la prevalencia.

### Criterios dependientes de la enfermedad

Los criterios dependientes de la enfermedad son los siguientes:

- La enfermedad debe ser común y grave.
- Ha de conocerse la historia natural de la enfermedad.
- El tratamiento, en el estado presintomático, tiene que reducir la morbimortalidad en mayor medida que el tratamiento después de la aparición de los síntomas.

### Criterios dependientes del test

Los criterios dependientes del test son:

- De fácil aplicación.
- Coste razonable dentro del presupuesto de salud.
- Inocuo para la población.
- Confiabilidad o capacidad de repetición.
- **Validez:** se refiere a la capacidad del test de medir lo que realmente se desea medir. El concepto de validez incluye los de sensibilidad, especificidad y valor predictivo; ya se han mencionado previamente los conceptos de validez interna y externa.

### Criterios dependientes de la población diana

Estos criterios se detallan a continuación:

- El riesgo de ser afectado por la enfermedad debe ser alto.
- La información demográfica tiene que estar disponible en la comunidad.
- La comunidad ha de sentir la necesidad de programas de salud pública.

### 1.5. Coeficiente Kappa

El **coeficiente Kappa** es la medida del grado de concordancia no aleatoria entre varios observadores o entre distintas mediciones de la misma variable. Varía de -1 a +1.

- $k = -1$ . Discrepan más mediciones de lo esperado por azar.
- $k = 0$ . Concordancia debida al azar.
- $k = +1$ . Concordancia completa.

**Ejemplo:** dos oftalmólogos revisaron la misma serie de 100 fotografías de fondo de ojo para determinar el grado de retinopatía. Se desea medir el grado de coincidencia, concordancia entre los dos observadores.

El índice Kappa sirve para medir la concordancia cuando la variable de estudio es cuantitativa. Sin embargo, si de lo que se trata es de comparar concordancias, el índice más apropiado para cuantificar la concordancia entre diferentes mediciones de una variable numérica es el llamado **coeficiente de correlación intraclass**.

- Los valores predictivos dependen de la prevalencia. El positivo se refiere a la probabilidad de ser enfermo cuando el test es positivo y el VPN, a la probabilidad de estar sano cuando el test ha sido negativo.
- Los test sensibles sirven para descartar enfermedad. Los específicos, para confirmar enfermedad.

➤ Las curvas COR representan la relación recíproca entre sensibilidad y especificidad. Cuanto mayor es el área bajo la curva COR, mejor será el test.

➤ La prevalencia va a determinar el tipo de test más útil en una comunidad. Si la prevalencia es alta, se realizará un test sensible. Si es baja, uno específico.

## Casos clínicos

La probabilidad de padecer estenosis coronaria en hombres de más de 65 años, con angina de pecho de esfuerzo típica, es mayor del 90%. A un paciente de estas características se le practica un ECG de esfuerzo que resulta negativo. La sensibilidad es del 80% y su especificidad del 90%. ¿Cuál de las siguientes afirmaciones es correcta?

- 1) Se puede descartar por completo estenosis coronaria porque la prueba es negativa.
- 2) No se puede excluir la estenosis coronaria porque el valor predictivo positivo de la prueba es alto.
- 3) No se puede excluir la estenosis coronaria porque el valor predictivo negativo de la prueba es bajo.
- 4) Se podría descartar el diagnóstico de estenosis coronaria si la especificidad de la prueba fuera del 100%.

RC: 3

## Case Study

You are comparing the clinical diagnosis of gastroduodenal ulcer and its prevalence in a series of 10,000 autopsies. When you begin to analyze the data you plot the following table:

	Ulcer	Non-ulcer
Ulcer	130	20
Non-ulcer	170	9.680

Which is the closest to the sensitivity of the clinical diagnosis of ulcer?

- 1) 3%.
- 2) 43%.
- 3) 87%.
- 4) 100%.

Correct answer: 2



Haz clic en la imagen para abrir tu plataforma

## Conceptos y uso de la epidemiología



*Se deben reconocer los criterios de causalidad; en especial, la demostración experimental.*

### 2.1. Concepto de epidemiología

Etimológicamente, la palabra epidemiología procede del griego y significa tratado sobre el pueblo (*epi*: sobre, *demo*: pueblo, *logos*: tratado).

Se puede definir la **epidemiología** como la ciencia que estudia la distribución y los determinantes del fenómeno salud-enfermedad en las poblaciones humanas. La comunidad sustituye al individuo a la hora de buscar la respuesta a ciertas preguntas sobre la etiología y la prevención de la enfermedad, así como acerca de los recursos necesarios para los cuidados de salud de esa población.

La epidemiología se ocupa de los dos aspectos fundamentales siguientes:

- Estudiar la distribución de las enfermedades en relación con las variables lugar, tiempo y persona. Es lo que se denomina **epidemiología descriptiva**.
- Buscar los factores que determinan la distribución encontrada e identificar las asociaciones causales. Es lo que hace la llamada **epidemiología analítica**.

Las principales posibilidades de aplicación de la epidemiología son las que se enumeran a continuación:

- Establecer el diagnóstico de salud de una comunidad.
- Conocer los factores causales de una enfermedad y las probabilidades de enfermar.
- Evaluar la eficacia/efectividad/eficiencia de los procedimientos preventivos y terapéuticos de los servicios de salud.

Los pasos fundamentales dentro del método epidemiológico son los siguientes:

1. Observar un fenómeno.
2. Elaborar una hipótesis.
3. Probar la hipótesis.
4. Emitir un informe o ley.

### 2.2. Indicadores de riesgo. Causas de enfermedad

#### Conceptos

- **Riesgo**: es la probabilidad de desarrollar una enfermedad por estar expuesto a ciertos factores.
- **Causa**: es la condición que, sola o acompañada de otras condiciones, inicia o permite iniciar una secuencia de acontecimientos que producen un determinado efecto.
- **Factor de riesgo**: variable endógena o exógena al individuo, controlable, que precede al comienzo de la enfermedad y que está asociada a un incremento de la probabilidad de aparición de la misma, y a la que se puede atribuir "responsabilidad" en la enfermedad (por ejemplo: tabaco como factor de riesgo de cáncer de pulmón).
- **Marcador de riesgo**: variable no controlable, endógena al individuo (constitucional), y cuya presencia anuncia a los individuos particularmente vulnerables (por ej.: sexo femenino como marcador de riesgo de cáncer de mama).
- **Indicador de riesgo**: variable sin relación causal con el problema, pero cuya presencia alerta precozmente sobre el padecimiento. (por ejemplo: manchas de Koplik como signo precursor de la aparición del sarampión).

#### Criterios epidemiológicos de causalidad

La existencia de asociación estadística no es sinónimo de causalidad, sino que para que se dé deben cumplirse los denominados criterios de causalidad, que son los que se definen a continuación:

- **Fuerza de asociación**: es un criterio muy importante, consistente en cuántas veces más riesgo de enfermedad tienen las personas expuestas al factor estudiado que las no expuestas (véase el Capítulo 4).
- **Efecto dosis-respuesta**: es decir, que la asociación tenga coherencia interna, que al incrementarse la presencia del factor de riesgo aumente la de la enfermedad.
- **Secuencia temporal**: la exposición al presunto factor de riesgo debe preceder al efecto.
- **Coherencia externa o consistencia de los resultados**: los resultados deben ser concordantes con los obtenidos en otros estudios (reproductibilidad).

- **Ausencia de distorsiones metodológicas en el estudio:** debidas a sesgos.
- **Ausencia de explicaciones alternativas:** es decir, ausencia de otras hipótesis que expliquen los resultados.
- **Plausibilidad biológica:** que los resultados sean compatibles con el marco de los conocimientos científicos previos.
- **Efecto de la cesación o reversibilidad:** es decir, que exista una disminución del riesgo tras la reducción de la exposición al presunto factor de riesgo.
- **Demostración experimental:** es la prueba causal por excelencia, pero su realización plantea en ocasiones problemas éticos.

## Modelos causales

Los modelos causales son los siguientes:

- **Determinista o unicausal:** este modelo dice que siempre que se presente la causa (causa suficiente) y sólo cuando se presente la cau-

sa (causa necesaria), ocurre el efecto. Hay especificidad de causa (el efecto tiene una única causa) y también de efecto (cada causa produce un único efecto).

- **Multicausal:** postula que hay pluralidad de causas y multiplicidad de efectos, formando redes causales, de forma que diferentes causas producen el mismo efecto y una única causa ocasiona varios efectos.
- **Determinista modificado (Rothman):** en este modelo se distinguen tres tipos de causas:
  - **Causa suficiente:** es aquella que inevitablemente produce el efecto cuando se presenta.
  - **Causa complementaria o contribuyente:** son aquellas causas que forman parte de una causa suficiente.
  - **Causa necesaria:** es la que debe estar presente inevitablemente para que se produzca el efecto. Por tanto, ha de formar parte de todas las causas suficientes de una enfermedad.

### Ideas clave

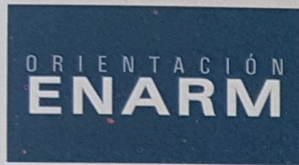
- La demostración experimental supone la mejor evidencia de causalidad.

- El principal inconveniente de la demostración experimental son los problemas éticos.



## Medidas de frecuencia de la enfermedad

Haz clic en la imagen para abrir tu plataforma



*En este tema es importante conocer los conceptos de incidencia, de prevalencia y de densidad de incidencia.*

Su valor es más alto a mayor incidencia y duración de la enfermedad.

$$Pv = \frac{\text{N.º total de casos de enfermedad}}{\text{Población en ese momento}}$$

### 3.1. Formas básicas de medida

Las formas básicas de medida son las siguientes:

- **Número:** es un valor absoluto. Al no conocer el denominador no se puede saber la importancia relativa del problema.
- **Razón:** es el cociente en el que el numerador no está incluido en el denominador, es decir, cociente entre dos entidades que poseen caracteres distintos. Sus valores oscilan entre 0 e infinito.
  - Fórmula:  $R = a/b$
  - Ejemplo: razón de hombres/mujeres.
- **Proporción:** es el cociente en el cual el numerador está incluido en el denominador. Sus valores oscilan entre 0 y 1. Es adimensional.
  - Fórmula:  $P = a/a+b$
  - Ejemplo: proporción de aprobados en un examen.
- **Tasa:** mide la ocurrencia de un suceso durante un periodo de tiempo en una población determinada. Es un cociente en el que el numerador está incluido en el denominador pero donde, a diferencia de la proporción, el tiempo también está incluido en éste. Tiene, por tanto, dimensión temporal.
  - Fórmula:  $T = a/\text{persona} \times \text{tiempo}$ .
  - Ejemplo: 300 ca. pulmón/100.000 personas-año.

### 3.2. Medidas de frecuencia de la enfermedad

Las medidas de frecuencia de la enfermedad se llevan a cabo con los siguientes indicadores:

- **Prevalencia:** es el número total de casos entre el total de la población. No hace distinción entre casos antiguos y nuevos y nos describe la situación en un momento determinado del tiempo.

- **Incidencia acumulada:** se define como el número de casos nuevos de una enfermedad que se producen en un periodo de tiempo. Es la probabilidad de que un individuo desarrolle la enfermedad en ese periodo.

$$IA = \frac{\text{N.º casos nuevos en un periodo de tiempo}}{\text{Población en riesgo al inicio de ese periodo}}$$

La incidencia acumulada se puede relacionar con la prevalencia según las siguientes fórmulas:

- **$Pv < 10\%$ :**

$$Pv = I \times \text{duración media de la enfermedad}$$

- **$Pv > 10\%$ :**

$$Pv = \frac{I \times D}{1 + (I \times D)}$$

- **Densidad de incidencia:** se define como el número de casos nuevos de la enfermedad que se producen en un periodo de tiempo, teniendo en cuenta la información que aporta el sumatorio de riesgo de cada uno de los individuos.

$$DI = \frac{\text{N.º casos nuevos}}{\text{Personas} \times \text{tiempo}}$$

**Personas x tiempo:** es una unidad de medida de la población en riesgo. Se obtiene sumando la cantidad de tiempo que cada individuo ha estado expuesto al factor de estudio.

**Ejemplo:**


1	1 año
2	4 años
3	5 años
4	2 años
5	8 años
<años>	TOTAL: 20 años

$$DI = \frac{2 \text{ casos}}{20 \text{ personas-año}} = 0,1 \text{ personas-año}$$

La densidad o tasa de incidencia, por tanto, expresa los casos nuevos de enfermedad que se producen en un periodo de tiempo que corresponde al sumatorio de los periodos individuales de riesgo. Es decir, que si se lee que la densidad de incidencia es de 2 por 1.000 personas-año, quiere decir que aparecen de media dos casos por cada 1.000 individuos en un año de exposición, sumando los tiempos de exposición de todos los individuos de la cohorte.

Este concepto se conoce también con el nombre de fuerza de la morbilidad.

Hay que observar que en las preguntas en las que a partir de datos se pide calcular incidencias, en el denominador habrá que restar los enfermos al inicio del periodo, mientras que no habrá que restarlos si lo que se pregunta es la prevalencia.



	Cálculo	Significado
Prevalencia	$\frac{\text{Casos existentes}}{\text{Población total}}$	Proporción de individuos enfermos en una población en un momento concreto
Incidencia acumulada	$\frac{\text{Casos nuevos observados}}{\text{Población susceptible al comienzo del periodo}}$	Riesgo individual de enfermar
Densidad de incidencia	$\frac{\text{Casos nuevos observados}}{\text{Suma de los periodos de riesgo}}$	Velocidad con que determinada enfermedad aparece en una población

**Tabla 1.** Medidas de frecuencia de la enfermedad

**Ideas clave** 

- La incidencia es sinónimo de riesgo, porque se refiere a los nuevos casos debidos a una enfermedad.
- Como incidencia es riesgo, todos los parámetros que incluyen el término riesgo (RR, RAE, RRR, RAR) se calculan utilizando incidencias.
- La prevalencia o probabilidad pretest se refiere a la cantidad de enfermos en una población, no sólo a los casos nuevos.
- Es importante conocer los aspectos de incidencia, prevalencia y densidad de incidencia.



Haz clic en la imagen para abrir tu plataforma

# Medidas de asociación



*Sin duda, lo más importante es el concepto del riesgo relativo.*

Las medidas de asociación sirven para cuantificar cuánto y cómo se asocia un factor de riesgo a una enfermedad, a partir de la frecuencia con la que aparece el evento resultado en un grupo expuesto a un factor de riesgo, frente al grupo que no está expuesto a dicho factor. Es decir, se emplean para decir si un factor analizado es protector o de riesgo y se utilizan para cuantificar la asociación entre factor de riesgo y evento resultado.

## 4.1. Riesgo relativo (RR)

El riesgo relativo es la medida de asociación en los estudios de cohortes. Mide la "fuerza de la asociación" entre el factor de riesgo y la enfermedad. Puede variar entre 0 e infinito.

El riesgo relativo responde a la pregunta: ¿cuánto más frecuente es la enfermedad entre los expuestos a un factor de riesgo respecto a los no expuestos?

Su significado varía dependiendo del valor que tome:

- $RR > 1$ : factor de riesgo (FR).
- $RR = 1$ : indiferente: la incidencia es igual en expuestos y en no expuestos.
- $RR < 1$ : factor de protección.

$$RR = \frac{\text{Incidencia en expuestos}}{\text{Incidencia en no expuestos}}$$

- Incidencia en expuestos:  $I_e = a/a + b$
- Incidencia en no expuestos:  $I_o = c/c + d$

$$\frac{I_e}{I_o} = \frac{a/a + b}{c/c + d}$$

	Enfermos	Sanos	Total
FR	a	b	a + b
No FR	c	d	c + d
Total	a + c	b + d	a + b + c + d

Tabla 1. Cálculo del RR

## 4.2. Odds ratio (razón de desventaja) (OR)

La **odds ratio** es una razón. Constituye la medida básica de los estudios casos-controles. Su significado es idéntico al del RR, aunque no puede calcularse como este, ya que en los estudios casos-controles no es posible valorar la incidencia de la enfermedad. Para que la OR sea un buen estimador del RR es necesario que los controles sean representativos de la población de la cual han surgido los casos y que la enfermedad tenga una prevalencia baja, inferior al 10%.

$$OR = \frac{\text{Odds de exposición en los casos}}{\text{Odds de exposición en los controles}} = \frac{\frac{\text{Casos expuestos}}{\text{Casos no expuestos}}}{\frac{\text{Controles expuestos}}{\text{Controles no expuestos}}}$$

	Casos (enfermos)	Controles (sanos)	Total
FR	a	b	a + b
No FR	c	d	c + d
TOTAL	a + c	b + d	a + b + c + d

Tabla 2. Cálculo de la OR

$$OR = \frac{a/c}{b/d} = \frac{a \times d}{b \times c}$$

## 4.3. Razón de prevalencia

La razón de prevaencia es la medida de asociación de los estudios transversales. Su interpretación es similar a la del riesgo relativo, es decir,

el número de veces más que padecen la enfermedad los expuestos frente a los no expuestos.

$$R. Pv = \frac{\text{Enfermos expuestos}}{\text{Enfermos no expuestos}}$$

### Ideas clave

- El RR es la medida de asociación de los estudios prospectivos, es decir, de cohortes y de los ensayos clínicos.
- Es importante conocer los valores posibles, así como el significado de RR menor, igual o mayor de 1.

- Cuando en una pregunta, además del RR, se proporcione un intervalo de confianza, hay que fijarse en este último. Si el intervalo incluye el 1, entonces se dice que el RR no ha sido significativo.
- La OR es la medida de asociación de los diseños retrospectivos o de casos y controles. Su mayor utilidad es cuando la enfermedad de estudio es poco prevalente.

### Casos clínicos

En un estudio de seguimiento a largo plazo de pacientes con fibrilación auricular crónica en la práctica clínica habitual, se ha documentado un riesgo relativo de padecer una hemorragia intracraneal para anticoagulantes orales (en comparación con antiagregantes plaquetarios) de 3 (intervalos de confianza del 95% 1,5-6). ¿Cuál de las siguientes afirmaciones es cierta?

- 1) Los intervalos de confianza del riesgo relativo estimado confirman que las diferencias no alcanzan la significación estadística.
- 2) El riesgo relativo estimado significa que la incidencia de hemorragia intracraneal es un 3% superior en el grupo tratado con anticoagulantes orales.
- 3) El riesgo relativo estimado significa que existe una alta proporción (superior al 3%) de pacientes tratados con anticoagulantes orales que presentaron hemorragia intracraneal.
- 4) El riesgo relativo estimado significa que la incidencia de hemorragia intracraneal entre los pacientes tratados con anticoagulantes orales ha sido tres veces superior a la de los tratados con antiagregantes plaquetarios.

RC: 4

Para determinar si la ingesta total de líquidos reduce el riesgo de padecer cáncer de vejiga urinaria, se estudió durante 10 años la frecuencia de este cáncer en un grupo de 47.909 hombres que no lo tenían al inicio del estudio. De todos ellos se había obtenido información sobre el consumo de líquidos, y se comparó la frecuencia de enfermedad según el volumen de líquidos consumidos, agrupado por quintiles. Entre los resultados, destacaba que la ingesta total de líquidos se asoció con el cáncer de vejiga, de forma que para el quintil superior (ingestión > 2.531 ml de líquidos por día) se calculó un riesgo relativo de 0,51 (intervalo de confianza al 95%: 0,32 a 0,80) al compararlo con el quintil inferior (< 1.290 ml por día). Señale la respuesta correcta:

- 1) Los autores han realizado un estudio de casos y controles.
- 2) El valor de "p", obtenido a partir del estadístico de la comparación de las medidas de frecuencia de cáncer de vejiga entre los dos grupos de ingesta citados, debe ser mayor de 0,005.
- 3) Se ha detectado un menor riesgo de padecer cáncer de vejiga en las personas que consumían mayor volumen de líquidos por día.
- 4) El intervalo de confianza nos da una idea del grado de precisión estadística del estudio, de forma que cuanto más amplio, menos error aleatorio.

RC: 3

### Case Study

You are reading a cohort study, which evaluates the effect of nonsteroidal anti-inflammatory drugs (NSAIDs) on decompensating heart failure (HF). The relative risk (RR) is 1.25 (0.7-2.2). Which of the following is true:

- 1) The use of NSAIDs causes decompensating heart failure, because the RR confidence interval does not attain 0.
- 2) The relationship between the use of NSAIDs and the CI appears to be accounted for by causality, but before asserting that the relationship is causal, it should be carefully evaluated whether the study has bias or confounding factors.

- 3) Cohort studies are unable to establish causal relationships between variables.
- 4) We cannot say that causality is responsible for the relationship between the two variables, as the RR confidence interval includes the value 1.

Correct answer: 4

We performed a retrospective cohort study to determine whether patients taking antipsychotics have an increased risk of sudden death than the population without antipsychotics. Once the adjustment for potential confounders was made we obtained a relative risk of 2.39 (confidence interval 1.77 to 3.22

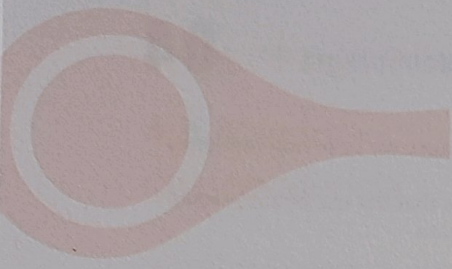


with a 95% confidence). What is the correct interpretation of the result?

- 1) The result is consistent with an increased risk associated with the use of antipsychotics, but not statistically significant.
- 2) The result suggests that antipsychotics protect against the risk of sudden death.

- 3) The result is not interpretable because we have not randomly allocated treatments.
- 4) The result supports the hypothesis that the use of antipsychotics increases the risk of sudden death.

Correct answer: 4



# 05

# Medidas de impacto o del efecto



Los conceptos más importantes son el NNT, la RAR y la RRR.

### 5.1. Diferencia de incidencias o riesgo atribuible (RA)

Es una medida que informa sobre el exceso de riesgo en los individuos expuestos frente a los no expuestos al factor de riesgo. Indica la incidencia acumulada en el grupo de expuestos que se debe exclusivamente al factor de riesgo. Representa el descenso en el número de casos nuevos entre los expuestos si se evitara el FR.

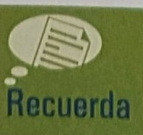
$$RA = I_e - I_o$$

### 5.2. Fracción atribuible en expuestos (FAE) o fracción etiológica del riesgo

La fracción atribuible en expuestos es la proporción de la enfermedad que se debe a la exposición; es decir, la proporción de casos nuevos de enfermedad, entre los expuestos, que se evitaría si se eliminase el factor de riesgo. Por tanto, estas dos medidas son muy importantes a la hora de evaluar, en términos de salud pública, cuál puede ser el impacto con determinadas medidas, calcular los costes y determinar prioridades.

$$FAE = \frac{I_e - I_o}{I_e}$$

Mientras que el riesgo atribuible mide cuántos casos se deben exclusivamente a un factor de riesgo, la fracción atribuible hace referencia al porcentaje de enfermedad atribuible directamente a dicho factor.



### 5.3. Reducción absoluta del riesgo (RAR)

Cuando se evalúa una medida preventiva, como un factor protector o el resultado de un fármaco frente a placebo en un ensayo clínico, el beneficio de dicha medida se cuantifica como una reducción del número de enfermos nuevos. Este concepto está representado por la reducción absoluta del riesgo (RAR).

$$RAR = I_o - I_e$$

### 5.4. Reducción relativa del riesgo (RRR)

De nuevo se emplea para calcular el beneficio de una medida preventiva.

En este caso, se expresa como porcentaje de casos evitados por haber sido expuestos a dicho factor protector:

$$RRR = \frac{I_o - I_e}{I_o}$$

### 5.5. Número de sujetos necesario para tratar

A partir de los parámetros anteriores, puede calcularse el número de sujetos que deberían exponerse al factor preventivo o terapéutico para conseguir una curación.

$$NNT = \frac{100}{I_o - I_e}$$

## Ideas clave

- El NNT siempre se pregunta de la misma forma: "¿a cuántos pacientes se debería tratar para conseguir un evento?".
- El NNT es tanto mejor cuanto menor sea.

- La RRR y la RAR sirven para cuantificar la relevancia clínica de una medida terapéutica. Su cálculo se realiza mediante una resta de incidencias.
- El RA y la FAE cuantifican cuánto de la enfermedad se debe a un factor de riesgo, en el grupo de sujetos expuestos a dicho factor.

## Casos clínicos

En un ensayo clínico controlado, aleatorizado y enmascarado a doble ciego en pacientes con hipertensión arterial, se comparó el riesgo de aparición de accidente vascular cerebral (AVC) en un grupo tratado con clortalidona y en otro tratado con placebo. A los cinco años, en el grupo tratado con clortalidona, al compararlo con el grupo placebo, se constató una reducción del riesgo de AVC del 6% al 4%, una reducción relativa del riesgo del 34% y un riesgo

relativo de 0,66. Suponiendo que estos datos se pudieran extrapolar a la práctica habitual, ¿a cuántos pacientes y durante cuánto tiempo deberíamos tratar con clortalidona para evitar un AVC?

- 1) A 50 pacientes durante un año.
- 2) A 34 pacientes durante cinco años.
- 3) A 50 pacientes durante cinco años.
- 4) A 34 pacientes durante cinco años.

RC: 3

# Tipos de estudios epidemiológicos

ORIENTACIÓN  
**ENARM**

*Se encuentran dos tipos de cuestiones a tener en cuenta y que son características. Por una parte, las que se refieren a temas concretos o aspectos teóricos de los diferentes tipos de estudios. Por otra, un tipo de cuestiones más "clínico": a partir de un estudio se pide que se indique de qué tipo de estudio se trata.*

## 6.1. Estudios descriptivos

Los objetivos de los estudios descriptivos son los siguientes:

- Describir las características y la frecuencia de un problema de salud, en función de las características de la **persona** (edad, sexo, estado civil...), del **lugar** (área geográfica...) y del **tiempo** de aparición del problema y su tendencia.
- Servir de base para estudios analíticos.

Los tipos de estudios descriptivos son los que se citan a continuación:

- **Series de casos clínicos:** describen las características de un grupo de enfermos. Son estudios longitudinales ya que contienen información adquirida a lo largo del tiempo. Su principal ventaja es que permiten generar nuevas hipótesis, mientras que el mayor inconveniente es que no presentan grupo control, por lo que cualquier FR puede ser un hallazgo casual.
- **Estudios ecológicos:** los conglomerados pueden estar constituidos por grupos poblacionales, comunidades, regiones, o países. La característica principal de este tipo de estudios es que se cuenta con información sobre la exposición o el evento para el conglomerado en su totalidad, desconociéndose la información a nivel individual para cada uno de los miembros del grupo. Son estudios en los que la unidad de análisis son grupos de individuos, no individuos (por ej: clases de una escuela, ciudades, regiones).

Son útiles cuando no se pueden hacer mediciones de exposición individuales (contaminación del aire, ruidos, etc.). En este tipo de estudios es común asignar la misma exposición (exposición promedio) a todo el conglomerado, no considerando la variación individual de cada miembro del mismo. Igual sucede con la medición del evento; dado que sólo se cuenta con el número de eventos registrados para el conglomerado, no es posible discernir los eventos que se presentaron en los sujetos expuestos de los que ocurrieron en los no expuestos, por lo que es frecuente atribuir la totalidad de eventos a la exposición promedio que se registró en el interior del conglomerado.

Los estudios ecológicos permiten estudiar grandes grupos poblacionales en poco tiempo y con un coste relativamente muy bajo ya que, en general, utilizan estadísticas existentes recolectadas con otros fines. Sin embargo, dado que son uno de los peores estudios en la escala de causalidad, deben ser considerados únicamente para sugerir hipótesis, que tendrán necesariamente que ser verificadas mediante otros estudios más rigurosos de base individual, realizados posteriormente.

- **Ventajas:** permiten describir diferencias en poblaciones que habrán de ser estudiadas con más detalle posteriormente.
- **Limitaciones:** los datos son promedios de poblaciones. Se usan medidas aproximadas de exposición (impuestos por alcohol, ventas de cajetillas de cigarrillos...) y de enfermedad (mortalidad en vez de incidencia...), lo que limita el valor de los hallazgos. Los principales problemas de este tipo de estudios son que se ignora la variabilidad individual de los integrantes de los conglomerados y que no es posible corregir por diferencias en otras variables que pudieran estar también asociadas con la exposición y el evento en estudio.

- **Estudios transversales o de prevalencia:** son estudios descriptivos y transversales, ya que analizan la relación entre una enfermedad y algunas variables en un momento concreto del tiempo. Buscan hallar una posible relación entre un FR y una enfermedad, que luego habrá de ser verificada por estudios analíticos.

- **Características:** es de "corte" o transversal, ya que enfermedad y características se miden simultáneamente.
- **Ventajas:** no tienen problemas éticos, son de duración mínima, su coste es bajo y son de fácil reproducibilidad. Son útiles para el estudio de enfermedades crónicas en la población.
- **Inconvenientes:** no es útil para estudiar enfermedades raras, ya que no permite ver el mecanismo de producción de la enfermedad y no sirve para comprobar una hipótesis previa de causalidad (es imposible conocer la secuencia temporal porque la información sobre el factor de riesgo y la enfermedad se recogen a la vez).

- **Análisis de las medidas de enfermedad:** con el estudio de prevalencia, la medida que se obtiene es la razón de prevalencia de individuos expuestos.

## 6.2. Estudios analíticos

Los estudios analíticos intentan establecer una relación de causalidad entre el factor de riesgo y la enfermedad; se pueden clasificar en experimentales y observacionales. En los primeros es el investigador el que asigna el factor de estudio (qué fármaco, vacuna, campaña de educación..., cuánto tiempo, cuándo, cuánta dosis recibirán los individuos...), mientras que en los segundos, el investigador se limita a observar qué es lo que sucede en un grupo de individuos, sin manipular el estudio.

### Estudios analíticos experimentales

Se dice que un estudio es experimental cuando cumple las dos condiciones siguientes:

- Asignación por parte del investigador del factor de estudio.
- Aleatorización de la muestra de modo que los participantes son adscritos al azar a uno u otro grupo de estudio.

### Limitaciones

Los problemas éticos son el principal inconveniente de este tipo de estudios. No es admisible exponer a unos sujetos a un factor de riesgo que presumiblemente es el causante de una enfermedad.

### Ventajas

- Son los estudios que mejor valoran la utilidad de una intervención y aportan mayor evidencia causal.
- Permiten un gran control sobre cualquier efecto que no sea el estudiado.
- Posibilitan el empleo de técnicas de enmascaramiento.

Los tipos de estudios analíticos experimentales son los siguientes:

- **Ensayo clínico aleatorio:** es, con mucho, el estudio experimental más frecuente. La asignación aleatorizada del factor de estudio (un fármaco o una intervención sanitaria) se hace sobre los individuos. Es el mejor para demostrar causalidad y la eficacia de una actuación. Por ejemplo, la asignación aleatorizada de un antirretroviral o un placebo a un grupo de pacientes con VIH (ampliar estudio en el Capítulo 8).

No siempre que se evalúan dos tratamientos se está ante un ensayo clínico. Si los pacientes no fueron asignados por el investigador y este no introdujo el factor de estudio, nos encontramos ante un estudio de cohortes.



- **Ensayo de campo:** es un estudio que valora la eficacia de una medida preventiva. En general, estos estudios son más caros que los ensayos clínicos y requieren mayor número de individuos. Las principales diferencias respecto a los ensayos clínicos son:
  - Se hacen sobre individuos sanos.
  - Valoran la eficacia de las medidas preventivas.

Por ejemplo, la evaluación de la eficacia de una campaña de vacunación en 10.000 niños, de forma que 5.000 reciben aleatoriamente una vacuna y los otros 5.000, un placebo.

### Estudios analíticos cuasiexperimentales

Se diferencian de los estudios experimentales puros en que no hay asignación al azar (aleatorización). Sus tipos son los siguientes:

- **Ensayo comunitario de intervención:** es una variedad de los ensayos de campo:
  - Se trabaja con individuos sanos.
  - Valora la eficacia de medidas preventivas.
  - No se aplica aleatorización individual.

Por ejemplo, una campaña sanitaria de agua fluorada en una comunidad y de agua clorada en la otra, comparando la frecuencia de caries en los dos grupos.

- **Ensayos antes-después:** en este tipo de estudios, el fármaco (o medida en general) se administra a los individuos y se compara el resultado con la situación basal. Los estudios antes-después tienen la ventaja de que son más fáciles de hacer, pero poseen el inconveniente grave de que, al no disponer de grupo de control, los resultados son difíciles de interpretar.

Por ejemplo, a un grupo de pacientes cardiopatas se les administra un IECA, comparando la TA media en el conjunto antes y después de administrar el tratamiento.

- **Estudios controlados no aleatorios:** se realizan cuando la asignación aleatoria no ofrece ventajas o no se puede hacer. Por ejemplo, se quiere hacer un ensayo en pacientes cardiopatas. A un grupo se le administra IECA (se trata de pacientes no dislipémicos) y al otro (que además son dislipémicos), IECA más estatinas. Aquí no hay asignación aleatorizada, sino que la inclusión en uno u otro grupo de estudio se ha hecho tomando como base los factores de riesgo que presentan los pacientes.

### Estudios analíticos observacionales

Los tipos de estudios analíticos observacionales son los que se enumeran a continuación.

#### Estudios de cohortes (Figura 1)

Partiendo de un grupo de individuos expuestos al factor de riesgo (cohorte expuesta), y de otro conjunto comparable en todo pero cuyos individuos no están expuestos al FR (cohorte no expuesta), se estudia la incidencia de la enfermedad en ambas cohortes. Por ejemplo, se sigue a un grupo de fumadores y a otro de no fumadores a lo largo del tiempo, observando el número de casos nuevos de cáncer de pulmón que aparecen en cada grupo de pacientes.

- **Características:**
  - Son estudios longitudinales, de seguimiento.
  - Es prospectivo (excepto en los estudios de cohortes históricas).
  - Va de la causa al efecto (enfermedad).

- **Ventajas:**
  - Es el mejor estudio para comprobar hipótesis previas de causalidad cuando, por razones éticas, no es posible realizar un estudio experimental.
  - Es el mejor para el estudio de la "multiefectividad del factor de riesgo" (todos los efectos del factor de riesgo).

- La posibilidad de sesgos es baja.
- Sirve para el estudio de exposiciones raras.

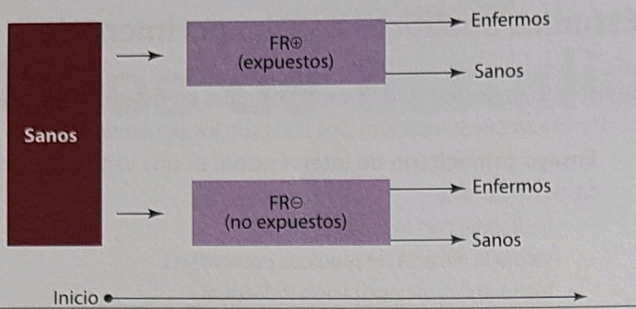


Figura 1. Estudio de cohortes

**Inconvenientes:**

- No es bueno para el estudio de enfermedades raras.
- No es adecuado para el estudio de enfermedades de largo periodo de incubación.
- El coste es alto.
- No sirve para el estudio de la "multicausalidad de la enfermedad".
- Es difícilmente reproducible.

**Análisis de las medidas de la enfermedad:** los estudios de cohortes son los que permiten saber cuál es la incidencia de la enfermedad. Las medidas que se obtienen son:

- **Riesgo relativo:** es la medida de la fuerza de la asociación.
- **Diferencia de incidencias o riesgo atribuible:** informa sobre el exceso de riesgo de enfermar.
- **Fración atribuible:** estima la proporción de la enfermedad entre los expuestos que es debida al factor de riesgo.

Es necesario darse cuenta de que como en un estudio de cohorte todos los sujetos están sanos en el momento inicial, los enfermos que vayan apareciendo serán nuevos; es decir, incidentes. Por ello, la medida de frecuencia es la incidencia y la de asociación, el riesgo relativo.



**Estudios de cohortes históricas (retrospectivas)**

El investigador identifica, mediante registros, una cohorte expuesta en el pasado a un factor de riesgo, y otra cohorte no expuesta. Mediante dichos registros (por ej.: historias clínicas) sigue la evolución de ambas cohortes, comparando los resultados.

Por ejemplo, se selecciona una muestra de 100 historias clínicas de un hospital que se divide en dos grupos en función del antecedente de tabaquismo o no. Se observa en cada grupo la frecuencia de cáncer de laringe.

**Estudio de casos-contróles (Figura 2)**

Partiendo de un grupo de individuos enfermos (casos), y de otro comparable a ellos en todo, pero que no tienen la enfermedad (contróles), se estudia la exposición, en ambos, a distintos factores de riesgo. Por ejemplo, se comparan los factores de riesgo a los que ha estado expuesto un grupo de pacientes con cáncer de páncreas frente a los factores a los que ha estado un grupo de individuos sanos.

**Características:**

- Es un estudio longitudinal.
- Es retrospectivo.
- Va del efecto (enfermedad) a la causa.

**Ventajas:**

- Es de corta duración.
- Es ideal para el estudio de enfermedades raras.
- Es el mejor para el estudio de enfermedades de largo periodo de inducción.
- El coste es bajo.
- Es el más adecuado para el estudio de la multicausalidad de la enfermedad (los posibles factores de riesgo de una determinada enfermedad).
- Es el mejor para formular nuevas hipótesis etiológicas.

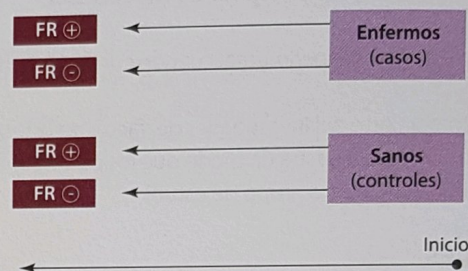


Figura 2. Estudio de casos-contróles

**Inconvenientes:**

- No es bueno para comprobar hipótesis previas de causalidad.
- No permite el estudio de la "multifactorialidad del factor de riesgo".
- La posibilidad de sesgos es muy alta, su control difícil.

**Análisis de las medidas de la enfermedad:** en los estudios de casos-contróles no es posible obtener información sobre la incidencia de la enfermedad, ya que se parte de una población seleccionada. Tampoco se tiene información acerca de la prevalencia, ya que el número de pacientes sólo depende de los que se elijan. Debido a ello, la fuerza de la asociación no se puede calcular directamente, como en el estudio de cohortes, sino de forma indirecta mediante la *odds ratio*.

En la **Tabla 1** se exponen las diferencias fundamentales entre los dos tipos de estudios analíticos observacionales.

Características	Cohortes	Casos-contróles
Población en riesgo	Definida	Indefinida
Cálculo de la incidencia	Sí	No
N.º de efectos medibles	Varios	Uno
N.º de exposiciones posibles a estudiar	Una	Varias
Enfermedades raras	Poco útil	Muy útil
Exposiciones raras	Muy útil	Poco útil
Sesgos posibles	Pocos	Muchos
N.º de sujetos en estudio	Muchos	Pocos
Coste	Elevado	Bajo
Duración	Larga	Corta
Evidencia causal	Regular	Mala

**Tabla 1.** Diferencias entre los estudios de cohortes y los de casos-contróles

## Ideas clave

- Los estudios experimentales son aquellos en los que el investigador forma parte "activa" del estudio al asignar el factor de estudio. La aleatorización de la muestra es otra condición indispensable.
- Los estudios observacionales son aquellos en los que el investigador se limita a "observar" a los sujetos enfermos-sanos (estudio de casos y controles), a sujetos expuestos-no expuestos (cohortes) o a observar cuántos sujetos están enfermos en una comunidad (transversal).
- Los estudios de cohortes son prospectivos porque se empieza a partir del factor de riesgo (FR hacia enfermedad). Es decir, que a

tiempo 0, nadie presenta el evento final de estudio. Son los mejores estudios de entre los observacionales para verificar causalidad. La medida de asociación es el RR.

- Los estudios de casos-controles son retrospectivos porque la selección de pacientes se hace en función de si están enfermos o no (enfermedad hacia FR). En este estudio a tiempo 0, ya hay sujetos portadores del evento final de estudio. Son los mejores para enfermedades raras y su principal limitación es la relativa validez por la facilidad para cometer sesgos. La medida de asociación es la OR.
- Los estudios transversales o de prevalencia se caracterizan por la ausencia de seguimiento. En estos estudios sólo "se ve" a los sujetos una única vez. La medida de frecuencia de enfermedad es la prevalencia.

## Casos clínicos

Para tratar de establecer una relación causal entre el consumo de benzodiacepinas durante el embarazo y el riesgo de fisura palatina en el recién nacido, se seleccionaron madres de recién nacidos con fisura palatina y se compararon con madres de recién nacidos sanos en cuanto a los antecedentes de toma de benzodiacepinas. ¿Cuál es el tipo de diseño de estudio empleado?

- 1) Casos y controles.
- 2) Estudio de cohortes.
- 3) Ensayo clínico aleatorizado.
- 4) Estudio ecológico.

RC: 1

¿A qué tipo de estudio corresponde el estudio de Framingham, que empezó en 1949 para identificar factores de riesgo de enfermedad coronaria, en el que, de entre 10.000 personas que vivían en Framingham de edades comprendidas entre los 30 y los 59 años, se seleccionó una muestra representativa de ambos sexos, reexaminándose cada dos años en busca de señales de enfermedad coronaria?

- 1) Ensayo clínico.
- 2) Estudio de cohortes.
- 3) Estudio de prevalencia.
- 4) Estudio de casos y controles.

RC: 2

## Case Study

A study without manipulation in the main variable that starts with patients without disease is a:

- 1) Quasi-experimental study.
- 2) Experimental study.

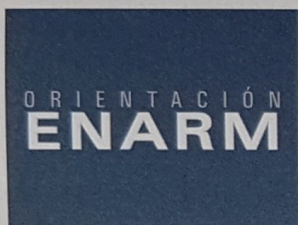
- 3) Cohort study.
- 4) Case-control study.

Correct answer: 3

# Validez y fiabilidad de los estudios epidemiológicos



Haz clic en la imagen para abrir tu plataforma



*Suelen presentarse estudios en los que se ha cometido algún error sistemático y puede no ser sencillo identificar el posible sesgo que se ha cometido.*

Cuando se realiza un estudio de investigación clínica, casi nunca se trabaja con poblaciones completas. Lo que se hace es, partiendo de observaciones realizadas en un grupo reducido de personas (la llamada muestra), generalizar o extrapolar los resultados obtenidos a colectivos más amplios. El hecho de que no se trabaje con poblaciones completas, sino con muestras, puede introducir en las observaciones errores producidos por el azar, a los que se les llama errores aleatorios. Existen además otros tipos de errores no relacionados con el hecho de trabajar con muestras, y que son conocidos como errores sistemáticos o sesgos.

## 7.1. Tipos de error

Los tipos de error son los siguientes:

- **Error aleatorio:** es el error que puede atribuirse a la variabilidad aleatoria que conlleva siempre un proceso de muestreo. El azar hace que la muestra con la que se va a trabajar no sea representativa. El error aleatorio no afecta a la validez interna de un estudio, pero reduce la posibilidad de elaborar conclusiones sobre la relación exposición-enfermedad, aunque no altera el sentido de la asociación. Los errores aleatorios, a diferencia de los errores sistemáticos, se pueden minimizar aumentando el tamaño de la muestra.
- **Errores sistemáticos o sesgos:** son los errores producidos cuando hay un fallo en el diseño o en la ejecución del estudio que hace que los resultados de la muestra sean diferentes de la población de la que proceden. No se relacionan con el tamaño de la muestra y, cuando no se controlan, tienden a invalidar las condiciones de un estudio; es decir, la existencia de sesgos conduce a la elaboración

de conclusiones incorrectas sobre la relación entre una exposición y una enfermedad.

La mayoría de los sesgos acontecen habitualmente en los estudios de casos-controles (retrospectivos).



## Tipos de errores sistemáticos

### Sesgo de selección

Se produce cuando se asignan sujetos al grupo control que difieren significativamente, en alguna característica clave, del grupo problema. Este tipo de sesgos se puede controlar mediante un proceso de aleatorización en la constitución de los distintos grupos de estudio.

Son ejemplos de este tipo de sesgo los siguientes:

- **Sesgo de autoselección o del voluntario:** la participación o auto-derivación del individuo al estudio compromete su validez.
- **Sesgo diagnóstico o sesgo de Berkson:** es el que se produce cuando, para saber qué ocurre en la población, se elige una muestra hospitalaria de esa población y el factor de riesgo que se está estudiando se asocia a una mayor probabilidad de hospitalización. También se incluye en este tipo de sesgo aquel que puede surgir al seleccionar como control a pacientes con alguna enfermedad que también se asocia al factor de exposición estudiado.
- **Sesgo del obrero sano:** la salida del trabajador enfermo del mercado laboral compromete la validez del estudio.
- **Falacia de Neyman:** se produce en los estudios de casos-controles al seleccionar casos prevalentes (ya existentes) en vez de casos incidentes (casos nuevos). Esto conlleva que en los casos sea menos frecuente la exposición a aquellos FR que disminuyen la supervivencia; esto es, los casos representarían a individuos con rasgos de mayor resistencia a la enfermedad o formas menos graves de esta.  
Por ejemplo: se quiere estudiar si la actividad física tiene algún efecto sobre la frecuencia de la insuficiencia coronaria. Se compara a per-



sonas con actividad y a personas sedentarias. Un sesgo de selección sería cuando la inactividad de los sedentarios fuese a consecuencia de padecer la enfermedad cardíaca.

### Sesgo de información

Tiene lugar cuando hay un error sistemático en la medición de alguna variable clave del estudio. La clasificación incorrecta puede o no afectar de igual manera a todos los grupos de estudio. Los sesgos de información debidos al investigador o a los participantes en el estudio se controlan mediante técnicas de enmascaramiento.

- **Clasificación incorrecta diferencial:** la probabilidad de clasificación errónea de la exposición (en un estudio de casos-controles) afecta de forma desigual a sanos y enfermos o la clasificación de enfermo o no enfermo (en un estudio de cohortes) se hace de manera distinta, según se esté o no expuesto al factor de estudio. Este tipo de sesgo produce una subestimación o sobreestimación del efecto del factor de exposición en la enfermedad. Dentro de este grupo cabe destacar:
  - **Sesgo de memoria:** se produce en los estudios de casos-controles. El hecho de padecer la enfermedad hace que se esté más motivado para recordar posibles antecedentes de exposición.
  - **Sesgo de atención o efecto Hawthorne:** los participantes en un estudio pueden modificar su comportamiento si saben que están siendo observados.

- **Clasificación incorrecta no diferencial:** la probabilidad de clasificación errónea ocurre en todos los grupos de estudio de forma similar. El error de clasificación no diferencial ocasiona una infraestimación (hace que la medida de asociación tienda a 1) del efecto del factor de exposición estudiado en la enfermedad. Por ejemplo: se quiere estudiar el efecto del tabaco sobre la aparición de bronquitis crónica, comparándose a un grupo de fumadores y a otro de no fumadores. Existe la posibilidad de que los fumadores sean diagnosticados más fácilmente de bronquitis crónica que los no fumadores, simplemente por el hecho de que fumar se considera que está asociado a esta patología (Tabla 1).

### Factor de confusión (confounding)

Un factor de confusión es una variable extraña al estudio que modifica los resultados que se obtienen. Todo factor de confusión debe cumplir tres condiciones:

- Ser un factor de riesgo para la enfermedad.
- Estar asociado con la exposición.
- No ser un paso intermedio entre la exposición y la enfermedad.

Para prevenir los sesgos de confusión, existen distintas técnicas:

- **Fase de diseño:** aleatorización (estudios experimentales), apareamiento y restricción.
- **Fase de análisis estadístico:** análisis estratificado (se dividen los datos globales en dos grupos, según la presencia o no del factor de confusión) y análisis multivariante.

**Ejemplo:** se desea saber si existe relación entre el alcohol y el cáncer de pulmón. La OR cruda es de 2,4.

Se sospecha que el tabaco puede actuar como factor de confusión. Al estratificar, se observa lo siguiente:

- OR cruda: 2,4.
- OR en fumadores: 1,5.
- OR en no fumadores: 1,5.

El tabaco es un factor de confusión, puesto que el valor de la OR se ha modificado en los estratos (de 2,4 a 1,5). La verdadera OR entre alcohol y cáncer de pulmón es de 1,5.

## 7.2. Validez y fiabilidad

Se definen los siguientes conceptos:

- **Validez:** el estudio cuantifica lo que realmente se propone medir. Es el grado de ausencia de error sistemático. También recibe el nombre de exactitud.
  - **Validez interna:** es el grado de validez del resultado para los pacientes del estudio. Se dice que un estudio tiene validez interna cuando los resultados del mismo son aplicables a los individuos del citado estudio.
  - **Validez externa:** un estudio tiene validez externa cuando los resultados del mismo son aplicables a otros individuos distintos de los del estudio.
- **Fiabilidad:** es el grado de reproducibilidad de un estudio, es decir, el grado de similitud que presentarían los resultados si se repitiese el estudio en condiciones similares. Es decir, es el grado de ausencia de error aleatorio.

	Características	Afecta a la validez interna	Solución o control
Error aleatorio	<ul style="list-style-type: none"> <li>• Simétrico</li> <li>• Incorregible</li> <li>• Impredecible</li> <li>• Estimable por estadística</li> </ul>	NO	↑ tamaño muestral
Sesgo o error sistemático	<ul style="list-style-type: none"> <li>• Asimétrico</li> <li>• Corregible</li> <li>• Predecible</li> <li>• Prevención y control por epidemiología</li> </ul>	SÍ	<ul style="list-style-type: none"> <li>• Sesgo de selección: aleatorización</li> <li>• Sesgo de clasificación: enmascaramiento</li> <li>• Factor de confusión:                             <ul style="list-style-type: none"> <li>- Aleatorización</li> <li>- Apareamiento</li> <li>- Restricción</li> <li>- Estratificación</li> <li>- Análisis multivariante</li> </ul> </li> </ul>

Tabla 1. Tipos de error

### 7.3. Validez de un test diagnóstico

Para conocer si una prueba diagnóstica es útil, se comparan sus resultados con los de otra prueba que actúa como patrón de referencia (*gold standard*). El grado en el que las medidas de la prueba evaluada se correlacionan con las de referencia se denomina validez de criterio (**Figura 1**).

La prueba de referencia debe haber sido aceptada como tal por la comunidad científica, se ha de aplicar a toda la serie de casos estudiados y no puede incorporar información procedente de la prueba que evalúa.

El valor real de la prueba sólo podrá ser establecido si el estudio se lleva a cabo en condiciones semejantes a la práctica clínica habitual; es decir, incorporando un amplio espectro de pacientes a los que en condiciones normales se les aplicaría dicho procedimiento diagnóstico.



Figura 1. Validez y fiabilidad

### Ideas clave

- Los errores aleatorios son errores producidos por el azar. Son una falta de precisión, no afectan a la validez interna y se corrigen aumentando el tamaño de la muestra.
- Los errores sistemáticos o sesgos no se deben al azar, afectan a la validez de los estudios y no se corrigen aumentando el tamaño de la muestra.
- Los sesgos de selección se deben al hecho de elegir mal la muestra del estudio, bien por escoger controles hospitalarios (Berkson), por disponer de pocos casos enfermos al estudiar enfermedades de rápida mortalidad (Neyman), o por seleccionar voluntarios.

- Los errores de clasificación se deben a un error al incluir a los pacientes en los diferentes grupos de estudio, ya sea porque los sanos (controles) no recuerdan el antecedente de la exposición (sesgo de memoria), o porque varían su comportamiento al sentirse observados.
- En los ensayos clínicos, se pueden controlar los sesgos de clasificación mediante las técnicas de ciego o enmascaramiento; y los de selección, por medio de la aleatorización.
- El sesgo de confusión se debe a la existencia de un factor de confusión, que es esencialmente un factor que está desigualmente distribuido entre los dos grupos de pacientes del estudio y que se relaciona etiológicamente con la enfermedad final. Los sesgos de confusión se controlan mediante asignación aleatoria, apareamiento, restricción, análisis multivariante y análisis estratificado.

### Casos clínicos

En un estudio de casos y controles se encontró una asociación entre la ingesta de alcohol y el desarrollo de cáncer esofágico. Dado que el consumo de tabaco se encontró asociado tanto a la ingesta de alcohol como al desarrollo de cáncer de esófago, el consumo de tabaco en este estudio debe considerarse como:

- 1) Un factor (sesgo) de confusión.
- 2) Un sesgo de información.
- 3) Un sesgo de memoria.
- 4) Un sesgo de falacia ecológica.

RC: 1

En un estudio de casos y controles se estudia la relación entre cáncer de cavidad bucal (efecto) y consumo de bebidas alcohólicas (exposición). ¿Cómo definiría el fenómeno que puede producir el hecho de que el hábito de fumar se asocie a la aparición de cáncer bucal y también a un mayor consumo de bebidas alcohólicas?

- 1) Sesgo de información.
- 2) Sesgo por factor de confusión.

- 3) Un error aleatorio.
- 4) Sesgo de observación.

RC: 2

Suponga que, en la lectura de un ensayo clínico cuya variable principal de respuesta es la mortalidad a cinco años, observa usted que los pacientes de un grupo tienen una media de edad al inicio del tratamiento de 64 años y los del otro de 52 años, no siendo estadísticamente significativas estas diferencias. ¿Cómo cree que podría afectar este hecho a los resultados del ensayo?

- 1) De ninguna forma, ya que las diferencias iniciales en la edad se deben al azar.
- 2) Invalida el estudio, ya que los grupos no tienen inicialmente el mismo pronóstico.
- 3) Invalida el estudio, ya que significa que la asignación al azar de los pacientes a los grupos (aleatorización) no ha producido grupos inicialmente homogéneos y, por tanto, ha sido incorrecta.
- 4) Haría necesario utilizar alguna técnica estadística que permita comparar la mortalidad entre grupos, teniendo en cuenta las posibles diferencias entre ellos en las características basales.

RC: 4

### ORIENTACIÓN ENARM

*Uno de los temas más importantes en el ensayo clínico, ya sea por sus conceptos teóricos ya sea por el tipo de ensayo que se trate.*

## 8.1. Introducción

Un ensayo clínico es un experimento cuidadosa y éticamente diseñado con el fin de dar respuesta a preguntas que tienen que ver con la utilidad de procedimientos diagnósticos, terapéuticos y profilácticos en el ser humano.

Los ensayos clínicos son estudios prospectivos y experimentales en los que, una vez seleccionada la muestra, se divide aleatoriamente en dos grupos de pronóstico comparable que idealmente sólo se diferencian en la intervención terapéutica que van a recibir.

Teniendo en cuenta los objetivos perseguidos en el desarrollo de un medicamento, se distinguen cuatro fases:

- **Ensayo en fase I:** es la primera vez que un fármaco se administra a humanos. Generalmente, se realiza con voluntarios sanos ( $n = 20-80$ ) y sin grupo control. Suelen ser EC abiertos, no aleatorizados y, muchas veces, no comparativos. El principal objetivo es evaluar la toxicidad y conocer la dosis única aceptable no tóxica. Se limitan al terreno experimental y/o a sus efectos farmacocinéticos y farmacodinámicos.
- **Ensayo en fase II:** el principal objetivo es aportar información sobre la relación dosis/respuesta, proporcionando una información preliminar acerca de la eficacia. Se lleva a cabo en pacientes ( $n = 100-200$ ). No necesariamente tienen que ser comparativos. Generalmente, este tipo de EC es controlado y con asignación aleatoria del tratamiento. Algunas clasificaciones distinguen a su vez una fase IIa (estudios piloto realizados en pocos pacientes con criterios de inclusión/exclusión más estrictos) y una fase IIb (que evalúa la eficacia y la seguridad en un mayor número de pacientes y representa una demostración más rigurosa de la eficacia del nuevo compuesto).
- **Ensayo en fase III:** es el prototipo del ensayo clínico. Suele ser comparativo con la terapéutica de referencia o con un placebo. Es la inves-

tigación clínica más extensa y rigurosa sobre un tratamiento médico. Sirve para establecer la eficacia de un nuevo fármaco y la existencia de efectos adversos frecuentes. Un paciente debe conocer que probablemente ha recibido placebo, pero no si lo ha recibido con certeza o no. Lo ideal es comparar frente al mejor tratamiento, a la mejor alternativa existente. Se realiza en un grupo mayor de pacientes (varias decenas de miles), seleccionados con unos criterios de inclusión/exclusión más laxos, que normalmente se encuentran polimedicados y que permitan hacer una extrapolación de los resultados al resto de la población (validez externa). Es posible, a su vez, distinguir las fases IIIa y IIIb. La primera incluye EC que se llevan a cabo una vez determinada la eficacia terapéutica, pero antes de ser enviada a las Autoridades Reguladoras la solicitud de aplicación del nuevo compuesto, y constituyen la evidencia de efectividad terapéutica que se exige para la tramitación del expediente del nuevo compuesto. Con frecuencia son considerados como pivotal trials. La segunda comprende estudios que se realizan una vez aceptada la tramitación y antes de la aprobación y comercialización del nuevo compuesto.

- **Ensayo en fase IV:** también se denomina farmacovigilancia, y consiste en el seguimiento, posterior a la comercialización del fármaco, de un número muy elevado de pacientes con el fin de detectar efectos adversos poco frecuentes o de aparición tardía. La fase IV también sirve para evaluar interacciones entre medicamentos y sugerir nuevas indicaciones de fármacos ya aceptados para otro fin.

Así, la farmacovigilancia es la recogida de datos sobre seguridad de fármacos una vez que se autoriza su comercialización. Su objetivo es detectar reacciones adversas poco frecuentes que no han sido identificadas durante el desarrollo del ensayo, dado que este se ha realizado en una población limitada (generalmente  $n < 2.000$ ).

En general, se acepta cualquier reacción adversa. Sin embargo, las de mayor interés son las reacciones adversas graves, inesperadas o raras y aquellas de fármacos de comercialización reciente (menos de tres años).

## 8.2. Etapas en la realización de un ensayo clínico

A continuación se desarrollan los pasos a seguir para la realización de un ensayo clínico en fase III con dos grupos de intervención.

## Selección de la cohorte de estudio

Consiste en la elección de un grupo de sujetos adecuados para la evaluación del medicamento, de los cuales se extrapolarán los resultados del ensayo clínico. Mediante la formulación de los criterios de inclusión, se establecen las características clínicas y sociodemográficas de los pacientes en los que se empleará el fármaco. Los criterios de exclusión se usan para rechazar a aquellos pacientes con criterios de inclusión pero que presentan alguna contraindicación, condiciones que pueden afectar a la variable resultado o alguna característica que haga complicado su estudio.

En función de los criterios de inclusión, se puede dividir los EC en dos grandes grupos:

- **EC pragmáticos:** se acercan lo más posible a la población general. Los criterios de inclusión son muy laxos (prácticamente los únicos son los diagnósticos).
- **EC explicativos:** los criterios de inclusión son muy estrictos, por lo que la muestra no es exactamente igual a la población general.

Si los criterios de inclusión son muy estrictos, el EC presenta una muestra más homogénea y se necesita menor tamaño muestral para detectar las diferencias. Además, hay mayor validez interna. Sin embargo, existe una clara limitación a la hora de generalizar los resultados y de reclutar a los pacientes. Se utilizan, sobre todo, en las fases II y III. En cambio, cuando no son estrictos, el reclutamiento se verá facilitado, se podrán generalizar los resultados con mayor facilidad y tendrá mayor validez externa. Sin embargo, la muestra será heterogénea, por lo que se necesita un mayor tamaño muestral y los resultados son confusos si sólo es eficaz en subgrupos de pacientes. Se emplean en las fases III y IV.

En esta etapa, también hay que determinar el tamaño muestral. Este ha de ser suficiente para obtener un IC del 95% de la eficacia cuyos límites sean clínicamente relevantes.

## Medición de variables basales

La medición de variables basales consiste en determinar una serie de variables en los sujetos que cumplen los criterios de inclusión y que han aceptado participar en el estudio (consentimiento informado), con los siguientes propósitos:

- Definir las características de la población a la que podrá extrapolarse el resultado del ensayo clínico.
- Aportar una serie de datos que permitan posteriormente verificar que la aleatorización ha sido eficaz y hacer el análisis estratificado (por ejemplo: por edad), en caso de diferencias entre los dos grupos.
- Demostrar que el evento resultado, cuya frecuencia se pretende disminuir con el fármaco estudiado, no está presente al comienzo del estudio.
- Registrar posibles predictores del evento interés, de modo que se pueda evaluar su interacción con el fármaco estudiado, si la aleatorización no los ha distribuido homogéneamente entre los distintos grupos de estudio.

## Aleatorización

La aleatorización consiste en asignar por azar, sin que influya ningún factor, los pacientes de la muestra a los dos grupos de intervención de modo que, si el tamaño muestral es suficientemente grande, se consiga una distribución homogénea de las variables predictoras en ambos grupos.

Se puede realizar de tres formas diferentes:

- **Aleatorización simple:** cada paciente tiene la misma probabilidad de ser asignado a cualquiera de los grupos de tratamiento. Con este método, existe riesgo de desigualdad numérica entre grupos.
- **Aleatorización por bloques:** se establecen bloques de aleatorización, de modo que en cada uno de ellos la mitad de los pacientes reciba el tratamiento experimental y la otra mitad, el control. Con este tipo de aleatorización se evita la desigualdad numérica entre los distintos grupos, por lo que es especialmente útil cuando el tamaño muestral no es muy grande.
- **Aleatorización estratificada:** los pacientes son divididos en grupos (estratos) homogéneos respecto a alguna variable de interés pronóstico y posteriormente son asignados aleatoriamente a uno de los dos grupos de intervención (**Figura 1**).

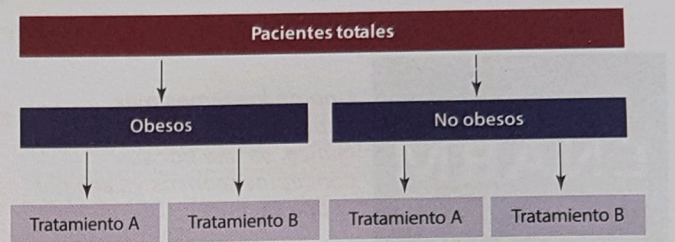


Figura 1. Aleatorización estratificada

En general, la aleatorización debe conseguir asignar los tratamientos de forma desconocida e impredecible, tiene que ser un proceso reproducible y estar documentado, ha de basarse en propiedades matemáticas conocidas, no debe ser prevista la secuencia de los tratamientos y, finalmente, ha de ser posible detectar los fallos de la asignación. Para evitar sesgos en la aleatorización es importante que quien decide la inclusión de los pacientes en el ensayo y la persona que lleva a cabo la aleatorización desconozcan la secuencia de aleatorización hasta que se aplique a cada uno de los pacientes reclutados. En caso de no hacerse ocultación de la secuencia de aleatorización, existe la posibilidad de seleccionar incorrectamente a los sujetos del ensayo y perder la comparabilidad de los dos grupos de tratamiento. La ocultación de la secuencia de aleatorización se puede conseguir mediante un sistema de aleatorización centralizado en un lugar distinto a aquel en el que se lleva a cabo el ensayo o mediante sobres numerados secuencialmente, opacos y lacrados, que contengan el tratamiento y que sólo se abran tras la inclusión del paciente en el EC.

## Aplicación de la intervención

Es importante evitar que tanto investigadores como otros proveedores de cuidados traten de forma diferente a los dos grupos de intervención. Para evitar este hecho, llamado sesgo de cointervención, y que la medición de la variable resultado pueda estar sesgada por el mismo motivo, se utiliza el enmascaramiento. Los tipos de enmascaramiento son los siguientes:

- **Simple ciego:** el paciente desconoce el grupo de tratamiento al que está asignado.
- **Doble ciego:** paciente, personal sanitario y evaluador de la variable resultado lo desconocen.
- **Triple ciego:** además de los anteriores, el analista de los datos no sabe el tipo de tratamiento de cada uno de los grupos.

Tanto las reacciones adversas específicas como la falta de eficacia y los cambios analíticos específicos pueden desenmascarar un EC. Otro aspecto

to a considerar en esta fase del ensayo clínico es que siempre que exista un tratamiento eficaz, hay que administrarlo al grupo control, así que lo que se determinará con el ensayo clínico será la eficacia relativa del nuevo fármaco. En caso de no existir alternativa terapéutica previa, se medirá la eficacia en términos absolutos.

### Análisis de los resultados (Tabla 1)

En este punto se deben tener en cuenta los siguientes aspectos:

- Las pérdidas de sujetos incluidos en el ensayo clínico ocurridas antes de la aleatorización van a afectar a la capacidad de generalización de los resultados, mientras que las pérdidas postaleatorización pueden afectar a la validez interna.
- El análisis estadístico de los ensayos clínicos es muy parecido al de los estudios de cohortes, pero es más frecuente el uso de métodos no paramétricos y, al tener importancia no sólo que ocurra la variable resultado, sino cuándo ocurre, es frecuente el empleo de análisis de supervivencia.
- Comparaciones múltiples: al aumentar el número de comparaciones, se incrementa la posibilidad de resultados falsamente positivos.
- Análisis de subgrupos: puede ocasionar problemas, especialmente cuando no se tiene previsto desde el principio. Produce comparaciones múltiples, aumenta la probabilidad de resultados espurios y, por tanto, sus resultados deben interpretarse con precaución.
- Análisis por intención de tratar frente a análisis por protocolo:
  - Por protocolo:** incluye sólo a aquellos pacientes que han cumplido los requisitos del protocolo y/o han finalizado el estudio.
  - Por intención de tratar:** el análisis incluye a todos los pacientes que han sido seleccionados y en el grupo en el que fueron asignados, aunque no hayan finalizado el estudio o hayan cambiado de grupo.

<b>Pragmáticos</b>	Muestra heterogénea, validez externa
<b>Explicativos</b>	Muestra homogénea, validez interna
<b>Cruzado</b>	Cada grupo toma ambos fármacos con periodo de lavado
<b>Paralelo</b>	Cada grupo toma un único fármaco
<b>Por intención de tratar</b>	Todos los sujetos incluidos son analizados
<b>Por protocolo</b>	Sólo se analizan los que terminan el estudio
<b>De tamaño predeterminado</b>	Tamaño muestral conocido a priori
<b>Secuenciales</b>	Incorporación progresiva de pacientes al estudio
<b>De superioridad</b>	Demostrar que un fármaco es mejor que otro
<b>De no inferioridad</b>	Se supone una eficacia similar, el nuevo fármaco tiene otras ventajas
<b>Enmascarados</b>	Empleo de técnicas de ciego para evitar sesgos
<b>Abiertos</b>	No enmascarados

Tabla 1. Clasificación de los ensayos clínicos

- Los análisis intermedios se realizan durante las fases III y IV del estudio. Están justificados para evitar que los pacientes del grupo control no se beneficien del tratamiento. Sin embargo, el realizar muchos análisis intermedios aumenta el riesgo de cometer un error tipo I, debe estar previsto en el protocolo del estudio, se ha de mantener la confidencialidad de los resultados y tiene que haber una penaliza-

ción estadística porque, de nuevo, se están realizando comparaciones múltiples.

- La magnitud del efecto del tratamiento en los ensayos clínicos se puede expresar de varias formas:
  - El riesgo relativo:** cociente entre el riesgo de sufrir un determinado evento en el grupo expuesto a un determinado tratamiento y el de sufrir el mismo evento en el grupo control (no expuesto al tratamiento).
  - La reducción absoluta del riesgo (RAR):** que es la diferencia entre el porcentaje de eventos en el grupo control y el mismo porcentaje en el grupo experimental.
  - La reducción relativa del riesgo (RRR):** se define como la diferencia entre el porcentaje de eventos en el grupo control y el mismo porcentaje en el grupo experimental, dividido por el porcentaje de eventos en el grupo control.
  - El **número necesario de pacientes a tratar (NNT):** que se obtiene como el inverso de la RAR multiplicado por 100, y es el número de pacientes que se deben tratar para prevenir un suceso indeseable adicional.

### 8.3. Tipos de ensayos clínicos

Los tipos de ensayos clínicos son los siguientes:

- Diseño clásico o en paralelo:** el grupo control recibe el tratamiento a la vez que el grupo experimental con el fin de controlar el efecto de factores pronósticos que pudieran cambiar a lo largo del tiempo. El análisis de los datos consiste en comparar la diferencia en la variable resultado entre ambos grupos con la variabilidad esperable dentro de cada grupo por el mero efecto del azar. Es muy apropiado en los casos de enfermedades agudas (Figura 2).

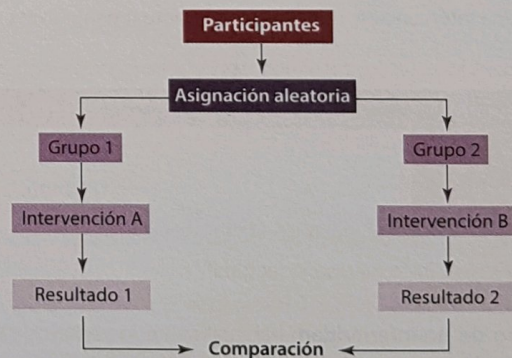


Figura 2. Ensayo clínico paralelo

- Diseño cruzado o intrapaciente:** consiste en que un mismo paciente va a recibir los dos tratamientos objeto de comparación en dos momentos distintos, de modo que el propio paciente sirve de control a sí mismo, lo que permite llevar a cabo el estudio con un tamaño muestral menor al del diseño clásico al reducir la variabilidad (Figura 3).

Uno de los problemas es el efecto carry-over o de arrastre (el efecto de uno puede influir sobre la toma del otro), por lo que debe haber un periodo de lavado entre la administración de ambos fármacos, o entre las intervenciones terapéuticas, de modo que el primero haya sido totalmente aclarado del organismo antes de administrar el segundo.

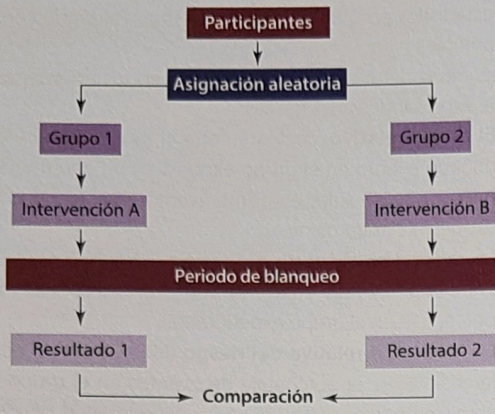


Figura 3. Ensayo clínico cruzado

- **Diseño secuencial:** consiste en introducir pares de pacientes aleatorizados a los dos tratamientos hasta que la diferencia entre los distintos tratamientos favorezca a uno u otro (exceso de preferencias), momento en el que el ensayo clínico se detiene.
- **Diseño factorial:** en ese tipo de diseño se evalúan simultáneamente dos tratamientos distintos en una misma muestra de sujetos, asignándose aleatoriamente a los sujetos a uno de los cuatro grupos posibles: A, B, A + B y placebo (Tabla 2). Con este diseño es posible evaluar la interacción entre los dos tratamientos, aunque hay que asumir que entre ambos no hay interacciones. La evaluación de dos o más intervenciones puede hacerse mediante un diseño de tipo paralelo. Sin embargo, esto es posible que requiera un gran tamaño muestral y sea ineficiente, especialmente si hay también interés en considerar combinaciones de las intervenciones. El diseño más apropiado en esta situación es de tipo factorial, donde se pueden asignar de manera aleatoria dos o más intervenciones de forma independiente, de tal modo que los sujetos pueden no recibir ninguna intervención, una de ellas o eventualmente todas.

Tratamiento	A	Placebo
B	A + B	B
Placebo	A	Placebo

Cada paciente se asigna a uno de los 4 grupos resultantes

Tabla 2. Ensayo clínico de diseño factorial

- **Diseño de no-inferioridad:** habitualmente los ensayos clínicos se realizan con el fin de demostrar la superioridad de un fármaco con respecto a otro. Sin embargo, puede ser útil probar que dos fármacos son al menos igualmente eficaces, con lo cual se podrá emplear, por ejemplo, el más barato. Es decir, importa que NO sea inferior, de manera que si es igual o superior se considera que el fármaco experimental es "no-inferior" comparado con el grupo control. Se denomina equivalente terapéutico a un fármaco diferente en su estructura química del original, pero del que se espera un efecto terapéutico y un perfil de efectos adversos similares cuando se administra a un paciente. La equivalencia terapéutica se define a partir de los estudios de no-inferioridad. En estos estudios es importante tener en cuenta dos aspectos:
  - **Valor delta:** es la máxima diferencia clínica que se acepta para definir dos tratamientos como equivalentes, que es distinta para cada tipo de fármaco analizado. Habitualmente se considera un

25-50% de las diferencias detectadas en estudios de superioridad previos.

- **Análisis de las variables:** el análisis que suele realizarse es "por protocolo". Este análisis permite aumentar las diferencias entre los tratamientos, lo que dificulta poder concluir que dos tratamientos son equivalentes, manteniendo la posición más cauta en la interpretación de los resultados.

## 8.4. Estudios de bioequivalencia

Los estudios de bioequivalencia son ensayos clínicos en los que su objetivo es la demostración de la similitud en biodisponibilidad de dos formulaciones de un mismo principio activo a partir de la comparación de sus características farmacocinéticas.

En la mayoría de los casos, el diseño de los estudios de bioequivalencia es el de un ensayo clínico cruzado y con asignación aleatoria de dos secuencias de tratamiento.

Con el fin de reducir la variabilidad de la muestra, se suelen elegir para participar en el estudio voluntarios sanos (no persiguen, por tanto, comparar la eficacia de dos formas de tratamiento, dependiendo el tamaño de la muestra de la variabilidad interindividual (diferente farmacocinética para cada sujeto) que se haya observado en estudios piloto previos.

La demostración de bioequivalencia se obtiene mediante la comparación de los perfiles farmacocinéticos de los fármacos estudiados. Para ello, después de la administración de cada formulación, es necesario saber qué cantidad de fármaco existe en el organismo y cómo va variando a lo largo del tiempo.

El procedimiento más habitual consiste en la obtención de sucesivas de muestras de sangre. Es mucho menos frecuente que sea necesaria la determinación del fármaco en orina o en otras muestras biológicas.

## 8.5. Ensayo clínico de tamaño único

El ensayo clínico de tamaño único (n: 1) se basa en modificar el tratamiento de una enfermedad de un paciente a lo largo de una serie determinada de periodos de tiempo. El fin del estudio es confirmar la eficacia o no del tratamiento en un único paciente.

Como en este caso hay un único paciente, no es posible compararlo con otros, por lo que aquí se comparan periodos de tratamiento para el mismo paciente. Por ejemplo, el paciente L.G. recibe durante un tiempo el tratamiento A y luego el B. Es decir, en un ensayo n: 1, la unidad experimental es el periodo de tratamiento y no el paciente. Se compara la eficacia del tratamiento A frente al B en el señor L. Por supuesto, la secuencia de tratamientos puede hacerse al azar y con técnicas de enmascaramiento. El ensayo clínico n: 1 es muy útil para:

- Evaluar el tratamiento sintomático en enfermedades en las que no se haya encontrado una intervención eficaz en ensayos convencionales, como por ejemplo en enfermedades raras.
- En aquellas situaciones en las que exista una gran variabilidad interindividual en la respuesta al tratamiento, como en el caso de tratamiento con dicumarínicos.

- Tratamiento sintomático de enfermedades con importante componente subjetivo, como por ejemplo migrañas, dolores articulares en la artrosis, fibromialgia...

Es necesario que la condición estudiada sea crónica, con síntomas repetidos pero reversibles y con respuesta al tratamiento evaluable por el paciente (por ej., cefalea tensional). Se evalúa, por tanto, un tratamiento sintomático y no curativo.

Por las características del estudio es evidente que las conclusiones obtenidas no son generalizables a todos los pacientes con la misma enfermedad o síntoma.

### 8.6. Recomendaciones para la publicación de ensayos clínicos (recomendaciones CONSORT)

Los estudios aleatorizados representan el patrón de referencia para conocer la eficacia de un determinado tratamiento. Sin embargo, la calidad de los estudios puede no ser óptima. La declaración CONSORT (*Consolidated Standards of Reporting Trials*) persigue mejorar la metodología de los ensayos clínicos, ya que permite presentar de forma sistematizada y con total transparencia los resultados para mejorar su validez. Se basa en el seguimiento de una lista de variables en función de la evidencia (Tabla 3) y un diagrama de flujo (Figura 4) durante el proceso de escritura, revisión y análisis de los ensayos clínicos.

Se ha demostrado que la implementación de los criterios CONSORT mejora la calidad de la información publicada en los estudios aleatorios.

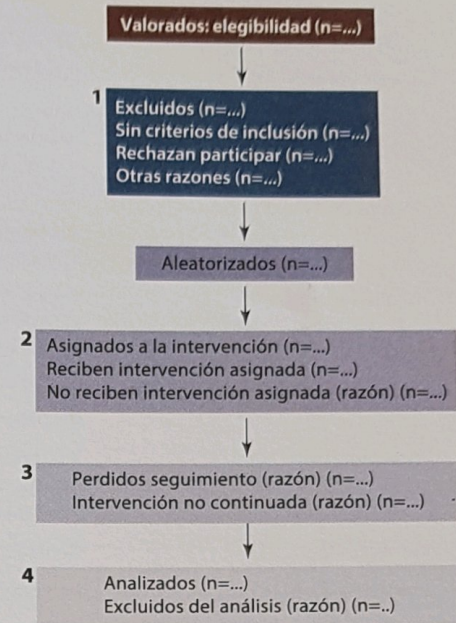


Figura 4. Diagrama de flujo de los sujetos participantes en las diferentes fases de un estudio aleatorizado: (1) Periodo de inclusión, (2) Asignación a la intervención, (3) Seguimiento, (4) Análisis

Variable (número)		Descripción	
Título y resumen (1)		Cómo se asignó a los pacientes a la intervención ("asignación aleatorizada", "aleatorizados" o "asignados aleatoriamente")	
Introducción	Antecedentes (2)	Antecedentes científicos y explicación de la hipótesis	
Métodos	Participantes (3)	Criterios de elegibilidad de los participantes y situación y lugar donde los datos se obtuvieron	
	Intervenciones (4)	Detalles precisos de las intervenciones planeadas en cada grupo y cómo y cuándo se realizaron finalmente	
	Objetivos (5)	Objetivos e hipótesis específicos	
	Resultados (6)	Definición clara de los objetivos de valoración de resultados primarios y secundarios y, en su caso, de los métodos utilizados para mejorar la calidad de las mediciones (por ej.: observaciones múltiples, entrenamiento de los valoradores)	
	Tamaño de muestra (7)	Cómo se determinó el tamaño muestral y, en su caso, explicación de cualquier análisis interno y de las reglas para la finalización del estudio	
	Aleatorización	Generación de la consecuencia (8)	Método utilizado para generar la secuencia de asignación aleatoria, incluyendo los detalles de cualquier restricción (por ej.: bloques, estratificación)
		Ocultación de la asignación (9)	Método utilizado para implementar la secuencia de asignación aleatoria (sistemas cerrados numerados, central telefónica) especificando si la secuencia permaneció oculta hasta el momento de la asignación a la intervención
		Implementación (10)	Quién generó la secuencia de asignación, quién incluyó a los participantes y quién asignó a los participantes a sus grupos respectivos
	Cegamiento (enmascaramiento) (11)	Especificar si los participantes, las personas que realizaron o asignaron la intervención y las que valoraron los resultados conocían a qué grupo estaban asignados los pacientes. En caso contrario, cómo se valoró el éxito del enmascaramiento	
	Métodos estadísticos (12)	Métodos estadísticos utilizados para comparar la variable de valoración de resultado primaria entre los grupos. Métodos empleados en los análisis adicionales, como los análisis de subgrupos y ajustes	

Tabla 3. Lista de variables CONSORT (continúa en la página siguiente)

	Variable (número)	descripción
Resultados	Flujo de participantes (13)	Flujo de participantes en cada fase del estudio (se recomienda encarecidamente usar un diagrama). De forma específica se debe detallar en cada grupo el número de pacientes asignados a cada tratamiento, los que recibieron el tratamiento previsto, los que completaron el protocolo de estudio y los que fueron analizados para la variable principal de resultado. Describir las desviaciones del protocolo inicial y sus razones
	Selección (14)	Fechas que definan los periodos de selección y seguimiento
	Datos basales (15)	Características demográficas y clínicas basales de cada grupo
	Números analizados (16)	Número de participantes (denominador) de cada grupo incluido en cada análisis especificando si el análisis se realizó por intención de tratar. Si es posible, indicar los resultados con números absolutos (por ej.: 10/20, no 50%)
	Resultados y estimaciones (17)	Para cada variable de valoración de resultado primaria y secundaria, indicar resumidamente los resultados en cada grupo y la estimación de la importancia del efecto y de su precisión (por ej.: IC del 95%)
	Análisis secundarios (18)	Valorar la presencia de multiplicidad presentando cualquier otro análisis realizado, incluyendo análisis de subgrupos y ajustes, indicando los que fueron preespecificados y los de carácter exploratorio
	Acontecimientos adversos (19)	Indicar todos los acontecimientos adversos importantes o los efectos secundarios en cada grupo de intervención
Discusión	Interpretación (20)	Interpretación de resultados en la que se tenga en cuenta la hipótesis del estudio, las posibles fuentes de sesgos o imprecisiones y los peligros asociados con la multiplicidad de análisis y de resultados
	Generabilidad (21)	Generabilidad (validez externa) de los resultados del estudio
	Evidencia global (22)	Interpretación general de los resultados en el contexto de la evidencia actualmente disponible

Tabla 4. Lista de variables CONSORT (continuación)

## Ideas clave

- El objetivo de la fase I del ensayo clínico es determinar las propiedades farmacodinámicas del fármaco, así como la dosis tóxica del mismo.
- En una fase II se establece la dosis eficaz del fármaco.
- La fase III se caracteriza porque se compara la eficacia de otro fármaco con la del que se está investigando. Tras esta fase, el fármaco se introduce en el mercado.
- La fase IV es la fase de la farmacovigilancia. En este caso, se registra la aparición de efectos secundarios infrecuentes. La notificación en la práctica clínica de posibles eventos adversos se realiza mediante la tarjeta amarilla.
- La correcta secuencia en un ensayo clínico es:
  1. Los pacientes cumplen los criterios de inclusión y exclusión.
  2. Firman el consentimiento informado (sabiendo que posiblemente reciban placebo, si uno de los tratamientos lo es).
  3. Sean asignados a uno u otro grupo de intervención (preferiblemente mediante aleatorización).
- Un aspecto importante es manejar las pérdidas que se producen en un ensayo clínico. Lo mejor es tenerlas previstas a priori. Si aún así se produjeran, sería conveniente realizar un nuevo estudio con el tamaño apropiado. Sin embargo, no todas las pérdidas afectan de igual manera al EC. Si las pérdidas se producen antes de la asignación del factor de estudio, no necesariamente se ve afectada la validez del mismo, al contrario que si se producen después de tal asignación.
- Los EC se dividen en:
  1. Pragmáticos (muestra heterogénea) y explicativos (muestra homogénea).
  2. Cruzado (cada grupo de sujetos toma los dos fármacos con un tiempo de lavado) y paralelo (cada grupo de tratamiento toma un único fármaco).
  3. Por intención de tratar (todos los sujetos incluidos en el estudio son analizados) y por protocolo (sólo se analizan los que finalizan el estudio).
  4. De tamaño predeterminado (se conocen a priori el tamaño de la muestra) o secuenciales (los pacientes van incorporándose al estudio progresivamente).
  5. De superioridad (interesa demostrar que un fármaco es mejor que otro) y de no-inferioridad (se supone una eficacia similar, la ventaja del nuevo fármaco es ser más barato, tener una administración más fácil...).
  6. Enmascarados (se han realizado técnicas de ciego para evitar sesgos) o abiertos (no enmascarados).



## Casos clínicos

En un estudio de seguimiento a largo plazo de pacientes con fibrilación auricular, en la práctica clínica habitual, se ha documentado un riesgo relativo de padecer un accidente vascular cerebral, para anticoagulantes orales (en comparación con antiagregantes plaquetarios), de 0,57 (intervalo de confianza al 95% 0,40-0,80). La incidencia de accidente vascular cerebral durante los dos primeros años de seguimiento es de un 2% y un 3,5% respectivamente. Suponiendo que estos resultados sean válidos, ¿a cuántos pacientes habría que tratar con anticoagulantes orales, en lugar de antiagregantes, para evitar un accidente vascular cerebral?

- 1) 50-60.
- 2) 60-70.
- 3) 70-80.
- 4) 80-90.

RC: 2

Spongua que, en la lectura de un ensayo clínico en el cual se han incluido 50 pacientes por grupo, observa usted que los resultados se refieren a 47 pacientes en un grupo y 42 en el otro. Con respecto a los pacientes que faltan, el autor comunica que fueron retirados por no seguir el tratamiento y que no hay diferencias estadísticamente significativas entre grupos en cuanto al porcentaje de pérdidas de pacientes. ¿Cómo cree que podría afectar este hecho a los resultados del ensayo?

- 1) La validez del estudio podría ser dudosa, ya que la pérdida de pacientes tal vez esté relacionada con los efectos de los tratamientos.

- 2) De ninguna forma, si en el diseño del estudio se habían previsto las pérdidas y, por tanto, se incluyó un número de pacientes mayor para compensarlas.
- 3) Aumentando la probabilidad de un resultado negativo falso al reducirse la muestra, pero sin que este hecho aumente las posibilidades de sesgos.
- 4) Si el estudio era a doble ciego, el hecho no tiene importancia, pero si el investigador conocía el tratamiento de cada paciente, entonces pudo haber sesgado los resultados.

RC: 1

### Los ensayos clínicos de fase II:

- 1) Se suelen realizar en grupos de pacientes no muy numerosos.
- 2) Regularmente son estudios observacionales.
- 3) Su objetivo principal es confirmar la efectividad de un medicamento en una patología concreta.
- 4) Su objetivo principal es evaluar la seguridad y tolerabilidad de un medicamento.

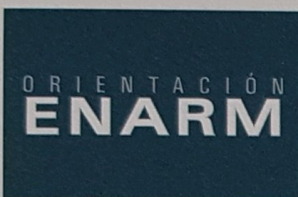
RC: 1

¿A qué tipo de ensayo clínico nos referimos cuando los criterios de inclusión se ajustan a las indicaciones, con criterios de exclusión menos restrictivos para incluir una amplia representación de la enfermedad en estudio y fundamentar así el registro de un medicamento?

- 1) Estudio en fase I.
- 2) Estudio en fase II.
- 3) Estudio en fase III.
- 4) Estudio en fase IV.

RC: 3

# Niveles de calidad de la evidencia científica



*Este es un tema importante y sencillo. Hay que conocer los diferentes conceptos, así como resolver tablas que incluyan los conceptos anteriores.*

## 9.1. Metanálisis

El metanálisis es un análisis estadístico que combina o integra los resultados de varios ensayos clínicos independientes que el analista considera que se pueden combinar. Añade a la mera revisión de los datos, un análisis estadístico de los mismos para dar un resultado conjunto de todos los estudios revisados. Es un estudio observacional (de ensayos clínicos), lo que implica que está sometido a los mismos sesgos de los citados estudios.

No puede tener mayor fortaleza metodológica que los estudios primarios. Si analiza ensayos clínicos (EC) de buena calidad, obtendrá resultados fiables y viceversa.

### Métodos estadísticos para la combinación de los resultados de diferentes estudios

Los métodos estadísticos más utilizados en la práctica pueden clasificarse en dos grupos, según se tenga en cuenta o no la heterogeneidad entre estudios en el análisis: los modelos de efectos fijos y los modelos de efectos aleatorios.

#### Modelos de efectos fijos

En el modelo de efectos fijos se asume que no existe heterogeneidad entre los estudios incluidos en la revisión, de modo que todos ellos estiman el mismo efecto, y las diferencias observadas se deben únicamente al azar. El efecto global puede estimarse como un promedio ponderado de los efectos individuales de cada estudio.

#### Modelos de efectos aleatorios

Por el contrario, con un modelo de efectos aleatorios se asume que los estudios incluidos en la revisión constituyen una muestra aleatoria de to-

dos los estudios existentes. Del mismo modo que en el modelo de efectos fijos, el efecto global se estima aquí como una media ponderada de los estimadores individuales.

Mientras que algunos autores defienden la utilización del modelo de efectos aleatorios en todos los casos, otros hacen hincapié en sus posibles deficiencias, como el hecho de que es menos preciso, proporcionando intervalos de confianza más amplios que el modelo de efectos fijos. Si las discrepancias son muy grandes, el modelo de efectos aleatorios se convierte en la alternativa al modelo más sencillo con efectos fijos para combinar los resultados. En el de efectos fijos, la generalización se limita a la población de estudios de características similares a los incluidos en el metanálisis. En el modelo de efectos aleatorios, por el contrario, los resultados pueden generalizarse a una población mayor de posibles estudios.

## 9.2. Factores determinantes de la calidad de la evidencia científica

Los aspectos del diseño de un estudio epidemiológico que están asociados a una mayor calidad y rigor científico son los siguientes:

- El sentido prospectivo del estudio (secuencia temporal).
- La asignación aleatoria a los grupos experimental y control.
- La existencia de un grupo control concurrente.
- El enmascaramiento de pacientes e investigadores participantes.
- La inclusión en el estudio de un número suficiente de pacientes como para detectar, en caso de haberlas, diferencias estadísticamente significativas entre los distintos grupos del estudio.

## 9.3. Escalas de gradación de la evidencia científica

Todas las escalas de gradación son muy similares entre sí y presentan al metanálisis de EC controlados y aleatorizados como el estudio más riguroso y que aporta mayor evidencia causal. La medicina basada en la evidencia persigue la integración de la maestría clínica individual con las mejores evidencias científicas disponibles (Tablas 1 y 2).

La calidad de la evidencia científica se clasifica en tres grados de recomendación para la adopción de una medida sanitaria:

- **Buena (grado A):** existe adecuada evidencia científica para recomendar o desaconsejar la adopción del procedimiento médico.
- **Regular (grado B):** hay cierta evidencia científica (no concluyente) para recomendar o desaconsejar la adopción del procedimiento médico.
- **Mala (grado C):** existe insuficiente evidencia científica para recomendar o desaconsejar la adopción del procedimiento médico.

Fuerza de la recomendación		
A	Existe adecuada evidencia científica para adoptar una práctica	I, II- 1
B	Existe cierta evidencia para recomendar la práctica	II- 1, II- 2
C	Hay insuficiente evidencia para recomendar o no recomendar la práctica	III
D	Existe cierta evidencia para no recomendar la práctica	II- 1, II- 2
E	Existe adecuada evidencia científica para no adoptar la práctica	I, II- 1

Tabla 1. Fuerza de recomendación clínica basada en los niveles de evidencia

Niveles de evidencia	
Grado I	Evidencia obtenida a partir de al menos un ensayo clínico aleatorizado y controlado bien diseñado
Grado II - 1	Evidencia obtenida a partir de ensayos clínicos no aleatorizados y bien diseñados
Grado II - 2	Evidencia obtenida a partir de estudios de cohortes (2a: prospectivo y 2b: retrospectivo)
Grado II - 3	Evidencia obtenida a partir de múltiples series comparadas en el tiempo, con o sin grupo control. Incluye resultados "dramáticos" producidos por experimentos no controlados (por ej., penicilina en los años cuarenta)
Grado III	Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de ética

Tabla 2. Gradación de la evidencia científica

### Ideas clave

➤ El metanálisis es un estudio en el que se realiza una combinación estadística de los resultados de ensayos previos similares.

➤ Si está realizado correctamente, es el estudio que aporta mayor evidencia científica.

➤ Una limitación importante se debe a que sólo se publican resultados favorables (sesgo de publicación).

# Estadística. Generalidades

ORIENTACIÓN  
**ENARM**

*En este tema sólo hay que saber dos cuestiones, qué son los tipos de muestreo y los tipos de variables.*

Se define la **estadística** como el método de razonamiento que permite interpretar un conjunto de datos cuyo carácter esencial es la variabilidad.

## 10.1. Tipos de estadística

Existen dos tipos de estadística:

- **Estadística descriptiva:** organización, presentación y síntesis de los datos de una manera científica y lógica.
- **Estadística inferencial:** bases lógicas mediante las cuales se establecen conclusiones relacionadas con poblaciones, a partir de los resultados obtenidos en muestras; trabaja con los datos que le proporciona la estadística descriptiva y es la parte de la estadística fundamental para la medicina clínica. Se basa en dos pilares básicos, que son el contraste de hipótesis y la estimación de intervalos de confianza.

## 10.2. Población y muestra

La **población** es el conjunto de todos los individuos, generalmente inaccesible, que tienen en común alguna característica observable y del que se pretende obtener una serie de conclusiones.

Se denomina **muestra** al conjunto menor de individuos, accesible y limitado, sobre el que se realiza el estudio con idea de obtener conclusiones generalizables a la población.

El proceso de selección de los individuos se puede realizar mediante distintas técnicas de muestreo. Los tipos de muestreos se detallan a continuación. Es importante conocer que el tipo de muestreo más importante es el muestreo probabilístico, en el que todos los elementos de la población tienen la misma probabilidad de ser incluidos en la muestra:

- **Muestreos probabilísticos:** en ellos se conoce la probabilidad de que un individuo sea elegido para la muestra. Son interesantes para usar estadística matemática con ellos. La elección de la muestra se hace aleatoriamente. La estadística inferencial exige este tipo de muestreo.
- **Muestreos no probabilísticos:** en ellos no se conoce la probabilidad. Son muestreos que muy posiblemente esconden sesgos. En principio, no se pueden extrapolar los resultados a la población.

Los muestreos con la menor posibilidad de sesgo (probabilísticos) son los siguientes: aleatorio simple, sistemático, estratificado y por grupos o conglomerados.

- **Muestreo aleatorio simple:** es aquel tipo de muestreo en el que cada individuo tiene las mismas posibilidades de ser elegido para formar parte de la muestra.
- **Muestreo aleatorio estratificado:** en él la población se divide en estratos. La muestra se obtiene seleccionado de cada estrato un número proporcional a la población que representa. Con este tipo de muestreo se asegura que la muestra tenga la misma proporción de una(s) variable(s) que la población de la que procede y que podría tratarse de una variable pronóstico diferente en los grupos de estudio.
- **Muestreo sistemático:** el proceso de selección empleado se basa en alguna regla sistemática simple, como por ejemplo, elegir uno de cada  $n$  individuos.
- **Muestreo por conglomerados:** en este tipo, la selección se realiza fundamentalmente con el objetivo de reducir costes, utilizando como base del muestreo al grupo de sujetos.

Por último, individuo será cada uno de los componentes de la población y de la muestra. Al número de individuos que forman la muestra se le llama tamaño muestral ( $n$ ).

## 10.3. Variables

Una variable es una característica observable que se desea estudiar en una muestra de individuos, pudiendo tomar diferentes valores. Los tipos de variables están recogidos en la **Tabla 1**.

Variable	Subtipo	Ejemplo
Cualitativa	<ul style="list-style-type: none"> <li>• Nominal dicotómica</li> <li>• Nominal no dicotómica</li> <li>• Ordinal</li> </ul>	<ul style="list-style-type: none"> <li>• Sexo</li> <li>• Raza</li> <li>• Nivel socioeconómico</li> </ul>
Cuantitativa	<ul style="list-style-type: none"> <li>• Discreta</li> <li>• Continua</li> </ul>	<ul style="list-style-type: none"> <li>• N.º episodios de asma/sem.</li> <li>• Uricemia</li> </ul>

Tabla 1. Tipos de variables

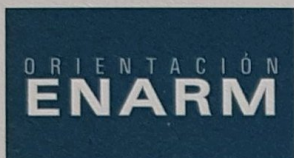
## Ideas clave

- La obtención de una muestra a partir de la población puede hacerse de forma probabilística (mediada por el azar) y no probabilística.
- Los tipos de muestreo probabilístico son el muestreo aleatorio simple, el estratificado, el sistemático y por conglomerados. El

estratificado debe realizarse cuando haya en la muestra una hipotética variable pronóstica que pudiera influir en el desarrollo de la enfermedad.

- Las variables son características observables. Pueden ser cuantitativas o cualitativas, en función de que la variable analizada pueda cuantificarse o expresarse en función de una característica.

# Estadística descriptiva



*Lo más importante de este tema es conocer la definición de los parámetros de tendencia central y de los de dispersión.*

Se puede considerar la estadística descriptiva como el conjunto de técnicas que facilitan la organización, resumen y comunicación de datos.

### 11.1. Representación gráfica

La representación gráfica de los datos obtenidos en cualquier estudio facilita un análisis visual. Según la naturaleza de las variables estudiadas, se utilizan diferentes tipos de representación.

#### VARIABLES CUALITATIVAS

Se emplean:

- Diagrama de barras (**Figura 1**).
- Diagrama sectorial (pastel) (**Figura 2**).

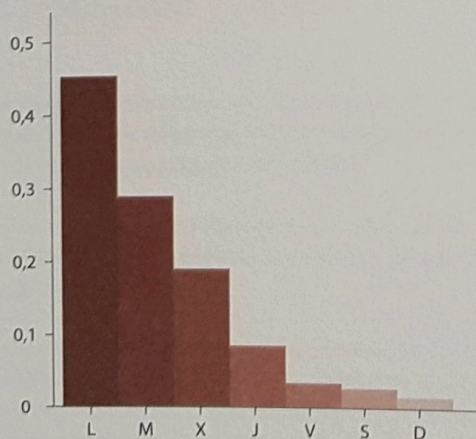


Figura 1. Diagrama de barras

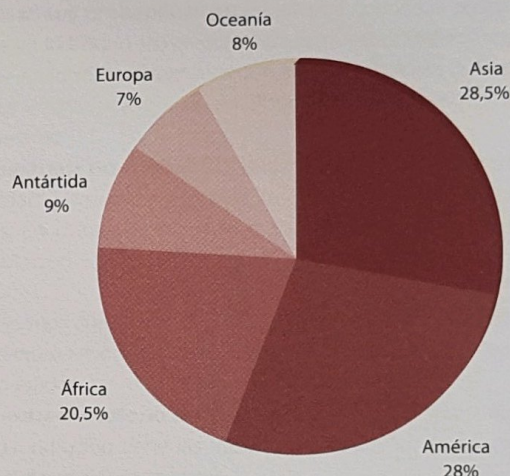


Figura 2. Diagrama sectorial

En ambos casos, se debe cumplir el principio de proporcionalidad de las áreas a las frecuencias absolutas. En los diagramas de sectores, el ángulo central es proporcional a la frecuencia absoluta correspondiente, por lo que también lo es su área. Los diagramas de barras tienen una base constante y una altura proporcional a la frecuencia absoluta correspondiente (también su área lo es a la frecuencia absoluta).

#### VARIABLES CUANTITATIVAS DISCRETAS

- **Diagrama de barras:** en un diagrama, sobre el valor que puede tomar la variable, se levanta una barra cuya altura mide exactamente la frecuencia absoluta del valor. También es posible representar la frecuencia relativa y/o los porcentajes.

#### VARIABLES CUANTITATIVAS CONTINUAS (Figura 3)

- **Histograma:** es un gráfico que está formado por rectángulos adyacentes que tienen por base cada uno de los intervalos y por altura, las frecuencias absolutas. La superficie de cada rectángulo es proporcional a la frecuencia de cada una de las clases, y el área total lo será al número de individuos en la muestra.
- **Polígono de frecuencias:** es una línea quebrada que une los puntos medios de las barras superiores de los rectángulos del histograma.

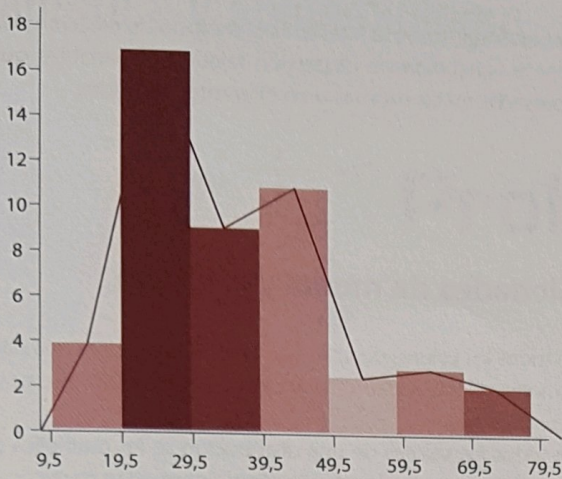


Figura 3. Histograma y polígono de frecuencias

## 11.2. Síntesis de los datos

### Medidas de localización o tendencia central

Las medidas de centralización indican alrededor de qué valores se agrupan los datos observados. Se distinguen las siguientes:

- **Media aritmética:** es la medida de centralización más común. Se calcula sumando los valores numéricos de todas las observaciones y dividiendo el total por el número de observaciones.

$$\bar{x} = \frac{\sum x_i}{n}$$

La media aritmética verifica la propiedad de equilibrar las desviaciones positivas y negativas de los datos respecto a su valor, es decir,  $\sum (x_i - \bar{x}) = 0$ . Actúa, por tanto, como centro geométrico o centro de "gravedad" para el conjunto de puntos.

- **Mediana:** es el valor numérico que divide al conjunto de datos ordenados en dos partes iguales, es decir, el 50% de los datos será menor que ella y el resto, mayor. En una distribución simétrica, la mediana coincide con la media aritmética, pero no así en una asimétrica.
- **Moda:** es el valor más corriente o el valor de la variable que se presenta con mayor frecuencia. Pueden existir distribuciones con más de una moda.
- **Parámetros de posición:** cuartiles, deciles, percentiles. Son valores que dividen el conjunto de las observaciones en cuatro, diez o cien partes iguales, respectivamente. Por ejemplo:

$$Q2 = D5 = Pc50 = \text{mediana}$$

El 50% de las observaciones será inferior al segundo cuartil, quinto decil o percentil 50.

- **Media geométrica:** muy utilizada en Microbiología y Serología, cuyos datos tienen una marcada asimetría positiva (hacia la derecha). Por ejemplo: títulos de anticuerpos.

### Comparación de medidas de centralización

Aunque desde un punto de vista puramente descriptivo, estas medidas proporcionan información complementaria, sus propiedades son muy

distintas: la media utiliza todos los datos y es, por tanto, preferible si los datos son homogéneos; tiene el inconveniente de que es muy sensible a observaciones atípicas, y un error de datos o un valor anormal puede modificarla totalmente.

Por el contrario, la mediana emplea menos información que la media, ya que sólo tiene en cuenta el orden de los datos y no su magnitud, pero, en contrapartida, no se ve alterada si una observación (o, en general, una pequeña parte de observaciones) es extrema o contiene errores grandes de medida o de transcripción.

En consecuencia, es siempre recomendable calcular la media y la mediana, puesto que ambas medidas diferirán mucho cuando la distribución sea muy asimétrica (Figura 4), lo que sugiere heterogeneidad en los datos.

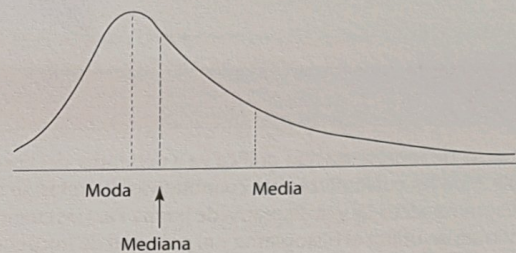


Figura 4. Distribución asimétrica

### Medidas de dispersión o variabilidad

Junto a las medidas de tendencia central, las medidas de dispersión o variabilidad completan la información sobre la distribución de la variable (indican si los valores de la misma están muy dispersos o se concentran alrededor de la medida de centralización) (Tabla 1).

	Medida de centralización	Medida de dispersión
Distribución homogénea	Media	Desviación típica o estándar
Distribución asimétrica	Mediana	Rango intercuartílico

Tabla 1. Formas de medida

- **Rango o recorrido:** diferencia entre el valor máximo y el mínimo observado en una serie.

$$R = \text{Máx.} - \text{Mín.}$$

- **Desviación media:** es la media de las desviaciones respecto a la media aritmética.

$$Dm = \frac{\sum |x_i - \bar{x}|}{n}$$

- **Varianza:** se define como la media del cuadrado de las desviaciones de los elementos respecto a la media aritmética.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

- **Desviación típica o estándar:** es la raíz cuadrada positiva de la varianza. Es, junto con esta, la medida de dispersión más utilizada.

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

La desviación típica es una medida complementaria de la media aritmética; mientras que esta da una idea de la magnitud general de la distribución, la desviación estándar muestra cómo se distribuyen los valores alrededor de la media.

- **Rango intercuartílico:** es la diferencia entre el percentil 75 y el 25. Es, junto con el rango, la medida de dispersión usada para los datos asimétricos.
- **Coefficiente de variación (CV):** es una medida de dispersión adimensional. Se define como el porcentaje que representa la desvia-

ción estándar sobre la media. Es el método de elección para comparar la variabilidad o dispersión relativa de variables que estén expresadas en las mismas o en diferentes unidades.

$$CV = \frac{s}{\bar{x}} \times 100$$

## Propiedades de media y varianza

- Si a todos los valores de una distribución se les suma una constante, su media queda aumentada en ese valor, mientras que su varianza no se modifica.
- Si a todos los valores de una distribución se les multiplica por una constante, su media y su desviación típica quedan multiplicadas por la constante, mientras que su varianza queda multiplicada por el cuadrado de esa constante.

## Ideas clave

- La forma de representación gráfica va a depender del tipo de variable. Para las cualitativas y las cuantitativas discretas, se emplea el diagrama sectorial y el diagrama de barras. Para las cuantitativas continuas se utiliza el histograma y el polígono de frecuencias.
- Para definir correctamente una variable en una muestra se necesitará un parámetro que sirva para aglutinar todos los valores de dicha muestra (parámetro de tendencia central) y otro que sirva para informar de lo agregados o no que están los valores (parámetro de dispersión).
- El tipo de distribución más frecuente es la normal o de Gauss, que es simétrica. En distribuciones simétricas se usará como pa-

rámetro de tendencia central la media y de dispersión, la desviación típica.

- En distribuciones asimétricas, se emplea la mediana como parámetro de tendencia central y el rango como parámetro de dispersión.
- En la distribución de Gauss, media, mediana y moda coinciden en el mismo valor.
- El coeficiente de variación es el parámetro que se emplea para decir si una distribución es homogénea o dispersa. Además, informa de cuál de entre dos distribuciones muestra una mayor variabilidad (dispersión).

## Casos clínicos

Un niño de 5 años pesa 20 kg. Según las tablas de Tanner-Whitehouse, este valor corresponde aproximadamente al percentil 75 de la distribución del peso en niños de dicha edad. ¿Qué significa esta expresión?

- 1) El 75% de los niños de 5 años pesan aproximadamente 20 kg.
- 2) El niño pesa un 75% más que el promedio de los niños de su edad.
- 3) El 75% de los niños de 5 años pesan menos de 20 kg.
- 4) El niño tiene un peso que es un 75% superior a la mediana del peso de los niños de su edad.

RC: 3

En un estudio en el cual se ha medido la intensidad de dolor en un grupo de 145 pacientes con artritis reumatoide mediante una escala de 0 (ausencia de dolor) a 10 (dolor de intensidad máxima), se informa de que la mediana es de 6. ¿Cuál el significado de este valor?

- 1) La mitad de los sujetos de la muestra tienen valores de intensidad de dolor iguales o inferiores a 6.
- 2) El valor 6 indica la intensidad de dolor que puede considerarse normal en la escala utilizada.
- 3) El valor 6 de la escala ha sido la puntuación obtenida con mayor frecuencia por los sujetos de la muestra.
- 4) El valor 6 es la media aritmética de las puntuaciones obtenidas por los sujetos de la muestra.

RC: 1



# Probabilidad

ORIENTACIÓN  
**ENARM**

*Todo el tema se puede resumir en dos fórmulas, que son la de la suma y la de la multiplicación.*

## 12.1. Probabilidad

Se supone una población finita con  $N$  elementos,  $k$  de los cuales tienen la característica  $A$ . Se conoce como probabilidad de la característica  $A$  en la población a:  $P(A) = k/N$ . La probabilidad se define, desde un punto de vista óptimo, en términos de frecuencia relativa.

$$P(A) = \frac{\text{Número de veces que ocurre } A}{\text{Número de veces que puede ocurrir } A}$$

De esta forma:

$$P(A) = \frac{\text{Casos favorables}}{\text{Casos posibles}}$$

### Propiedades (Figura 1)

- La probabilidad es un valor entre 0 y 1, es decir,  $0 \leq P(A) \leq 1$ .
- Se llama suceso seguro  $E$  al que ocurre siempre, entonces:  $P(E) = 1$ .

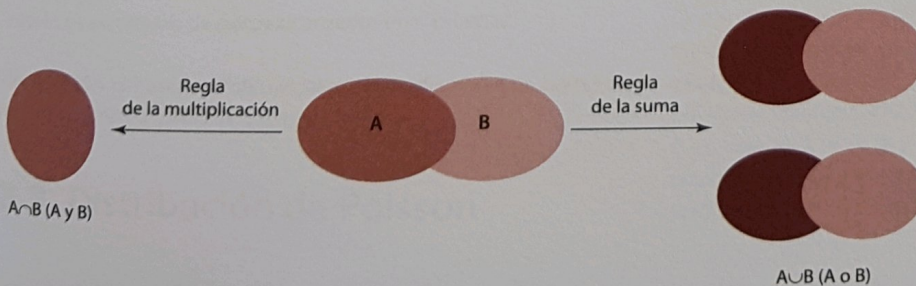


Figura 1. Reglas de la probabilidad

- Si  $\bar{A}$  es el suceso complementario de  $A$ , que sucede siempre que no lo hace  $A$ , entonces:  $P(\bar{A}) = 1 - P(A)$ .
- Dos sucesos son mutuamente excluyentes (incompatibles) cuando no pueden ocurrir a la vez (por ej., ser mujer y tener cáncer de próstata).

$$P(A \cap B) = P(A \text{ y } B) = 0$$

- Dos sucesos son independientes cuando la probabilidad de aparición de uno de ellos no está relacionada o no se modifica por la aparición del otro (por ej., efectos secundarios de un fármaco en dos pacientes).

$$P(A \cap B) = P(A \text{ y } B) = P(A) \times P(B)$$

## 12.2. Probabilidad condicionada

La probabilidad de  $A$  condicionada a la ocurrencia de  $B$  se define considerando únicamente los casos en los que aparece  $B$ , y viendo en cuántos de estos casos tiene lugar el suceso  $A$ . Es la probabilidad de que suceda  $A$ , una vez que ya ha sucedido  $B$ .

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B/A)}{P(B)}$$

## 12.3. Regla de la multiplicación

Es la probabilidad de que sucedan  $A$  y  $B$ .

$$P(A \cap B) = P(A/B) \times P(B)$$

Si dos sucesos son independientes, entonces:

$$P(A \cap B) = P(A) \times P(B)$$

## 12.4. Regla de la suma

Si dos sucesos son mutuamente excluyentes (o bien se da A o bien ocurre B), entonces:

$$P(A \cup B) = P(A) + P(B)$$

Esta regla nos indica la probabilidad de que suceda A o de que suceda B:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

### Ideas clave

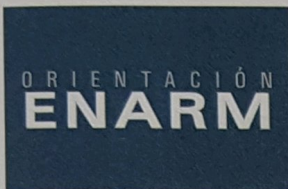
➤ La fórmula de la suma se debe aplicar cuando se pregunta cuál es la probabilidad de que ocurra un evento u otro (por ejemplo, probabilidad de que un sujeto tenga cáncer de pulmón o de laringe).

➤ La multiplicación se aplica cuando se pregunta por la probabilidad de que ocurra un evento y el otro (probabilidad de que un sujeto tenga cáncer de pulmón y de laringe).

# Principales leyes de distribución de variables aleatorias



Haz clic en la imagen para abrir tu plataforma



*Se debe saber interpretar los intervalos de una muestra a partir de la curva de Gauss y manejar perfectamente esta distribución.*

## 13.1. Principales leyes de distribución de variables aleatorias

Una vez que se ha elegido la muestra y recogido los datos, el siguiente paso es inferir las propiedades de la población a partir de la muestra. El instrumento conceptual que permitirá esta generalización es un modelo de la población, es decir, una representación simbólica de su comportamiento.

Existen tres leyes teóricas que tienen la particularidad de ser seguidas por una inmensa mayoría de los fenómenos biológicos habituales; estas leyes son las siguientes:

- Distribución binomial (distribuciones discretas).
- Distribución de Poisson o de los sucesos raros (distribuciones discretas).
- Distribución normal o de Gauss (distribuciones continuas).

## 13.2. Distribución binomial

Es una ley teórica de aplicación siempre que se conozca, de entrada, la probabilidad de aparición de un fenómeno biológico ( $p$ ). El resultado es dicotómico y puede ser evaluado como "éxito o con fracaso". La variable de interés es el número de éxitos alcanzados en  $n$  pruebas.

La distribución binomial constituye siempre un diagrama de barras discontinuo se aplica a variables discretas.

## 13.3. Distribución de Poisson

En general, se empleará la distribución de Poisson como aproximación de experimentos binomiales donde el número de pruebas es muy alto, pero

la probabilidad de éxito muy baja. Es, por consiguiente, la ley adecuada para los "sucesos raros", cuando  $p < 0,1$  y  $n \times p < 5$ .

## 13.4. Distribución normal o de Gauss

Este tipo de distribución es seguida por una inmensa cantidad de variables biológicas cuyas medidas se agrupan alrededor de un valor central, y que presentan una frecuencia cada vez menor a medida que se alejan de dicho valor.

### Características

Las características de la distribución normal o de Gauss son las siguientes:

- Corresponde a variables cuantitativas continuas.
- Se caracteriza por dos medidas: media y desviación típica.
- Es unimodal.
- Es simétrica alrededor de la media. Por tanto, media, mediana y moda coinciden.
- Tiene forma acampanada, sin un pico excesivo.
- Va desde  $-\infty$  a  $+\infty$  (asintótica al eje de abscisas).
- El área bajo la curva tiene un valor igual a 1.

El 50% de las observaciones se encuentran por debajo de la media aritmética y el 50% por encima. El 68% de las mismas están dentro del intervalo  $x \pm s$ ; el 95% dentro del intervalo  $x \pm 1,96s$  y el 99% dentro del intervalo  $x \pm 2,57s$ .

### Tipificación

La distribución normal viene definida por la media y por la desviación típica; pueden existir, por tanto, infinitas curvas de este tipo, tantas como valores puedan tomar la media y la desviación típica (**Figura 1**). La distribución normal tipificada tiene una media igual a 0 y una desviación típica cuyo valor es 1  $[N(0,1)]$  y se encuentra tabulada. Es posible transformar cualquier variable aleatoria que se distribuya normalmente a una normal tipificada mediante la siguiente ecuación:

$$Z = \frac{x_i - \bar{x}}{s}$$

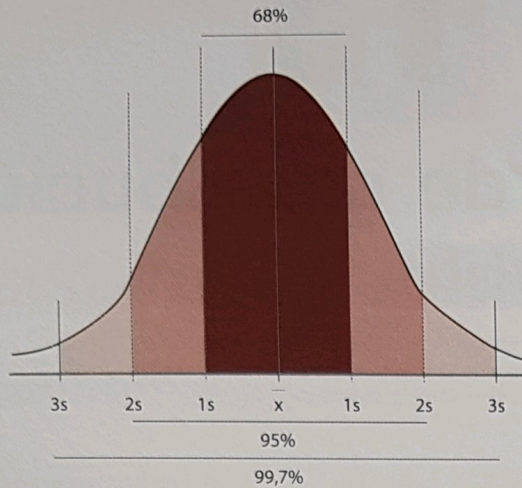


Figura 1. Distribución normal

## Ideas clave

- La distribución normal o de Gauss es una curva de probabilidad continua con forma de campana.
- Los tres parámetros de tendencia central: media, mediana y moda, coinciden en el mismo valor.
- Los límites de dicha campana son  $\pm \infty$  y su área es 1.
- Siempre se cumple que en el intervalo media  $\pm 1,96 s$  está incluido el 95% de la muestra.

## Casos clínicos

En una muestra aleatoria extraída de población sana se encuentra que una variable bioquímica tiene como media 90 y como desviación típica 10. La afirmación: "aproximadamente el 95% de los individuos sanos tiene un valor de esa variable comprendido entre 70 y 110" es correcta:

- 1) Siempre.
- 2) Nunca.
- 3) Sólo si la variable tiene distribución normal.
- 4) Sólo si la variable tiene distribución normal y la muestra es suficientemente grande.

RC: 4

En un estudio sobre 275 mujeres premenopáusicas, se informa de que su índice de masa corporal (IMC) sigue una distribución normal, con una media de 27 y una desviación típica de 4. ¿Cuál de las siguientes afirmaciones es cierta?

- 1) El rango de valores de IMC observado en las 275 mujeres va desde 23 a 31.

- 2) El rango de valores de IMC observado en las 275 mujeres va desde 19 a 35.
- 3) Aproximadamente, el 95% de las mujeres estudiadas tienen valores de IMC situados entre 19 y 35.
- 4) Aproximadamente, el 95% de las mujeres estudiadas tienen valores de IMC situados entre 23 y 31.

RC: 3

Un estudio realizado sobre una muestra de 200 sujetos informa de que su edad media es de 45 años, con una desviación estándar de 8 años. ¿Cuál de las siguientes afirmaciones es correcta?

- 1) La mayoría de sujetos tienen entre 37 y 53 años.
- 2) Aproximadamente, el 95% de los sujetos tienen entre 29 y 61 años.
- 3) Se tiene un 95% de confianza de que la verdadera edad media se sitúe entre 29 y 61 años.
- 4) Se tiene un 95% de confianza de que la verdadera edad media se sitúe entre 37 y 53 años.

RC: 2

# Estimación de parámetros. Estadística inferencial

ORIENTACIÓN  
**ENARM**

*Para abordar con garantías de éxito este tema se debe conocer muy bien el tema anterior.*

La tarea fundamental de la estadística inferencial es hacer inferencias acerca de la población a partir de una muestra extraída de la misma.

## 14.1. Estimación de medias

En este caso, se quiere conocer la media poblacional a partir de la media muestral. Para ello, se pueden aplicar:

- **Estimación puntual:** la media muestral es un buen estimador de la media poblacional ( $\mu$ ).
- **Estimación por intervalos:** el valor numérico de los estimadores puntuales determinados en muestras diferentes puede variar. Es decir, si se repiten los cálculos con otra muestra distinta de la misma población, el resultado de la media muestral podría ser diferente. Por tanto, sería mejor que, además, se aportase un intervalo que presumiblemente incluya también el parámetro de la población. Por ello, es preferible la estimación por intervalos, ya que entonces se indican límites de valores dentro de los cuales el parámetro poblacional tiene la probabilidad de estar. Al intervalo alrededor del estadístico muestral se le denomina intervalo de confianza, y a sus límites, límites de confianza. El cálculo de los límites de confianza comprende el empleo del error estándar de la media y los principios de la distribución normal.

El nivel de confianza  $1 - \alpha$  da la probabilidad que hay de que la media poblacional,  $\mu$ , pertenezca a ese intervalo; es frecuente expresarlo en porcentajes. Son equivalentes las expresiones "nivel de confianza  $1 - \alpha$ " y "significación  $\alpha$ "; así, suele hablarse de obtener un intervalo con un nivel de confianza del 95% o para una significación  $\alpha = 0,05$  (este valor de  $\alpha$  indica la probabilidad que hay de error en la estimación. A veces esa probabilidad se indica con la letra  $P$ ).

### Error estándar de la media

De una población pueden extraerse infinitas muestras, cada una de ellas con su media. Este conjunto de medias se distribuye según una curva

normal cuya media, la media de las medias, es la media poblacional ( $\mu$ ) y cuya desviación típica se conoce como "error estándar de la media"; por tanto, es la dispersión de las medias muestrales con respecto a la media poblacional. Se calcula mediante la fórmula (Figura 1):

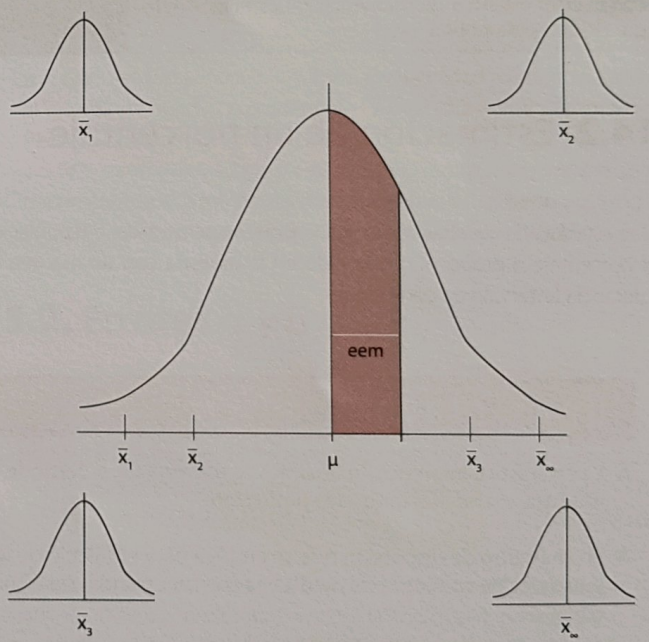


Figura 1. Cálculo del error estándar de la media

$$eem = \frac{\sigma}{\sqrt{n-1}}$$

$\sigma$  = desviación típica poblacional

Si se construye un intervalo de confianza del 95%, se concluye que en el 95% de las ocasiones la media poblacional estará entre:

$$\bar{x} \pm 1,96 \times \frac{\sigma}{\sqrt{n-1}}$$

Mientras que en un 5% de las ocasiones estará equivocado. Si se establece el nivel de confianza en el 99%, la media poblacional se situará en un 99% de las ocasiones entre:

$$\bar{x} \pm 2,57 \times \frac{\sigma}{\sqrt{n-1}}$$

En el caso de que se desconozca el valor del parámetro  $\sigma$  (que es lo más habitual en la práctica) o cuando  $n < 30$ , los límites de confianza se calcularían siguiendo:

$$\bar{x} \pm t \times \frac{s}{\sqrt{n-1}}$$

Donde el valor  $t$  corresponde a los valores tabulados de la distribución teórica  $t$  de Student.

No tiene nada que ver la desviación típica de los valores de la variable de la muestra o en la población con el error estándar de las medias muestrales alrededor de la media poblacional.



## 14.2. Estimación de un porcentaje

Si se estudia una variable cualitativa, el objetivo será aplicar a la población el porcentaje o proporción obtenido en la muestra con un margen de tolerancia (intervalo de confianza).

### Ideas clave

- La estadística inferencial o analítica se encarga de, a partir de la muestra, extrapolar datos a la población.
- El parámetro de dispersión que se emplea para el cálculo de los intervalos de confianza de medias es el error estándar de la media.

### Casos clínicos

Si la media es 25, la desviación estándar es 5 y el tamaño muestral es de 101, ¿cuál será el error estándar de la media?

- 1) 5/5.
- 2) 5/10.
- 3) 10/5.
- 4) 0,05.

RC: 2

En caso de muestras pequeñas ( $n < 100$ ), la estimación de un porcentaje poblacional sigue una distribución binomial y su cálculo está tabulado.

En caso de muestras grandes ( $n > 100$ ) y si  $n \times p > 5$  y  $n \times (1 - p) > 5$ , la distribución binomial se puede aproximar a la distribución normal y la estimación del porcentaje poblacional sería:

$$p \pm 1,96 \times eep, p < 0,05$$

$$p \pm 2,57 \times eep, p < 0,01$$

donde:

- $p$ : proporción (en tanto por uno) hallada en la muestra.
- $eep$ : error estándar del porcentaje.

$$eep = \sqrt{\frac{p(1-p)}{n}}$$

## 14.3. Estimación de medidas de asociación

Independientemente del tipo de estudio epidemiológico y del parámetro utilizado para medir la fuerza de asociación (RR, OR, RP), la estimación de las medidas de asociación ha de ser extrapolada a la población general mediante la construcción de un intervalo de confianza.

Si el IC incluye el 1 (efecto neutro del factor de estudio), se habla de un resultado no significativo.

- La curva de probabilidades de los intervalos de confianza sigue otra curva de Gauss, por lo que se puede decir que, en el intervalo media  $\pm 2$  eem, estará incluida la media poblacional con una probabilidad del 95%.
- Los intervalos así obtenidos se denominan intervalos de confianza del 95%.
- El 5% restante es la probabilidad de que el intervalo esté mal estimado. A este error se le llama  $\alpha$ .

Si la media es 25, la desviación estándar es 5 y el tamaño muestral es de 101. ¿Cuál será el intervalo que incluya el verdadero valor poblacional con una probabilidad del 95%?

- 1) 23-27.
- 2) 23,5-26,5.
- 3) 24-26.
- 4) 25.

RC: 3

# Contraste de hipótesis

ORIENTACIÓN  
**ENARM**

*Lo que hay que tener claro de este tema es el significado de los errores  $\alpha$  y  $\beta$ , y del poder de un test estadístico y sobre cómo se modifican los parámetros anteriores en función del tamaño de la muestra.*

## 15.1. Hipótesis nula, hipótesis alternativa y grado de significación estadística

Para la comprensión de este concepto, se partirá de un caso práctico. Se compara un tratamiento nuevo (N) contra la HTA con uno tradicional (T) en dos series de individuos. Tras cierto tiempo, N controla al 70% de los individuos y T al 30%. El problema consiste en saber si este hecho es debido a una mayor eficacia real de N o lo es simplemente por diferencias que el azar puede establecer (variaciones aleatorias del muestreo), sin que existan diferencias reales en el control de hipertensos entre ambos tratamientos.

Para abordar este problema, se consideran dos hipótesis:

- **Hipótesis nula ( $H_0$ ):** no existen diferencias entre los dos tratamientos ( $N = T$ ).
- **Hipótesis alternativa ( $H_1$ ):** sí existen diferencias entre los dos tratamientos ( $N \neq T$ ).

Estas dos hipótesis son mutuamente excluyentes, de forma que sólo hay dos decisiones posibles:

- Rechazar  $H_0 \rightarrow$  aceptar  $H_1$ .
- No rechazar  $H_0 \rightarrow$  no poder aceptar  $H_1$ .

Previamente al ensayo de una hipótesis, se fija la probabilidad máxima de que los resultados diferentes observados entre los dos grupos puedan ser debidos simplemente al azar ( $H_0$  cierta), que suele ser por convenio del 5%. A continuación, se calcula cuál es la probabilidad de que las diferencias que se han observado puedan ser explicadas por azar. Esta probabilidad es el valor de la  $p$  o "grado de significación estadística". Así, cuanto menor sea  $p$ , es decir, cuanto más baja sea la probabilidad de que el azar sea el responsable de las diferencias, mayor será la evidencia contra  $H_0$  y a favor de  $H_1$ .

De tal forma, si:

- $p < 0,05 \rightarrow$  diferencias reales. Poca probabilidad de que se deban al azar. Se acepta  $H_1$ . El resultado es estadísticamente significativo. Rechazo de  $H_0$ .
- $p > 0,05 \rightarrow$  no existe suficiente evidencia como para decir que ambos tratamientos son distintos. Las diferencias pueden deberse al azar, con una probabilidad mayor al nivel de exigencia. No se rechaza  $H_0$ .

## 15.2. Errores $\alpha$ y $\beta$

- **Error tipo I:** se rechaza  $H_0$  siendo cierta (se acepta que hay diferencias y, de hecho, no las hay). Es como un falso positivo: dar como significativo algo que no lo es. Se denomina  $\alpha$  a la probabilidad de cometer el error tipo I. El valor de  $p$  coincide con la probabilidad de cometer el error tipo I.

Cuando entre los distintos grupos de estudio se compara más de una variable de resultado, aumenta la posibilidad de resultados falsamente positivos. Para evitarlo, se incrementa la exigencia del nivel de significación como sigue :

$$p = \frac{0,05}{n.^{\circ} \text{ de comparaciones}}$$

- **Error tipo II:** no se rechaza  $H_0$  y ésta es falsa (no se aceptan las diferencias y sí las hay). Sería como un falso negativo: se da como no significativo algo que lo es. Se denomina  $\beta$  a la probabilidad de cometer un error tipo II.
- **Poder o potencia del test:** lo complementario del error  $\beta$  es la "potencia o poder estadístico de un test" ( $1 - \beta$ ) que se define como la capacidad que tiene un test de detectar una diferencia cuando esta existe en realidad; es decir, corresponde a la probabilidad que se tiene de demostrar la hipótesis alternativa, siendo cierta (Tabla 1).

		Realidad	
		Existe diferencia $H_0$ falsa	No existe diferencia $H_0$ cierta
Resultados del test	Hay diferencias significativas Rechazo $H_0$	$1 - \beta$ Poder estadístico o potencia del test	Error tipo I o error $\alpha$
	No hay diferencias significativas No rechazo $H_0$	Error tipo II o error $\beta$	$1 - \alpha$

Tabla 1. Contraste de hipótesis

### 15.3. Pruebas de significación estadística (Figura 1)

- **Asociación estadística entre dos variables:** el objetivo es demostrar o no la asociación estadística entre dos variables observadas en una muestra.

**Datos apareados:** en una única muestra se estudia si existe una diferencia significativa en la variable resultado antes y después de algún acontecimiento relevante.

- Variable resultado cualitativa de dos o más categorías: test de McNemar.
- Variable resultado cuantitativa: t de Student apareada.
- Variable resultado cuantitativa (varias medidas repetidas): ANOVA para medidas repetidas.

- **Test no paramétricos:** son pruebas que no dependen de la distribución de la variable y no precisan condición previa de normalidad (Figura 2). Estas pruebas no manejan los valores cuantitativos que toma la variable en cuestión, sino únicamente sus rangos. Son pruebas de menor precisión que las paramétricas, por lo que solamente demostrarán como significativas diferencias mayores. Por tanto, generalmente se prefiere utilizar pruebas paramétricas (t de Student, análisis de la varianza, etc.) y el uso de los test no paramétricos se limita a estudios con tamaños muestrales menores de 30, en los que las poblaciones no puedan ser consideradas como normales, o bien cuando la variable represente únicamente una medida aproximada del carácter, sin una valoración exacta. Si las muestras son mayores de 30, no existe inconveniente en utilizar pruebas paramétricas.

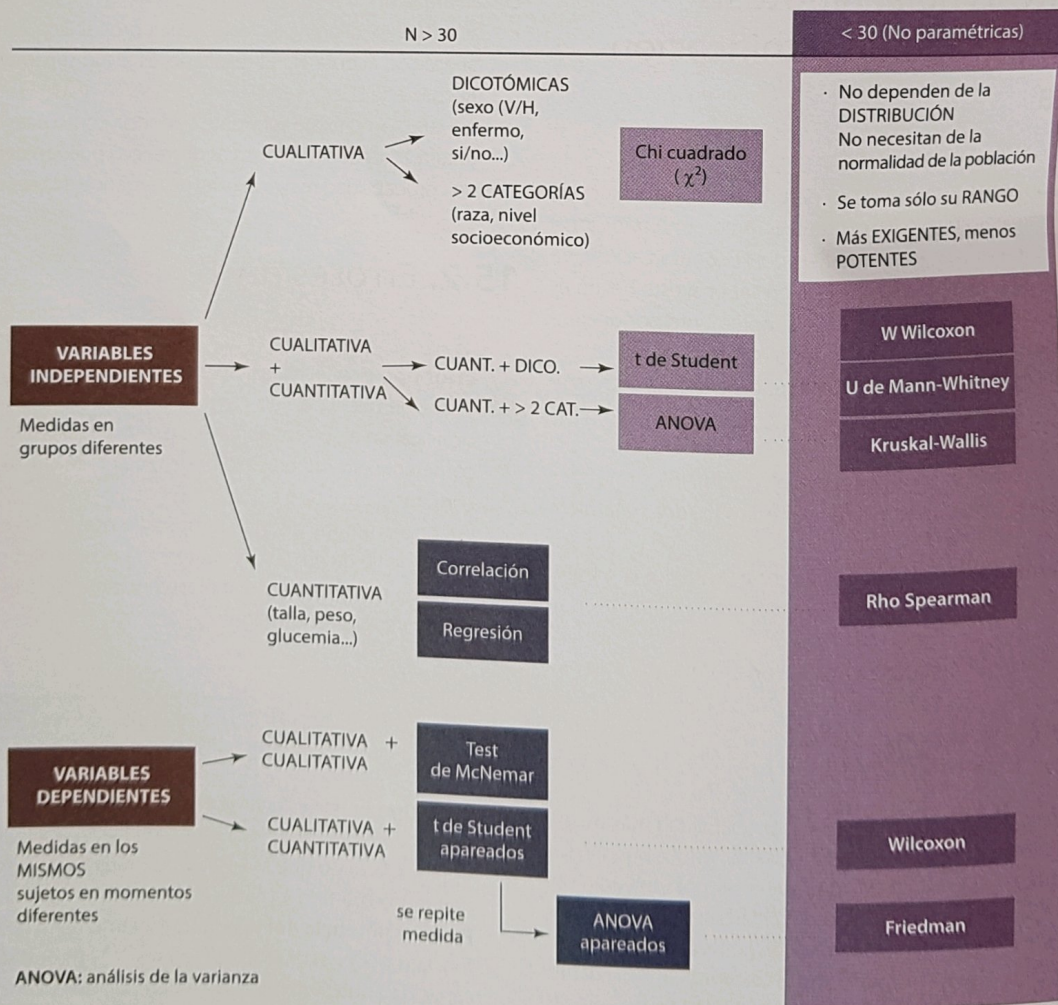


Figura 1. Pruebas de asociación estadística



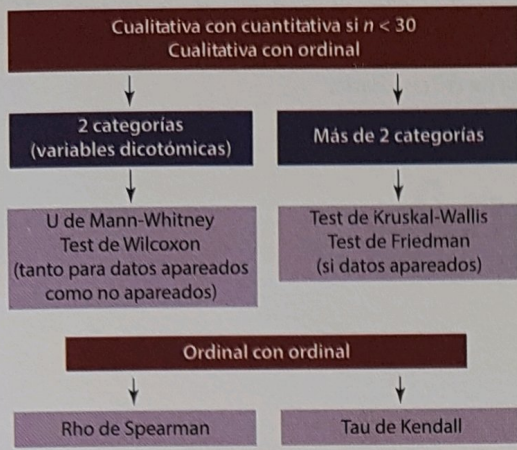


Figura 2. Test no paramétricos

- **Análisis multivariante:** es un conjunto de test estadísticos que se aplica cuando se intenta establecer la asociación estadística entre dos variables, pero controlando el efecto de terceros factores que podrían actuar como factores de confusión. Un tipo de test multivariante frecuentemente empleado es el test de regresión logística. Por ejemplo: se desea saber si existe una relación entre la infección de herida quirúrgica y el tipo de cirugía, controlando que la profilaxis antibiótica empleada no actúe como factor confusor.
- **Análisis de supervivencia:** en ocasiones, para saber si por ejemplo un fármaco es mejor que otro se necesita conocer la idea dinámica de cómo fallecen los pacientes incluidos en uno u otro grupo de tratamiento a lo largo del tiempo. Es decir, no sólo importa si fallecen o no, sino el tiempo en el que se produce el evento. Esta idea es extrapolable a otras situaciones como por ejemplo, tiempo hasta que se produce la curación, tiempo hasta que des-

aparecen los síntomas... Uno de los análisis estadísticos más frecuentemente empleados es el de Kaplan-Meier. La característica distintiva del análisis con este método es que la proporción acumulada que sobrevive se calcula para el tiempo de supervivencia individual de cada paciente. Para saber si hay diferencias en el tiempo en el que se produce el evento para los dos grupos de tratamiento hay que usar test específicos para tal fin como son el test de Breslow o el Logrank.

La Figura 3 resume los test estadísticos más empleados.

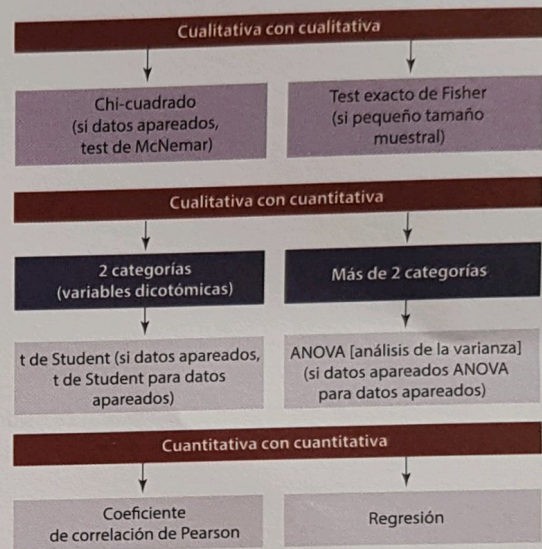


Figura 3. Test estadísticos más empleados

## Ideas clave

- El contraste de hipótesis se basa en intentar rechazar  $H_0$  (no hay diferencias) para así aceptar  $H_1$  (sí existen diferencias).
- Cualquier resultado observado en un estudio puede atribuirse al azar. El valor de  $p$ , o nivel de significación, se refiere a la probabilidad de que el azar explique un resultado obtenido.
- Si la  $p < 0,05$ , se admite poca probabilidad de azar y entonces se habla de que hay diferencias estadísticamente significativas.
- Si la  $p > 0,05$ , la probabilidad de azar es demasiado grande como para que se afirme que hay diferencias significativas en los resultados obtenidos.
- La probabilidad que existe de que yo afirme que hay diferencias, cuando en realidad no las hay, equivale a la probabilidad  $p$  y se llama error  $\alpha$  o tipo I.
- Por el contrario, si verdaderamente existen diferencias y no se detectan, se estará incurriendo de nuevo en un error, que en este caso es el  $\beta$  o tipo II.

- Ambos errores  $\alpha$  y  $\beta$  son dos formas diferentes de errores aleatorios, por lo que se podrán reducir si se aumenta el tamaño de la muestra.
- Si verdaderamente hay diferencias reales en los resultados obtenidos y estas diferencias son detectadas por el test, entonces se está en la situación de acierto que se conoce como poder o potencia del test.
- Los test estadísticos que se van a emplear dependerán del tipo de variable que se haya utilizado:
  - Cuali + cuali: Chi cuadrado.
  - Cuanti + cuali dicotómica: t de Student (equivalente no paramétrico: Wilcoxon, U de Mann-Whitney).
  - Cuanti + cuali no dicotómica: ANOVA (equivalente no paramétrico: Kruskal-Wallis, Friedman).
  - Cuanti + cuanti: correlación de Pearson (equivalente no paramétrico: Spearman).
- Los test paramétricos son mejores, pero sólo pueden realizarse cuando la distribución de estudio sea la normal.

## Casos clínicos

Un investigador está interesado en determinar si existe una asociación entre las cifras de tensión arterial diastólica (medida en mm de Hg) y los niveles de colesterol (medidos en mg/ml). Para ello, ha realizado estas mediciones a 230 voluntarios. ¿Qué prueba estadística es la más apropiada para examinar dicha asociación?

- 1) Regresión logística.
- 2) Prueba de la t.
- 3) Prueba de chi cuadrado.
- 4) Correlación de Pearson.

RC: 4

# Análisis de correlación y de regresión

ORIENTACIÓN  
**ENARM**

Los aspectos más importantes de este tema son la interpretación correcta del coeficiente  $r$  y de los valores  $\alpha$  y  $\beta$  de la regresión lineal.

## 16.1. Análisis de correlación y de regresión

Los análisis de correlación y de regresión sirven para estudiar la relación entre dos variables cuantitativas.

- **Análisis de regresión:** es útil para determinar la posible forma de relación entre variables  $y$ , por tanto, se puede usar para hacer predicciones o para estimar el valor de una variable que corresponde para un valor de la otra.

En este análisis se dispone de dos variables de interés,  $x$  e  $y$ . La variable  $x$  se llama variable independiente o explicativa y es controlada por el investigador. La variable  $y$  se denomina dependiente.

Con el análisis de regresión lineal, se obtiene una ecuación del tipo:

$$y = \alpha + \beta x$$

donde  $\alpha$  y  $\beta$  son los coeficientes de regresión.

El coeficiente  $\alpha$  representa el punto en el que la línea corta el eje vertical (valor de  $y$  para  $x = 0$ ). El coeficiente  $\beta$  es la pendiente de la recta que muestra la cantidad que varía  $y$  por una unidad de cambio de  $x$ .

La recta de regresión es más potente para hacer predicciones cuando el coeficiente de correlación ( $r$ ) está próximo a  $+1$  o a  $-1$ .

También es más eficaz cuando la predicción se obtiene a partir de valores de  $x$  próximos a la media (Figura 1).

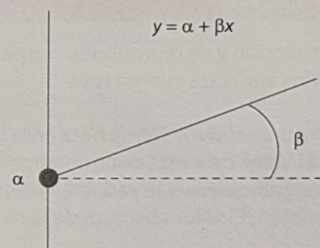


Figura 1. Recta de regresión lineal

- **Análisis de correlación:** estudia también la relación entre dos variables cuantitativas, aunque aquí no se habla de variable dependiente ni independiente. El coeficiente de correlación de Pearson mide la intensidad de la relación lineal entre las dos variables cuantitativas.

Las características del coeficiente de correlación son las siguientes:

- Varía entre  $-1$  y  $+1 \rightarrow -1 \leq r \leq +1$ .
- Para  $r = \pm 1$ , hay una relación perfecta entre  $x$  e  $y$ , es decir, todos los puntos  $(x,y)$  están en una línea recta.
- Un valor positivo de  $r$  indica que a medida que aumenta una variable, lo hace la otra, o que conforme disminuye una, también lo hace la otra. Un coeficiente de correlación negativo denota que a medida que disminuye una variable, aumenta la otra o viceversa.
- $r = 0$  indica que no hay correlación lineal.

La Figura 2 resume los distintos tipos de relación entre las variables.

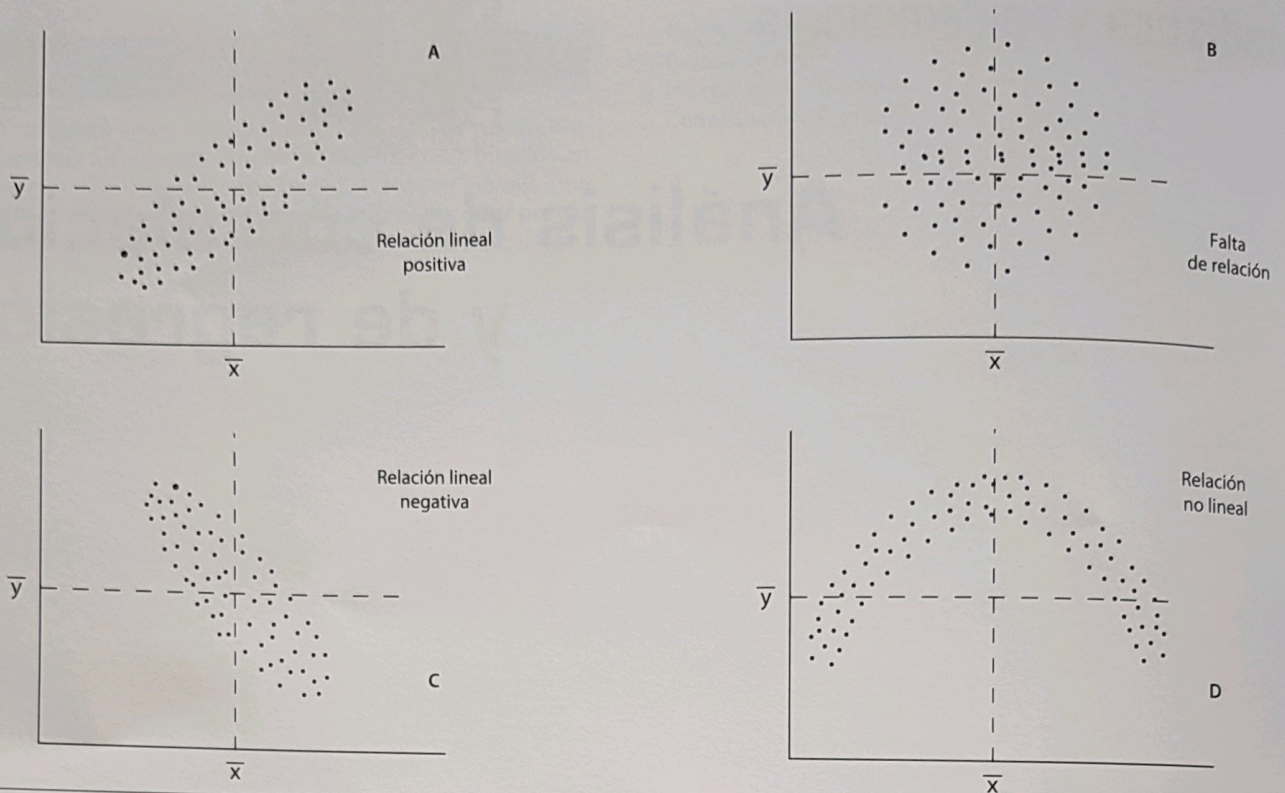


Figura 2. Distintos tipos de relación entre las variables

## Ideas clave

- Los test de correlación y de regresión sirven para establecer asociación entre dos variables cuantitativas.
- El análisis del test de correlación se basa en la interpretación del coeficiente  $r$ . El valor de  $r$  está comprendido entre  $-1$  y  $+1$ . Un valor positivo significa que si la variable "n" aumenta, la variable "m" lo hará también. El valor absoluto de la  $r$  establece lo fuerte

o débil de la correlación, considerándose fuerte cuando esta es mayor de 0,7.

- Una  $r$  de 1 o de  $-1$  implica una correlación cuya nube de dispersión se aproxima a la línea recta.
- Cuando la correlación es fuerte, podrá entonces hacerse el test de regresión lineal, cuya fórmula es  $y = \alpha + \beta x$ , donde  $\beta$  es la pendiente de la recta y  $\alpha$  el punto de corte de esta en el eje de ordenadas.

## Casos clínicos

Un estudio relaciona las concentraciones de colesterol ligado a las lipoproteínas de alta densidad (HDL-Col) con el número de cigarrillos fumados en la actualidad, en una muestra de 100 individuos. El coeficiente de correlación es  $-0,15$  ( $p < 0,05$ ). ¿Cuál de las siguientes afirmaciones es correcta?

- 1) El consumo de tabaco es la causa de que disminuya el HDL-Col.
- 2) El consumo de tabaco es la causa de la disminución del HDL-Col, porque la correlación es estadísticamente significativa.
- 3) El consumo de tabaco no es la causa del HDL-Col, porque la correlación es negativa.
- 4) Se ha observado una débil correlación negativa, estadísticamente significativa.

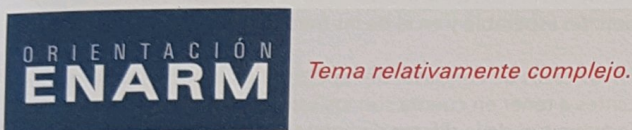
RC: 4



Haz clic en la imagen para abrir tu plataforma

# 17

## Tamaño muestral



### 17.1. Tamaño muestral

Para realizar un estudio sobre una población es necesario trabajar con una muestra, ya que abarcar a todos los miembros de una comunidad es prácticamente imposible.

El tamaño de la muestra se determina en el diseño del estudio de modo que sea apropiado para los objetivos buscados en el mismo y con los condicionamientos que se está dispuesto a asumir. Un número insuficiente de participantes impedirá encontrar las diferencias buscadas, concluyendo erróneamente que no existen, mientras que un número excesivo, aumenta innecesariamente el coste.

Manejar muestras es más barato, rápido, fácil y suficientemente preciso y exhaustivo que manejar poblaciones. Los datos obtenidos en la muestra se generalizarán a la población original utilizando ciertas pruebas estadísticas con una mínima probabilidad de error (< 5% habitualmente).

Hay dos campos en la inferencia estadística donde se va a desarrollar el cálculo del tamaño muestral:

- **Estimación de parámetros:** el tamaño dependerá de la variabilidad del parámetro que se desea estimar (esta ha de ser previamente conocida, o bien aproximarse a partir de datos preexistentes o estudios pilotos), de la precisión con que se pretenda obtener la estimación (amplitud deseada del intervalo de confianza, teniendo en cuenta que a mayor precisión, se necesitará un número más alto de sujetos) y del nivel de confianza.

Para estimar una proporción se debe conocer:

- El nivel de confianza o seguridad (1- $\alpha$ ). Para una seguridad del 95% = 2; para una seguridad del 99% = 2,6.
- La precisión que se desea para el estudio.
- Una idea del valor aproximado del parámetro que se quiere medir, revisando la literatura o los estudios previos. En caso de no tener dicha información, se empleará el valor  $p = 0,05$  (5%). En este caso, se necesita mayor  $n$ , pero garantizará una mejor estimación.

**Ejemplo:** ¿A cuántas personas se tendrían que estudiar para conocer la prevalencia de esquizofrenia en un área residencial de Madrid?

Seguridad = 95%; precisión = 5%; proporción esperada = próxima al 10%:

$$n: (22 \times p \times q) / d^2 = (4 \times 0,1 \times 0,9) / 0,05^2 = 144 \text{ donde}$$

- 22 ya que la seguridad es del 95%.
- $p$  = proporción esperada (en este caso, 10% = 0,1).
- $q = 1 - p$  (en este caso  $1 - 0,1 = 0,9$ ).
- $d$  = precisión (en este caso, se desea un 5%: 0,05).

Para **estimar una media** se debe conocer:

- El nivel de confianza o seguridad (1- $\alpha$ ), habitualmente 95%.
- La precisión con que se desea estimar el parámetro.
- Una idea de la varianza de la variable cuantitativa que se supone existe en la población.

**Ejemplo:** Si se desea conocer la media de la colesterolemia basal de una población, con una seguridad del 95% y una precisión de  $\pm 4$  mg/dl, y se tiene información por un estudio piloto o revisión bibliográfica de que la varianza es de 300 mg<sup>2</sup>/dl:

$$n: (22 \times 300) / 4^2 = 75$$

La precisión está en el denominador y al cuadrado, lo que significa que cuanto más preciso se pretenda ser, más pequeño será el número del denominador y, por ello, el tamaño muestral requerido será más grande.



En la estimación de parámetros, NO se necesita conocer el error  $\beta$  para estimar el tamaño muestral.



- **Contraste de hipótesis:** se utiliza para estudios comparativos y, en él, el tamaño de la muestra indica el número aproximado de sujetos que se necesitan para detectar una diferencia determinada, si existe, con unos márgenes de error previamente definidos. Los

factores a tener en cuenta para determinar el tamaño muestral son los siguientes:

- La magnitud de la diferencia a detectar que tenga interés clínicamente relevante. Se pueden comparar dos proporciones o dos medias.
- Tener una idea aproximada de los parámetros de la variable que se estudia (bibliografía, estudios previos).
- La variabilidad de la variable principal.
- La seguridad del estudio (riesgo de cometer un error  $\alpha$ : 5%).
- El poder estadístico ( $1 - \beta$ : 80% al menos) (riesgo de cometer un error  $\beta$ : 0,1- 0,2).
- La proporción de pacientes en los distintos grupos que responderá a cada tratamiento.

- La proporción de pérdidas ( $d$ ): se deberá multiplicar el tamaño calculado por:  $1/(1 - d)$ .
- La definición de si la hipótesis va a ser unilateral o bilateral.
  - › **Bilateral:** cualquiera de los dos parámetros a comparar (medias o proporciones) puede ser mayor o menor que el otro. No se establece dirección.  $H_0$  diferente de  $H_1$ .
  - › **Unilateral:** cuando se considera que uno de los parámetros debe ser mayor que el otro, indicando, por tanto, una dirección de las diferencias. Sólo  $H_0 > 0 < H_1$ .

La hipótesis bilateral es una hipótesis más conservadora y disminuye el riesgo de cometer un error de tipo I.

## Ideas clave

- Lo primero que se debe hacer cuando se pregunte sobre tamaño muestral es conocer si se trata de una estimación de parámetros o de un contraste de hipótesis, para así saber qué es lo que se necesita en el estudio.
- Si se trata de un intervalo de confianza, entonces se debe conocer si se trata de una variable cualitativa (proporción) o de una

cuantitativa (media). En ambos casos es preciso saber el nivel de confianza y la precisión con la que se desea el estudio. Evidentemente, en el caso de la proporción, se necesitará conocer la proporción esperable y en el de las medias, la varianza esperable.

- En los casos de contraste de hipótesis, los aspectos más importantes a tener en cuenta son los errores alfa y beta del estudio, la magnitud de la diferencia a detectar clínicamente relevante, la proporción de pacientes en los distintos grupos que responderán a cada tratamiento y la proporción de pérdidas.

## Casos clínicos

Para calcular el número mínimo de sujetos que es necesario incluir en un estudio sobre la prevalencia de un problema de salud, se han de realizar las siguientes asunciones, EXCEPTO:

- 1) Proporción de sujetos que se espera encontrar con el problema de salud.
- 2) El error tipo I.
- 3) El número de no respuestas esperado.
- 4) El error tipo II.

RC: 4

## Bibliografía

### Estadística y epidemiología

- ❑ Dawson/Trapp. *Bioestadística médica*. 4.ª ed. Manual Moderno, 2005.
- ❑ Hernández Aguado. *Manual de Epidemiología y Salud Pública*. 2.ª ed. Editorial médica Panamericana, 2011.
- ❑ Greenberg. *Epidemiología médica*. 4.ª ed. Manual Moderno, 2005.
- ❑ Grupo CTO. *Manual CTO de Estadística y epidemiología. Planificación y gestión*. 9.ª ed. CTO Editorial, Madrid, 2014.
- ❑ Ruiz, Morillo. *Epidemiología clínica*. Editorial médica Panamericana, 2004.



Haz clic en la imagen para abrir tu plataforma