

Epidemiología y Estadística

Índice

TEMA 1. ESTUDIO DE UN TEST. PARÁMETROS DE USO.	1
1.1. Estudio de un test.	1
1.2. Relación entre prevalencia y valores predictivos.	2
1.3. Aceptabilidad de un método diagnóstico.	2
1.4. Screening.	2
1.5. Coeficiente Kappa.	2
TEMA 2. CONCEPTO Y USO DE LA EPIDEMIOLOGÍA.	2
2.1. Concepto de epidemiología.	2
2.2. Indicadores de riesgo. Causas de enfermedad.	2
TEMA 3. MEDIDAS DE FRECUENCIA DE LA ENFERMEDAD.	3
3.1. Formas básicas de medida.	3
3.2. Medidas de frecuencia de la enfermedad.	3
TEMA 4. MEDIDAS DE ASOCIACIÓN O DEL EFECTO.	3
4.1. Riesgo relativo (RR).	4
4.2. Odds ratio (razón de desventaja).	4
4.3. Razón de prevalencia.	4
TEMA 5. MEDIDAS DE IMPACTO.	4
5.1. Diferencia de incidencias o riesgo atribuible (RA).	4
5.2. Fracción atribuible en expuestos o fracción etiológica del riesgo (FAE).	4
TEMA 6. TIPOS DE ESTUDIOS EPIDEMIOLÓGICOS.	4
6.1. Estudios descriptivos.	4
6.2. Estudios analíticos.	5
TEMA 7. VALIDEZ Y FIABILIDAD DE LOS ESTUDIOS EPIDEMIOLÓGICOS.	7
7.1. Tipos de Error.	7
7.2. Validez y fiabilidad.	8
7.3. Validez de un test diagnóstico.	8
TEMA 8. ENSAYO CLÍNICO.	8
8.1. Introducción.	8
8.2. Etapas en la realización de un ensayo clínico.	8
8.3. Tipos de ensayos clínicos.	9
TEMA 9. NIVELES DE CALIDAD DE LA EVIDENCIA CIENTÍFICA.	10
9.1. Factores determinantes de la calidad de la evidencia científica.	10
9.2. Escalas de gradación de la evidencia científica.	10
TEMA 10. ESTADÍSTICA. GENERALIDADES.	10
10.1. Tipos de estadística.	10
10.2. Población y muestra.	10
10.3. Variables.	11

TEMA 11. ESTADÍSTICA DESCRIPTIVA.....	11
11.1. Representación gráfica.	11
11.2. Síntesis de los datos.	12
TEMA 12. PROBABILIDAD.....	12
12.1. Probabilidad condicionada.	12
12.2. Regla de la multiplicación.	13
12.3. Regla de la suma.	13
TEMA 13. DISTRIBUCIONES DE PROBABILIDAD.	13
13.1. Distribución binomial.	13
13.2. Distribución de Poisson.....	13
13.3. Distribución normal o de Gauss.....	13
TEMA 14. ESTIMACIÓN DE PARÁMETROS.	13
14.1. Estimación de medias.....	13
14.2. Estimación de un porcentaje.....	14
14.3. Estimación de medidas de asociación.	14
TEMA 15. CONTRASTE DE HIPÓTESIS.....	14
15.1. Hipótesis nula, hipótesis alternativa y grado de significación estadística.	14
15.2. Errores alfa y beta.	14
15.3. Pruebas de significación estadística.	15
TEMA 16. DISTRIBUCIÓN CHI-CUADRADO.	15
TEMA 17. ANÁLISIS DE CORRELACIÓN Y REGRESIÓN.	16
TEMA 18. TAMAÑO MUESTRAL.....	16

TEMA 1. ESTUDIO DE UN TEST. PARÁMETROS DE USO.

I.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 tiene o no la enfermedad.

La eficacia de una prueba diagnóstica depende de su capacidad para detectar correctamente la presencia o ausencia de la enfermedad que se estudia, lo que se expresa matemáticamente en cuatro índices como: sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo.

Estos 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 I.

		ENFERMO	SANO	
Resultado de la prueba diagnóstica que evaluamos	POSITIVO	VP a	FP b	Total positivos a+b
	NEGATIVO	FN c	VN d	Total negativos c+d
		Total de enfermos a+c	Total de sanos b+d	TOTAL 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.

- **Sensibilidad (S).** Se define como la probabilidad de que un individuo enfermo tenga un test +. La sensibilidad indica la proporción del total de enfermos que el test es capaz de detectar (MIR 03-04, 143; MIR 00-01F, 238; MIR 98-99F, 205).

$$S = \frac{\text{Individuos enfermos con test } \oplus}{\text{Todos enfermos}} = \frac{VP}{VP+FN}$$

- **Especificidad (E).** Probabilidad de que un individuo sano tenga un test -. La especificidad indica la proporción de individuos sanos confirmados como tales por el resultado negativo del test (MIR 98-99F, 208).

$$E = \frac{\text{Individuos sanos con test } \ominus}{\text{Todos sanos}} = \frac{VN}{VN+FP}$$

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

$$TFN = \frac{\text{Individuos enfermos con test } \ominus}{\text{Todos enfermos}} = \frac{FN}{VP+FN} = 1-S$$

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

$$TFP = \frac{\text{Individuos sanos con test } \oplus}{\text{Todos sanos}} = \frac{FP}{VN+FP} = 1-E$$

- **Razón de probabilidad positiva.** Compara la probabilidad de que un paciente enfermo presente un resultado positivo comparado con la probabilidad de que el resultado positivo se presente en un individuo sano (MIR 99-00F, 209; MIR 98-99, 198; MIR 98-99, 201).

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

$$RP+ = S/I-E$$

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

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

$$RP- = I-S/E$$

- **Valor predictivo positivo.** Se trata de la proporción de verdaderos positivos entre aquellos que han sido identificados como positivos en una prueba de test (MIR 03-04, 135; MIR 01-02, 205; MIR 99-00, 239; MIR 98-99, 199; MIR 98-99F, 204; MIR 97-98, 81; MIR 96-97, 138; MIR 96-97, 139; MIR 95-96, 30; MIR 95-96F, 60; MIR 94-95, 60).

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

- **Valor predictivo negativo.** Se trata de la proporción de verdaderos negativos entre aquellos que han sido identificados como negativos en un test (MIR 03-04, 135; MIR 02-03, 28; MIR 00-01, 217; MIR 94-95, 236).

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

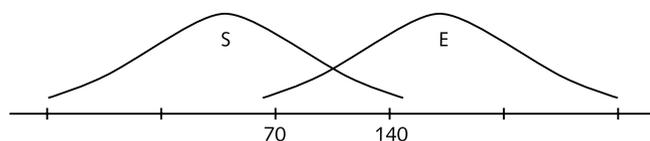
- **Valor global (eficiencia) del test.** Indica la proporción de resultados válidos entre el conjunto de resultados.

$$VG = \frac{VP+VN}{VP+VN+FP+FN}$$

- * **Sensibilidad + Tasa de falsos negativos = 100%.**

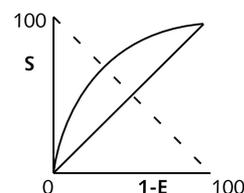
- * **Especificidad + Tasa de falsos positivos = 100%** (MIR 96-97, 152; MIR 95-96F, 54).

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



El punto de corte escogido determinará la sensibilidad y especificidad de la prueba (si cogemos 70, la prueba será muy sensible y poco específica; si cogemos 140, será poco sensible y muy específica) (MIR 96-97, 142).

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 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).



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 especificidad).

De esta forma, la sensibilidad y la especificidad son valores interdependientes, de forma que si aumenta la sensibilidad disminuye la especificidad y viceversa. Si se adoptan criterios de diagnóstico muy estrictos disminuye la sensibilidad (hay menos enfermos que cumplen estos criterios), y paralelamente aumenta la especificidad (pocos sanos cumplen estos criterios) (MIR 01-02, 206; MIR 99-00F, 217; MIR 99-00F, 203; MIR 98-99F, 209; MIR 96-97, 147; MIR 96-97F, 216; MIR 96-97F, 197).

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, aumenta el valor predictivo positivo, mientras que disminuye el valor predictivo negativo (MIR 99-00F, 203; MIR 97-98, 63; MIR 97-98, 87; MIR 97-98F, 69; MIR 95-96, 151).

Si la prevalencia de la enfermedad disminuye, aumenta el valor predictivo negativo y disminuye el valor predictivo positivo (MIR 05-06, 215; MIR 96-97, 147; MIR 95-96, 95).

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 te escaparán pocos, aunque al precio de bastantes "falsos positivos". Elegiremos un test sensible cuando:

- La enfermedad sea grave y no pueda pasar desapercibida.
- La enfermedad sea tratable.
- Los resultados falsamente positivos no supongan un traumatismo psicológico en los individuos examinados (MIR 98-99, 200).

Si lo que quieres es "asegurar" el diagnóstico, debes usar un test cuya especificidad sea máxima.

Utilizaremos 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 (MIR 02-03, 49; MIR 01-02, 203; MIR 00-01, 210; MIR 96-97, 130; MIR 95-96F, 49).

1.4. Screening.

Estrategia de detección precoz de la enfermedad. Lo ideal es aplicar primero un test muy sensible (detecta todos los casos posibles de enfermedad → se obtienen muchos FP y pocos FN) y en una segunda fase aplicar un test muy específico (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 criterios dependientes de la enfermedad, del test y de la población diana.

1) Criterios dependientes de la enfermedad.

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

2) Criterios dependientes del test.

- De fácil aplicación.
- Coste razonable dentro del presupuesto de salud.
- Inocua para la población.
- Confiabilidad o capacidad de repetición.
- Validez. Se refiere a la capacidad del test de medir lo que realmente deseamos medir. El concepto de validez incluye los de sensibilidad, especificidad y valor predictivo.

3) Criterios dependientes de la población diana.

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

(MIR 98-99F, 213; MIR 97-98, 66; MIR 96-97, 145; MIR 95-96, 109; MIR 95-96F, 188).

1.5. Coeficiente Kappa.

El coeficiente Kappa es una medida del grado de concordancia no aleatoria entre varios observadores o entre distintas mediciones de la misma variable. Varía de -1 a +1 (MIR 97-98, 85; MIR 96-97F, 210).

- $k = -1$. 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.

TEMA 2. CONCEPTO Y USO DE LA EPIDEMIOLOGÍA.

2.1. Concepto de epidemiología.

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

Podríamos 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, y sobre los recursos necesarios para los cuidados de salud de esa población.

La epidemiología se ocupa de dos aspectos fundamentales:

1. Estudiar la distribución de las enfermedades en relación con las variables "lugar", "tiempo" y "persona". Es lo que se denomina Epidemiología Descriptiva.
2. Buscar los factores que determinan la distribución encontrada e identificar asociaciones causales. Es lo que hace la llamada Epidemiología Analítica.

Las principales posibilidades de **aplicación** de la Epidemiología son:

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

Los principales pasos dentro del **método epidemiológico** serían:

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 podemos atribuir "responsabilidad" en la enfermedad. P. ej.: 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. P. 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. P. ej.: 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 deben cumplirse los denominados criterios de causalidad.

- **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 (ver Tema 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 (reproducibilidad).
- **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 (MIR 99-00, 240; MIR 99-00F, 207; MIR 98-99, 202; MIR 96-97F, 212).

MODELOS CAUSALES.

- **Determinista o unicausal.** Este modelo dice que siempre que se presente la causa (causa suficiente) y sólo cuando se presente la causa (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 produce varios efectos.
- **Determinista modificado (Rothman).** En este modelo se distinguen 3 tipos de causas:
 1. *Causa suficiente.* Es aquella que inevitablemente produce el efecto cuando se presenta.
 2. *Causa complementaria o contribuyente.* Son aquellas causas que forman parte de una causa suficiente.
 3. *Causa necesaria.* Es la que debe estar presente inevitablemente para que se produzca el efecto, por tanto, debe formar parte de todas las causas suficientes de una enfermedad.

TEMA 3. MEDIDAS DE FRECUENCIA DE LA ENFERMEDAD.

3.1. Formas básicas de medida.

Número. Es un valor absoluto. Al no conocer el denominador no se puede conocer la importancia relativa del problema.

Razón. 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 de 0 a infinito.

- Fórmula: $R = a/b$
- Ejemplo: razón de hombres/mujeres.

Proporción. Cociente en que el numerador está incluido en el denominador. Sus valores oscilan de 0 a 1. Es adimensional.

- Fórmula: $P = a/a+b$
- Ejemplo: proporción de aprobados en un examen.

Tasa. Mide la ocurrencia de un suceso durante un período de tiempo en una población determinada. Es un cociente en el que el numerador está incluido en el denominador pero, a diferencia de la proporción, el tiempo también está incluido en el denominador. Tienen, 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.

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. Es mayor a mayor incidencia y mayor duración de la enfermedad (MIR 96-97, 141; MIR 96-97F, 213).

$$Pv = \frac{\text{Nº total de casos de enfermedad}}{\text{Población en ese momento}}$$

Incidencia acumulada. Se define como el número de casos nuevos de una enfermedad que se producen en un período de tiempo. Es la probabilidad de que un individuo desarrolle la enfermedad en ese período (MIR 01-02, 195; MIR 00-01F, 234; MIR 98-99F, 212; MIR 96-97, 155; MIR 95-96, 44; MIR 95-96, 37).

$$IA = \frac{\text{Nº casos nuevos en un período de tiempo}}{\text{Población en riesgo al inicio de ese período}}$$

Ésta se puede relacionar con la prevalencia según las siguientes fórmulas.

$$Pv < 10\% \quad Pv = I \times \text{Duración media de la enfermedad}$$

$$Pv > 10\% \quad Pv = \frac{I \times D}{1 + (I \times D)}$$

Densidad de incidencia. Se define como el número de casos nuevos de la enfermedad que se producen en un período de tiempo, teniendo en cuenta la información que aporta cada uno de los individuos.

$$DI = \frac{\text{Nº casos nuevos}}{\text{Persona} \times \text{tiempo}}$$

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 de incidencia es una tasa que expresa la velocidad con la que se desarrolla una enfermedad en una población y, para ser válida, todos los sujetos debieran haber sido seguidos un tiempo suficiente para desarrollarla. Este concepto se conoce también con el nombre de **fuerza de la morbilidad**.

Tabla 2. Medidas de frecuencia de la enfermedad.

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

TEMA 4. MEDIDAS DE ASOCIACIÓN O DEL EFECTO.

Miden la diferencia de ocurrencia de una enfermedad entre dos grupos de personas: personas expuestas (a un factor de riesgo) y personas no expuestas.

4.1. Riesgo relativo (RR).

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 (MIR 05-06, 198; MIR 03-04, 136; MIR 03-04, 149; MIR 03-04, 152; MIR 02-03, 29; MIR 00-01, 193; MIR 98-99F, 211; MIR 96-97, 129).

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 no expuestos.
- RR < 1 → Factor de protección.

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

	ENFERMOS	SANOS	TOTAL
FR	a	b	a + b
No FR	c	d	c + d
Total	a + c	b + d	a+b+c+d

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

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

(MIR 01-02, 212; MIR 99-00F, 211; MIR 99-00F, 198; MIR 97-98F, 83).

4.2. Odds ratio (razón de desventaja).

Es una razón. Es la medida básica de los estudios casos-controles. Su significado es idéntico al del RR, aunque no puede calcularse como él, ya que en los estudios casos-controles no puede calcularse 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 que han surgido los casos y que la enfermedad tenga una incidencia baja, inferior al 10%. (MIR 02-03, 36; MIR 00-01, 196; MIR 99-00F, 209).

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

	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

$$OR = \frac{a/c}{b/d} = \frac{a \times d}{b \times c} \text{ (Cociente de productos cruzados)}$$

4.3. Razón de prevalencia.

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.

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

TEMA 5. MEDIDAS DE IMPACTO.

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 (MIR 03-04, 127; MIR 03-04, 152; MIR 00-01F, 249; MIR 94-95, 237).

$$RA = I_e - I_o$$

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

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 eliminásemos el factor de riesgo (MIR 02-03, 39; MIR 98-99, 193).

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

TEMA 6. TIPOS DE ESTUDIOS EPIDEMIOLÓGICOS.

6.1. Estudios descriptivos.

Los objetivos de los estudios descriptivos son:

1. Describir las características y la frecuencia de un problema de salud, en función de las características de PERSONA (edad, sexo, estado civil...), de LUGAR (área geográfica...) y de TIEMPO de aparición del problema y su tendencia.
2. Servir de base para estudios analíticos.

1. **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.

2. **Estudios ecológicos.** Pueden ser transversales o longitudinales. Son estudios en los que la unidad de análisis son grupos de individuos, no individuos (p.ej. clases de una escuela, ciudades, regiones). Son útiles cuando no se pueden hacer mediciones de exposición individuales (contaminación del aire, ruidos, etc.) (MIR 05-06, 196).

- **Ventajas.** Permiten describir diferencias en poblaciones que habrán de ser estudiadas con más detalle posteriormente (MIR 00-01, 198).
- **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.

3. **Estudios transversales o de prevalencia.** Son estudios descriptivos y transversales, ya que estudian 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 (MIR 00-01, 195; MIR 00-01F, 240).

- **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 reproductibilidad. Son útiles para el estudio de enfermedades crónicas en la población (MIR 99-00F, 205; MIR 95-96, 37).
- **Inconvenientes.** No es útil para estudiar enfermedades raras, no permite ver el mecanismo de producción de la enfermedad y no sirve para comprobar una hipótesis previa de causalidad (no es posible conocer la secuencia temporal porque

la información sobre el factor de riesgo y la enfermedad se recogen a la vez (MIR 01-02, 202; MIR 97-98, 62; MIR 96-97F, 214; MIR 95-96, 29).

Análisis de las medidas de enfermedad. Con el estudio de prevalencia, la medida que se obtiene es la RAZON DE PREVALENCIA de individuos expuestos (MIR 96-97, 136; MIR 95-96F, 53).

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 estudios experimentales 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 observacionales, el investigador se limita a observar qué es lo que sucede en un grupo de individuos, sin manipular el estudio (MIR 05-06, 195).

6.2.1. Estudios analíticos experimentales.

TERMINOLOGÍA.

Decimos que un estudio es experimental cuando cumple dos condiciones:

- 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.

Problemas éticos. Es 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 (MIR 02-03, 27).
- Permiten un gran control sobre cualquier efecto que no sea el estudiado.
- Permiten el empleo de técnicas de enmascaramiento.

1. **Ensayo clínico aleatorio.** Es, con mucho, el estudio experimental más frecuente. La asignación aleatorizada del factor de estudio (un fármaco o intervención sanitaria) se hace sobre los individuos. Es el que ofrece la mejor evidencia de una posible relación causa-efecto y la eficacia de una actuación (MIR 05-06, 199; MIR 99-00, 229; MIR 98-99, 259; MIR 97-98, 78; MIR 97-98F, 74; MIR 96-97F, 211; MIR 96-97, 144). (Ampliar estudio en Tema 8).



Figura 1. Ensayo clínico aleatorio.

2. **Ensayo de campo.** Es un estudio experimental 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 (MIR 99-00, 242).

6.2.2. Estudios analíticos cuasiexperimentales.

Se diferencian de los estudios experimentales puros en que no hay asignación al azar (aleatorización).

1. **Ensayo comunitario de intervención.** Son una variedad de los ensayos de campo (MIR 00-01, 194).

- Se aplica a individuos sanos.
- Valora la eficacia de medidas preventivas.
- No se aplica aleatorización individual.

2. **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 tienen el inconveniente grave de que, al no disponer de grupo de control, los resultados son difíciles de interpretar.

3. **Estudios controlados no aleatorios.** Se realizan cuando la asignación aleatoria, o no ofrece ventajas, o no se puede hacer.

6.2.3. Estudios analíticos observacionales.

1. Estudios de cohortes.

DESCRIPCIÓN.

Partiendo de un grupo de individuos expuestos al factor de riesgo (**cohorte expuesta**), y de otros comparables en todo pero que no estén expuestos al FR (**cohorte no expuesta**), se estudia la incidencia de la enfermedad en ambas cohortes (MIR 05-06, 200; MIR 03-04, 132; MIR 01-02, 199; MIR 00-01F, 237; MIR 94-95, 239).

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.
- Sirven para el estudio de exposiciones raras.

INCONVENIENTES.

- No es bueno para el estudio de enfermedades raras (MIR 01-02, 198).
- No es bueno para el estudio de enfermedades de largo período de incubación.
- El coste es alto.
- No sirve para el estudio de la “multicausalidad de la enfermedad”.
- Son difícilmente reproducibles.

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.
- Fracción atribuible. Estima la proporción de la enfermedad entre los expuestos, que es debida al factor de riesgo. (MIR 99-00, 235)

2. **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 (p.ej. historias clínicas) sigue la evolución de ambas cohortes, comparando los resultados.

Por ejemplo: ¿Los trabajadores que son despedidos, sufren una mayor morbilidad que los que no son despedidos?

En 1989 se diseñó un estudio de cohortes retrospectivo, consistente en identificar a los trabajadores despedidos de una fábrica. De esa misma localidad, se identificó a otro grupo de trabajadores

que siguieron trabajando. Mediante las historias clínicas, se estudió la morbilidad de ambas cohortes.

3) Estudio de casos-controles.

DESCRIPCIÓN.

Partiendo de un grupo de individuos enfermos (**casos**), y de otros comparables a ellos en todo, pero que no tienen la enfermedad (**controles**), se estudia la exposición, en ambos, a distintos factores de riesgo (MIR 05-06, 197; MIR 03-04, 129; MIR 03-04, 137; MIR 01-02, 196; MIR 00-01, 197; MIR 95-96F, 56).

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 período de inducción.
- El coste es bajo.
- Es el mejor 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.

INCONVENIENTES.

- No es bueno para comprobar hipótesis previas de causalidad.
- No permite el estudio de la "multiefectividad 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-controles no puedes obtener información sobre la incidencia

de la enfermedad ya que partes de una población seleccionada. Tampoco tienes información sobre la prevalencia, ya que el número de enfermos sólo depende de los que tú elijas. 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 (MIR 99-00, 234).

Tabla 3. Diferencias entre los estudios de cohortes y los de casos y controles.

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

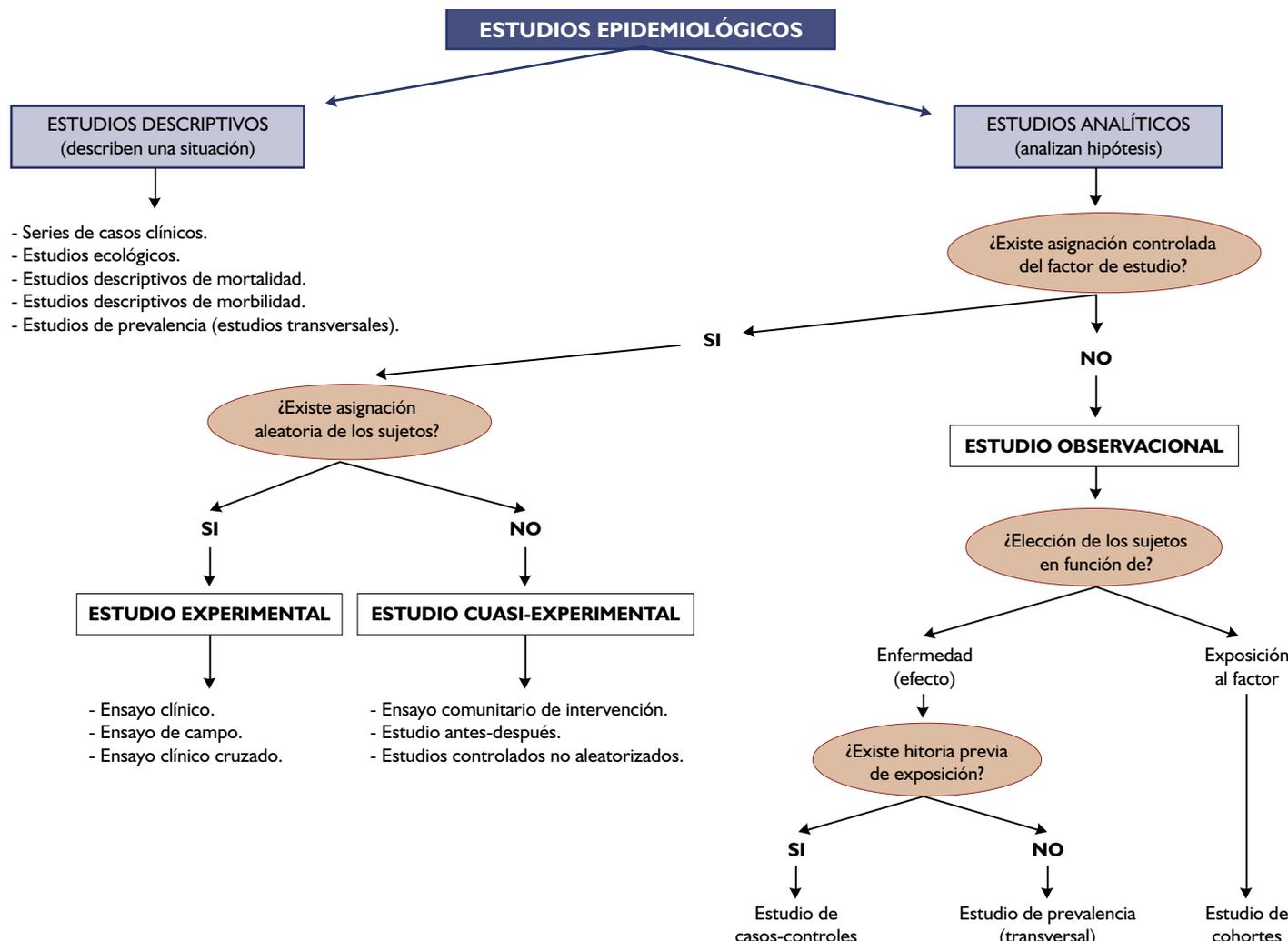


Figura 2. Tipos de estudios epidemiológicos.

TEMA 7. VALIDEZ Y FIABILIDAD DE LOS ESTUDIOS EPIDEMIOLÓGICOS.

Cuando realizamos un estudio de investigación clínica casi nunca trabajamos con poblaciones completas. Lo que hacemos es, partiendo de observaciones realizadas en un grupo reducido de personas (la llamada muestra), generalizar o extrapolar nuestros resultados a colectivos más amplios. El hecho de que no trabajemos con poblaciones completas, sino con muestras, puede introducir en nuestras observaciones errores producidos por el azar, a los que llamamos 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.

1. **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 vamos 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 (MIR 97-98F, 67; MIR 96-97F, 206).
2. **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.

TIPOS DE ERRORES SISTEMÁTICOS.

1. **Sesgo de selección** (MIR 02-03, 40; MIR 94-95, 235). Se produce cuando se asignan sujetos al grupo control que difieren significativamente, en alguna característica clave, del grupo problema. Este tipo de sesgos se pueden controlar mediante un proceso de aleatorización en la constitución de los distintos grupos de estudio (MIR 95-96, 34). Son ejemplos de este tipo de sesgo:
 - Sesgo de autoselección o del voluntario. La participación o autoderivación 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, eliges una muestra hospitalaria de esa población y el factor de riesgo que estamos estudiando se asocia a una mayor probabilidad de hospitalización. También se incluye en este tipo de sesgo aquel que puede surgir al elegir como control 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 casos y 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.

Ejemplo. Queremos estudiar si la actividad física tiene algún efecto sobre la frecuencia de insuficiencia coronaria. Comparamos personas con actividad y personas sedentarias. Un sesgo de selección sería cuando la inactividad de los sedentarios fuese a consecuencia de padecer la enfermedad cardíaca.

2. **Sesgo de información.** Se produce cuando hay un error sistemático en la medición de alguna variable clave del estudio. La clasificación incorrecta puede afectar de igual manera a todos los grupos de estudio o no. Los sesgos de información debidos al investigador o a los participantes en el estudio se controlan mediante técnicas de enmascaramiento (MIR 05-06, 210; MIR 00-01, 201).

- Clasificación incorrecta diferencial. La probabilidad de clasificación errónea de la exposición (en un estudio de casos y control) afecta de manera 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 sub 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 casos y controles. El hecho de padecer la enfermedad hace que estés 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 manera similar. El error de clasificación no diferencial produce una infraestimación del efecto del factor de exposición estudiado en la enfermedad (MIR 99-00F, 200).

Ejemplo. Queremos estudiar el efecto del tabaco sobre la aparición de bronquitis crónica, comparándose un grupo de fumadores y 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 la bronquitis crónica.

Tabla 4. Tipos de error.

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

3. **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 (MIR 03-04, 147; MIR 02-03, 43; MIR 01-02, 200; MIR 00-01F, 236; MIR 00-01, 208):
 - Ser un factor de riesgo para la enfermedad.
 - Estar asociado con la exposición.
 - No ser un paso intermedio entre la exposición y la enfermedad (MIR 97-98, 88).
 Para prevenir los sesgos de confusión tenemos distintas técnicas (MIR 01-02, 215):
 - Fase de diseño: aleatorización (estudios experimentales), apareamiento y restricción (MIR 00-01, 219; MIR 99-00F, 206).
 - Fase de análisis estadístico: análisis estratificado (dividir los datos globales en dos grupos según la presencia o no del factor de confusión) y análisis multivariante (MIR 98-99F, 217; MIR 98-99F, 214; MIR 96-97F, 207; MIR 94-95, 233).

Ejemplo. Queremos saber si existe relación entre el alcohol y el cáncer de pulmón. La OR cruda es de 2,4. Sospechamos que el tabaco puede actuar como factor de confusión. Al estratificar observamos que:

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

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). La verdadera OR entre alcohol y Ca. de pulmón es de 1.

7.2. Validez y fiabilidad.

Validez. El estudio mide 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 estudio son aplicables a los individuos del estudio.
- Validez externa. Se dice que un estudio tiene validez externa cuando los resultados del estudio son aplicables a otros individuos distintos de los del estudio (MIR 96-97, 151; MIR 95-96F, 63).

Fiabilidad. Es el grado de reproductibilidad de un estudio, es decir, el grado de similitud que presentarían los resultados si repitieses el estudio en condiciones similares. Es el grado de ausencia de error aleatorio (MIR 96-97F, 223).

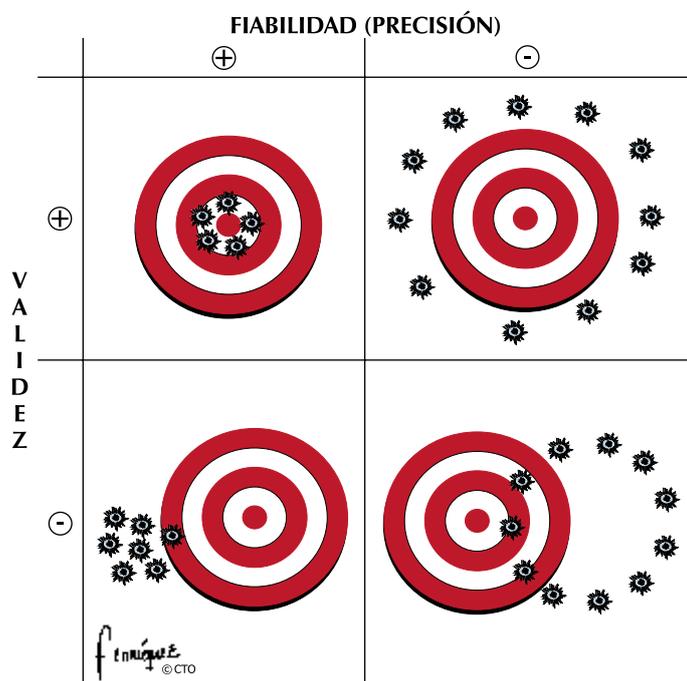


Figura 3. Tipos de error.

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 la de referencia se denomina **validez de criterio**.

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

El valor real de la prueba sólo podrá ser establecido si el estudio se realiza 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 (MIR 02-03, 50; MIR 01-02, 204; MIR 00-01F, 235; MIR 00-01, 212; MIR 96-97, 151).

TEMA 8. ENSAYO CLÍNICO

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 (MIR 99-00F, 216; MIR 96-97F, 198; MIR 96-97F, 211; MIR 95-96F, 48).

Teniendo en cuenta los objetivos perseguidos en el desarrollo de un medicamento, se distinguen 4 fases:

- Ensayo en fase I: es la primera vez que un fármaco se da a humanos. Generalmente se realiza con voluntarios sanos (n=20-80) y sin grupo control. El principal objetivo es evaluar la toxicidad y conocer la dosis única aceptable no tóxica (MIR 05-06, 205).
- Ensayo en fase II: el principal objetivo es aportar información sobre la relación dosis/respuesta, proporcionando una información preliminar sobre la eficacia. Se hacen en pacientes (n=100-200). No necesariamente tienen que ser comparativos (MIR 02-03, 23).
- Ensayo en fase III: es el prototipo del ensayo clínico. Suele ser comparativo con la terapéutica de referencia o con un placebo. Es la investigación 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 (MIR 05-06, 202).
- Ensayo en fase IV: también se denomina farmacovigilancia y consiste en el seguimiento postcomercialización 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 < 2000). El Sistema Español de Farmacovigilancia, recibe información de diferentes fuentes, entre ellas, y principalmente, la Notificación Espontánea de reacciones adversas (MIR 05-06, 214).

La notificación espontánea de reacciones adversas, supone la recogida y notificación mediante las tarjetas amarillas, de las reacciones adversas que aparecen durante la práctica clínica.

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 3 años). El programa de Notificación Espontánea está dirigido especialmente a médicos, aunque incluye a otros profesionales sanitarios.

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.

8.2.1. 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 enfermos en los que se empleará el fármaco. Los **criterios de exclusión** se utilizan para desdenar a aquellos enfermos con criterios de inclusión pero que presentan alguna contraindicación, condiciones que pueden afectar a la variable resultado o alguna característica que le haga complicado de ser estudiado (MIR 01-02, 211).

En función de los criterios de inclusión, podemos 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) (MIR 03-04, 257; MIR 00-01, 200).
- **EC explicativos:** los criterios de inclusión son muy estrictos, por lo que la población no es exactamente igual a la población general.

Cuando 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 podrá generalizar los resultados con mayor facilidad y tendrá mayor validez externa. Sin embargo, la muestra será heterogénea, se necesita un mayor tamaño muestral y los resultados son confusos si sólo es eficaz en subgrupos de pacientes. Se utilizan 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.

8.2.2. Medición de variables basales.

Consiste en determinar una serie de variables en los sujetos que cumplen los criterios de inclusión y 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 nos permitan posteriormente verificar que la aleatorización ha sido eficaz y hacer análisis estratificado (Ej. por edad) en caso de diferencias entre los dos grupos (MIR 01-02, 216; MIR 00-01F, 243).
- 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.

8.2.3. 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 (MIR 02-03, 42):

- **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 bloque 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 (MIR 02-03, 24).

En general, diremos que la aleatorización debe conseguir asignar los tratamientos de forma desconocida e impredecible. Debe 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, debe ser posible detectar los fallos de la asignación.

Para evitar sesgos en la aleatorización es importante que quien decide la inclusión de los paciente en el ensayo y quien lleva a cabo la aleatorización, desconozcan la secuencia de aleatorización hasta que se aplique a cada uno de los enfermos 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 (MIR 01-02, 211).

8.2.4. 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 esto, el llamado sesgo de cointervención, y que la medición de la variable resultado pueda estar sesgada por el mismo motivo utilizaremos el **enmascaramiento**. Los tipos de enmascaramiento son:

- **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 desconoce 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 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.

8.2.5. Análisis de los resultados.

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 (MIR 00-01, 203; MIR 00-01F, 242).
- 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 uso de análisis de supervivencia (MIR 05-06, 216).
- Comparaciones múltiples: al aumentar el número de comparaciones, aumenta 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 espúreos, y por tanto, sus resultados deben interpretarse con precaución (MIR 00-01F, 244).
- Análisis por intención de tratar frente a análisis por protocolo (MIR 05-06, 206; MIR 01-02, 208; MIR 98-99, 194; MIR 97-98, 86):
 - 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.
- 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 debe mantener la confidencialidad de los resultados y debe haber una penalización estadística porque, de nuevo, se están realizando comparaciones múltiples (MIR 02-03, 26).
- 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 riesgo de sufrir el mismo evento en el grupo control (no expuesto al tratamiento) (MIR 02-03, 29; MIR 01-02, 209).
 - La reducción absoluta del riesgo (RAR), que es la diferencia entre el porcentaje de eventos en el grupo control y el porcentaje de eventos en el grupo experimental (MIR 05-06, 207; MIR 00-01, 204).
 - La reducción relativa del riesgo (RRR) es la diferencia entre el porcentaje de eventos en el grupo control y el porcentaje de eventos 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 (MIR 05-06, 201; MIR 03-04, 127; MIR 02-03, 47; MIR 01-02, 207; MIR 00-01, 218; MIR 99-00F, 199; MIR 98-99, 197).

8.3. Tipos de ensayos clínicos.

- **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. Los hallazgos serán estadísticamente significativos cuando la variabilidad intergrupo sea suficientemente mayor que la intragrupal.

- **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 cada paciente sirve de control a sí mismo, lo cual, al reducir la variabilidad, permite llevar a cabo el estudio con un tamaño muestral menor al del diseño clásico. 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 período de lavado entre la administración de los dos fármacos, de modo que el primer fármaco haya sido totalmente aclarado del organismo antes de administrar el segundo (MIR 03-04; 134; MIR 03-04, 138; MIR 02-03, 45; MIR 01-02, 217; MIR 00-01F, 240).

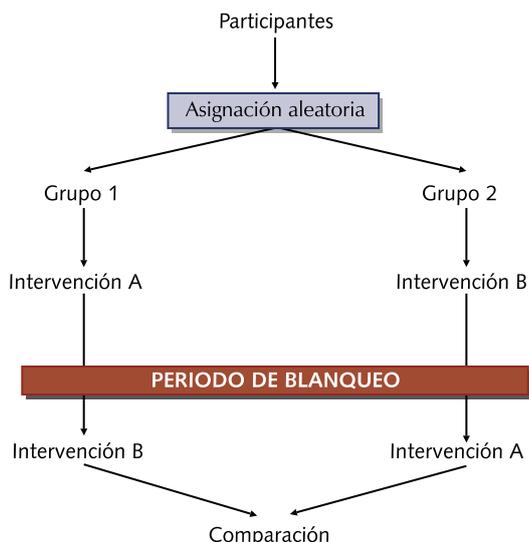


Figura 4. Ensayo clínico cruzado.

- **Diseño secuencial:** consiste en introducir pares de pacientes randomizados a los dos tratamientos hasta que la diferencia entre los distintos tratamientos favorece a uno u otro (exceso de preferencias), momento en el que el ensayo clínico se detiene (MIR 05-06, 211; MIR 02-03, 22; MIR 00-01F, 241).
- **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. Con este diseño se puede evaluar la interacción entre los dos tratamientos, aunque hay que asumir que los dos tratamientos no interaccionan entre sí (MIR 05-06, 204).

TEMA 9. NIVELES DE CALIDAD DE LA EVIDENCIA CIENTÍFICA.

9.1. 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:

- 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.

9.2. Escalas de gradación de la evidencia científica.

Todas son muy similares entre sí y presentan al metaanálisis de EC controlados y aleatorizados como el estudio más riguroso y que aporta mayor evidencia causal (MIR 02-03, 21; MIR 01-02, 197; MIR 99-00, 229; MIR 98-99, 259; MIR 97-98F, 74; MIR 97-98, 78; MIR 96-97, 144).

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): existe 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.

Tabla 5. Gradación de la evidencia científica.

Calidad de la evidencia científica	Tipo de diseño
BUENA	Metaanálisis de EC controlados y aleatorizados.
	EC controlado y aleatorizado de muestra grande.
	EC controlado y aleatorizado de muestra pequeña.
REGULAR	EC controlado no aleatorizado con controles concurrentes.
	EC controlado no aleatorizado con controles históricos.
	Estudio de cohortes prospectivos.
	Estudio de cohortes retrospectivos.
MALA	Estudio de casos y controles.
	Series clínicas, estudios transversales y ecológicos.
	Comités de expertos.

TEMA 10. ESTADÍSTICA. GENERALIDADES.

Concepto de estadística.

Método de razonamiento que permite interpretar un conjunto de datos cuyo carácter esencial es la variabilidad.

10.1. Tipos de estadística.

1. **Estadística descriptiva.** Organización, presentación y síntesis de los datos de una manera científica y lógica.
2. **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.

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**:

- a. **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.
- b. **Muestreo aleatorio estratificado.** Con este tipo de muestreo aseguras que la muestra tenga la misma proporción de una(s) variable(s) que la población de la que procede.
- c. **Muestreo sistemático.** El proceso de selección se basa en alguna regla sistemática simple, por ejemplo, elegir uno de cada "n" individuos.
- d. **Muestreo por etapas o polietápico.** La selección se realiza en dos o más etapas sucesivas y dependientes.

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 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.

1. Tipos de variables.

VARIABLE	SUBTIPO	EJEMPLO
Cualitativa	Nominal dicotómica	Sexo
	Nominal no dicotómica	Raza
	Ordinal	Nivel socioeconómico
Cuantitativa	Discreta	Nº episodios de asma/sem
	Continua	Uricemia

2. **Escalas de medición.** Son criterios usados para definir las diferentes categorías en que se pueden agrupar los valores de la variable. Las categorías de las variables deben cumplir dos condiciones:

- Mutuamente excluyentes.
- Colectivamente exhaustivas.

Hay cuatro niveles de escalas de medición.

- **Nivel nominal.** Clases o categorías en las que se está o no incluido. Pej. estado civil, raza.
- **Nivel ordinal.** Las categorías se ordenan de una forma determinada (de mayor a menor, de menor a mayor). Pej. intensidad del dolor (MIR 03-04, 146).
- **Nivel de intervalo.** Existe un orden numérico, la diferencia entre valores sucesivos es siempre la misma. Pej. temperatura.
- **Nivel de razón.** Es la escala que proporciona más información. Tiene las mismas características que la anterior, pero además posee el cero absoluto (no tiene el rasgo que nosotros estamos midiendo). Pej. glucemia, talla.

TEMA 11. ESTADÍSTICA DESCRIPTIVA.

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 facilita un análisis visual. Según la naturaleza de las variables estudiadas se utilizan diferentes tipos de representación.

- **Variables cualitativas.**
 - *Diagrama de rectángulos.*
 - *Diagrama sectorial (pastel).*

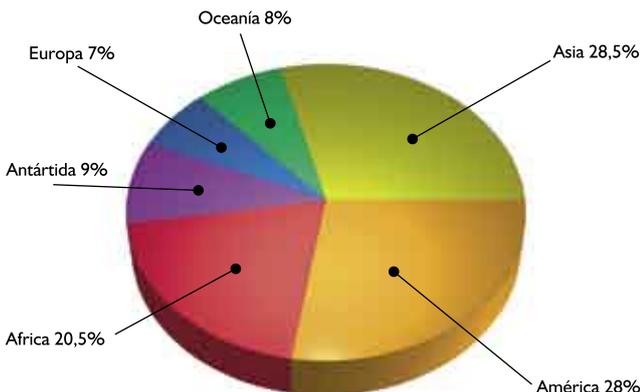


Figura 5. Diagrama sectorial.

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 rectángulos** tienen una base constante y una altura proporcional a la frecuencia absoluta correspondiente (también su área es proporcional a la frecuencia absoluta).

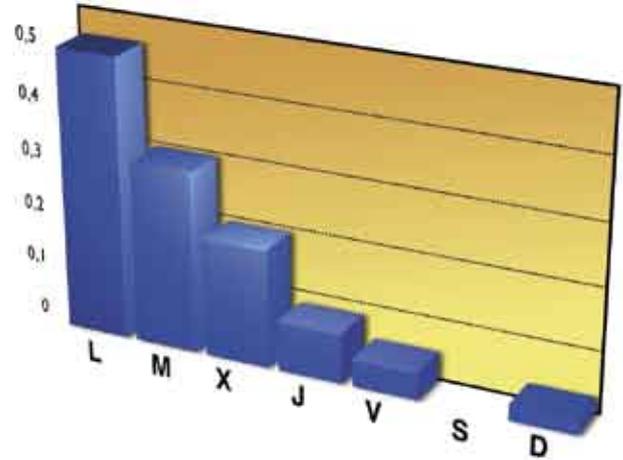


Figura 6. Diagrama de rectángulos.

- **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 se puede representar la frecuencia relativa y/o los porcentajes.
- **Variables cuantitativas continuas.**
 - *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.

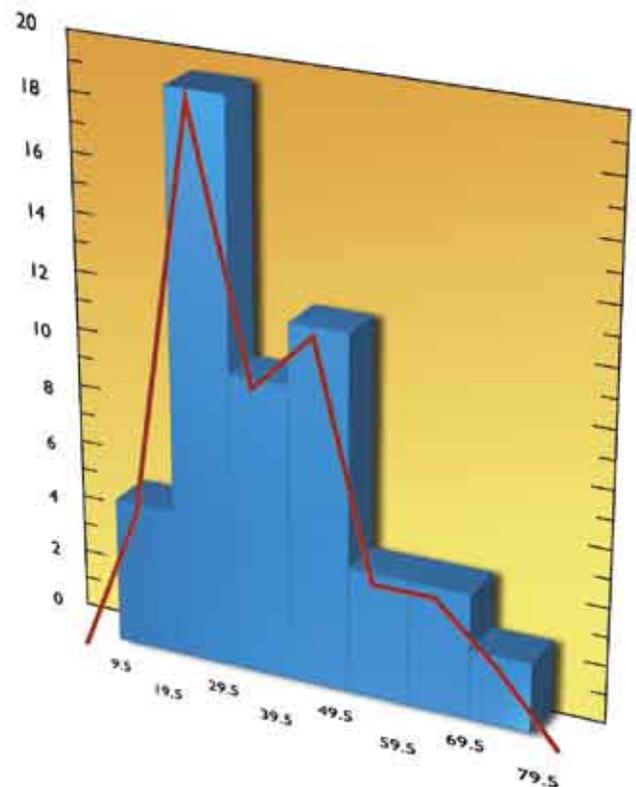


Figura 7. Histograma y polígono de frecuencias.

11.2. Síntesis de los datos.

1. Medidas de localización o tendencia central. Las medidas de centralización nos indican alrededor de qué valores se agrupan los datos observados. Distinguiamos:

- **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.

$$\bar{X} = \frac{\sum Xi}{n}$$

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 "gravidad" 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 50% de los datos mayor. En una distribución simétrica, la mediana coincide con la media aritmética, pero no en una asimétrica (MIR 98-99, 211; MIR 97-98, 77; MIR 96-97E, 204).
- **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.** Valores que dividen el conjunto de las observaciones en cuatro, diez o cien partes iguales. Por ejemplo: $Q2=D5=Pc50=$ Mediana. El 50% de las observaciones serán inferiores al segundo cuartil, quinto decil o percentil 50 (MIR 97-98E, 73; MIR 96-97E, 203; MIR 95-96, 31).
- **Media geométrica.** Muy utilizada en microbiología y serología, cuyos datos tienen una marcada asimetría positiva (hacia la derecha). Ejemplo: títulos de anticuerpos.

COMPARACIÓN DE MEDIDAS DE CENTRALIZACIÓN.

Aunque desde un punto de vista puramente descriptivo las tres 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** utiliza 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 (MIR 95-96, 39).

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

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

- **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.

$$DM = \frac{\sum |Xi - \bar{X}|}{n}$$

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

$$s^2 = \frac{\sum (Xi - \bar{X})^2}{n}$$

- **Desviación típica o estándar.** Es la raíz cuadrada positiva de la varianza. Es, junto con la anterior, la medida de dispersión más usada.

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

La desviación típica es una medida complementaria de la media aritmética; mientras que ésta 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. Es el porcentaje que representa la desviación 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 (MIR 96-97, 153).

$$C.V. = \frac{s}{\bar{X}} \times 100$$

Tabla 7. Formas de medida.

	MEDIDA DE CENTRALIZACIÓN	MEDIDA DE DISPERSIÓN
DISTRIBUCIÓN HOMOGÉNEA	Media	Desviación típica o estándar
DISTRIBUCIÓN ASIMÉTRICA	Mediana	Rango intercuartílico Rango

3. 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.

TEMA 12. PROBABILIDAD.

Supongamos una población finita con N elementos, k de los cuales tienen la característica A. Llamaremos 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.

- La probabilidad es un valor entre 0 y 1, es decir, $0 \leq P(A) \leq 1$.
- Llamaremos suceso seguro E, al que ocurre siempre, entonces: $P(E)=1$.
- Si \bar{A} es el suceso complementario de A, que ocurre siempre que no lo hace A, entonces: $P(\bar{A}) = 1 - P(A)$.
- Dos sucesos son **mutuamente excluyentes** (incompatibles) cuando no pueden suceder a la vez (P.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 depende, no se modifica por la aparición del otro (P. ej.: efectos secundarios de un fármaco en dos pacientes).

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

12.1. 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 ocurre el suceso A. Es la probabilidad de que suceda A, una vez que ya ha sucedido B.

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

12.2. 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:

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

(MIR 98-99, 210; MIR 96-97F, 224).

12.3. Regla de la suma.

Nos indica la probabilidad de que suceda A o de que suceda B.

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

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

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

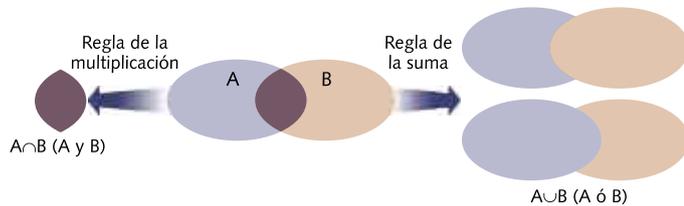


Figura 8. Reglas de la probabilidad.

TEMA 13. DISTRIBUCIONES DE PROBABILIDAD.

Una vez que hemos escogido 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.

Encontramos tres leyes teóricas que tienen la particularidad de ser seguidas por una inmensa mayoría de los fenómenos biológicos habituales:

- Distribución Binomial.
- Distribución de Poisson.
- Distribución normal o de Gauss.

13.1. 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 fracaso". La variable de interés es el número de éxitos alcanzados en "n" pruebas.

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

13.2. Distribución de Poisson.

Constituye un caso particular de la ley binomial para aquellas situaciones en que la probabilidad de aparición de un fenómeno es muy pequeña. Es, por tanto, la ley adecuada para los "sucesos raros", cuando $p < 0,1$ y $n \times p < 5$.

13.3. Distribución normal o de Gauss.

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 medio.

CARACTERÍSTICAS.

- 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 (MIR 96-97F, 220).
- Tiene forma acampanada, sin un pico excesivo.
- Va desde - infinito a + infinito (asintótica al eje de abscisas).
- El área bajo la curva = 1.

El 50% de las observaciones se encuentran por debajo de la media aritmética y el 50% por encima. El 68% de las observaciones se encuentran dentro del intervalo $\bar{X} \pm S$; el 95% dentro del intervalo $\bar{X} \pm 1,96 S$ y el 99% dentro del intervalo $\bar{X} \pm 2,57 S$ (MIR 98-99F, 219; MIR 98-99, 207; MIR 96-97F, 217; MIR 96-97, 146; MIR 95-96, 45).

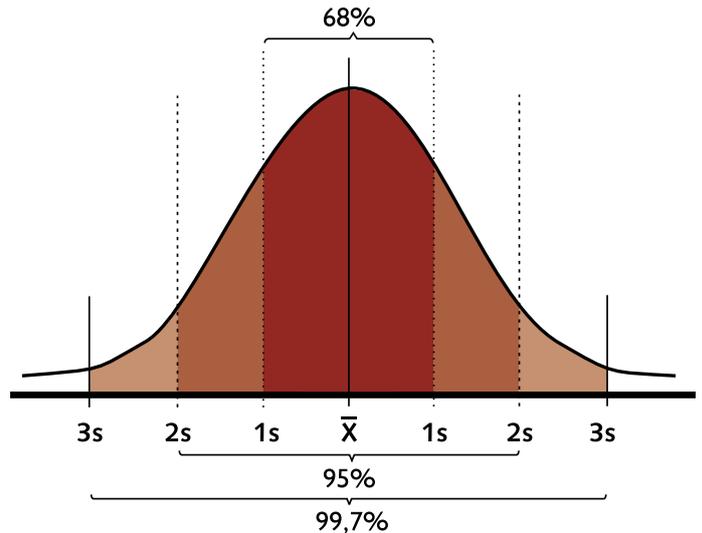


Figura 9. Distribución normal.

Tipificación. La distribución normal viene definida por la media y la desviación típica; pueden existir, por tanto, infinitas curvas de este tipo, tantas como infinitos valores puedan tomar la media y la desviación típica. La distribución normal tipificada tiene una media = 0 y una desviación típica = 1 [N(0,1)] y se encuentra tabulada. Nosotros podemos transformar cualquier variable aleatoria que se distribuya normalmente a una normal tipificada mediante la siguiente transformación:

$$z = \frac{X_i - \bar{X}}{S}$$

TEMA 14. ESTIMACIÓN DE PARÁMETROS.

El propósito general de la inferencia estadística es obtener conocimiento o información de una población a través de una muestra.

14.1. Estimación de medias.

En este caso, queremos conocer la media poblacional a partir de la media muestral. Para ello, podemos aplicar:

1. **Estimación puntual.** La media muestral es un buen estimador de la media poblacional (μ).
2. **Estimación por intervalos.** El valor numérico de los estimadores puntuales determinados en muestras diferentes puede variar, es decir, si repitiéramos los cálculos con otra muestra diferente de la misma población, el resultado de la media muestral podría ser diferente. Por tanto, sería mejor que, además aportáramos un intervalo que presumiblemente incluya también el parámetro de la población. Es por tanto 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.

Error estándar de la media. De una población pueden extraerse infinitas muestras, cada una de ellas con su media. Este conjunto de medias se distribuyen según una curva normal cuya media, la media de las medias, es la media poblacional (μ) y cuya desviación típica se conoce como "EL ERROR ESTÁNDAR DE LA MEDIA"; es, por tanto, la dispersión de las medias muestrales con respecto a la media poblacional (MIR 94-95, 227). Se calcula mediante la fórmula:

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

σ = desviación típica poblacional.

NO TIENE NADA QUE VER LA DESVIACIÓN TÍPICA DE LOS VALORES DE LA VARIABLE EN LA MUESTRA O EN LA POBLACIÓN, CON EL ERROR ESTÁNDAR DE LAS MEDIAS MUESTRALES ALREDEDOR DE LA MEDIA POBLACIONAL.

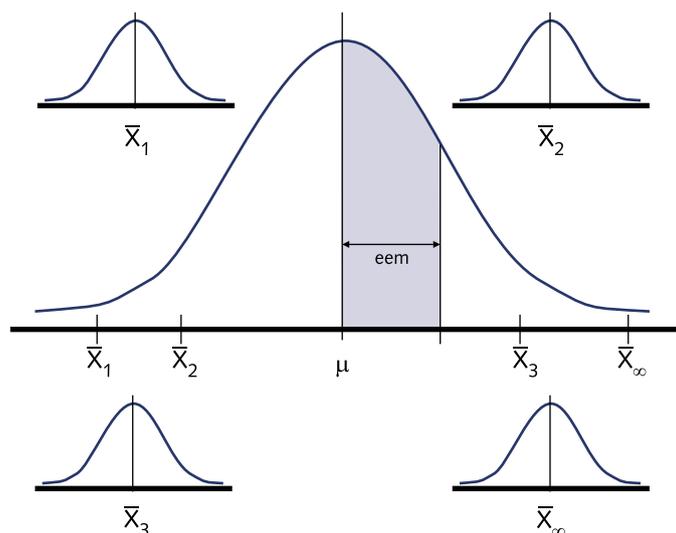


Figura 10. Cálculo del error estándar de la media.

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

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

Mientras que en un 5% de las ocasiones nos equivocaremos.

Si establecemos el nivel de confianza en el 99%, la media poblacional se situará en un 99% de las ocasiones entre:

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

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

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

Donde el valor t corresponde a los valores tabulados de la distribución teórica t-Student (MIR 05-06, 212; MIR 98-99F, 221; MIR 96-97F, 205; MIR 96-97F, 202; MIR 95-96, 28).

14.2. Estimación de un porcentaje.

Si hemos estudiado 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).

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 (MIR 98-99, 204):

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

$$p \pm 2,57 \times eep, p < 0,01, \text{ donde:}$$

- p : proporción (en tantos 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), éste ha de ser extrapolado 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 (MIR 00-01, 199; MIR 99-00F, 211).

TEMA 15. CONTRASTE DE HIPÓTESIS.

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

Para su comprensión partiremos 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 esto es debido a una real mayor eficacia de N o se debe simplemente a 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 nosotros hemos 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 menor 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 . Resultado estadísticamente significativo. Rechazo de H_0 .
- $p > 0,05 \Rightarrow$ No existe suficiente evidencia como para decir que ambos tratamientos son diferentes. Las diferencias pueden deberse al azar con una probabilidad mayor al nivel de exigencia. No se rechaza H_0 (MIR 05-06, 209; MIR 05-06, 208; MIR 05-06, 194; MIR 03-04, 153; MIR 99-00F, 212; MIR 98-99, 208; MIR 97-98, 75; MIR 97-98, 73; MIR 97-98, 68; MIR 97-98F, 65; MIR 96-97, 132; MIR 96-97F, 222; MIR 96-97F, 219; MIR 95-96F, 55; MIR 94-95, 224).

15.2. Errores alfa 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 es. Se denomina α a la probabilidad de cometer el error tipo I. El valor de la "p" coincide con la probabilidad de cometer el error tipo I (MIR 02-03, 35; MIR 01-02, 213; MIR 00-01, 202; MIR 99-00, 230).

Cuando entre los distintos grupos de estudio se comparan más de una variable de resultado, aumenta la posibilidad de resultados

falsamente positivos. Para evitarlo, se aumenta la exigencia del nivel de significación como sigue (MIR 00-01F, 244; MIR 97-98, 68):

$$P=0,05/n^\circ \text{ de comparaciones}$$

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: damos como no significativo algo que lo es. Se denomina β a la probabilidad de cometer un error tipo II (MIR 00-01, 206; MIR 98-99, 203; MIR 95-96F, 61).

Poder o potencia del test. Lo complementario del error β es la "potencia o poder estadístico de un test" $(1 - \beta)$: capacidad que tiene un test de detectar una diferencia cuando ésta existe en realidad, es decir, corresponde a la probabilidad que tengo de demostrar la hipótesis alternativa, siendo cierta (MIR 01-02, 210; MIR 98-99, 209; MIR 97-98F, 75; MIR 95-96, 46).

Tabla 8. Contraste de hipótesis.

		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 alfa
	NO HAY DIFERENCIAS SIGNIFICATIVAS. No rechazo H_0 .	Error tipo II o error beta	$1 - \alpha$

15.3. Pruebas de significación estadística.

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 Mc Nemar.
- Variable resultado cuantitativa: t de Student apareada.
- Variable resultado cuantitativa (varias medidas repetidas): ANOVA para medidas repetidas.

(MIR 03-04, 145; MIR 99-00F, 208; MIR 98-99, 213; MIR 98-99, 212; MIR 97-98, 69; MIR 97-98F, 77; MIR 97-98F, 70; MIR 96-97F, 200; MIR 95-96, 47; MIR 95-96F, 57; MIR 94-95, 234, MIR 94-95, 232).

Tabla 9. Test de contraste de hipótesis.

VARIABLE 1	VARIABLE 2	TEST DE HIPOTESIS
Dicotómica	Dicotómica	Chi-cuadrado. Test exacto de Fisher.
Cualitativa (> 2 categorías)	Cualitativa (>= 2 categorías)	Chi-cuadrado.
Dicotómica	Cuantitativa	t de Student.
Cualitativa (> 2 categorías)	Cuantitativa	Análisis de la Varianza.
Cuantitativa	Cuantitativa	Coficiente de correlación de Pearson. Regresión.

2) **Test no paramétricos.** Son pruebas que no dependen de la distribución de la variable y no precisan condición previa de normalidad. 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 lo 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 solamente una medida aproximada del carácter, sin una valoración exacta. Si las muestras son mayores de 30, no existe inconveniente en utilizar pruebas paramétricas (MIR 94-95, 230; MIR 05-06, 193).

Tabla 10. Test no paramétricos.

VARIABLE 1	VARIABLE 2	TEST DE HIPÓTESIS
Dicotómica independiente.	Cuantitativa (n <30) u ordinal.	U de Mann-Whitney. Test de Wilcoxon.
Dicotómica pareada o dependiente.	Cuantitativa (n <30) u ordinal.	Test de Wilcoxon.
Cualitativa (>2 categorías).	Cuantitativa (n <30) u ordinal.	Test de Kruskal-Wallis.
Cualitativa pareada (>2 categorías).	Cuantitativa (n <30) u ordinal.	Test de Friedman.
Ordinal.	Ordinal.	Rho de Spearman. Tau de Kendall.

3) **Análisis multivariante.** Son un conjunto de test estadísticos que se aplican cuando se intenta establecer la asociación estadística entre tres o más variables. Ejemplo: se desea saber si existe una relación entre la infección de herida quirúrgica, la administración de profilaxis y el tipo de cirugía (MIR 01-02, 214; MIR 95-96F, 50).

TEMA 16. DISTRIBUCIÓN CHI-CUADRADO.

Test de independencia. Se contrasta la hipótesis nula de que dos criterios de clasificación (variables cualitativas), cuando se observan en la misma serie de individuos u objetos, son independientes.

- H_0 . No hay asociación entre las variables.
- H_1 . Sí hay asociación entre las variables.

Test de homogeneidad. Se estudia una variable en dos o más muestras. La pregunta es: ¿son las muestras extraídas de poblaciones, que son homogéneas con respecto a alguna variable?

- H_0 . Las muestras se extraen de la misma población.
- H_1 . Las muestras provienen de distinta población.

Ejemplo. Estudiamos una muestra de 500 niños para saber si existe asociación entre el estado nutricional (pobre, bueno) y los resultados académicos (malos, satisfactorios). Los datos se muestran en la tabla:

	POBRE	BUENO	TOTAL
Malos	105	15	120
Satisfactorios	80	300	380
Total	185	315	500

Hipótesis.

- H_0 . El estado nutricional y los resultados académicos son independientes.
- H_1 . Las dos variables no son independientes.
- $\alpha= 0,05$

Test estadístico.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Decisión. Rechazamos H_0 si el valor de la χ^2 calculado es igual o mayor que el tabulado.

Cálculo del estadístico χ^2 .

- 1) Obtención de las frecuencias esperadas: para cada casilla, el valor de la frecuencia esperada es total filas x total columnas / total.

En este caso:

	POBRE	BUENO	TOTAL
Malos	44,4	75,6	120
Satisfactorios	140,6	239,4	380
Total	185	315	500

$$\chi^2 = (105-44,4)^2/44,4 + (80-140,6)^2/140,6 + (15-75,6)^2/75,6 + (300-239,4)^2/239,4$$

$\chi^2 = 177,7$, que es mayor que el valor de la tabla (3,84, $p < 0,05$).

Rechazamos H_0 , ya que el valor calculado es superior al tabulado, y concluimos que existe asociación entre el estado nutricional y los resultados académicos ($p < 0,05$).

En la tabla 2 x 2 de valores teóricos o esperados alguna casilla tiene menos de 5 elementos en vez de la prueba de Chi2, se deberá utilizar el test exacto de Fisher.

TEMA 17. ANÁLISIS DE CORRELACIÓN Y REGRESIÓN.

Sirven para estudiar la relación entre dos variables cuantitativas.

Análisis de regresión. Es útil para determinar la posible forma de la relación entre variables y, por tanto, se puede usar para hacer predicciones o estimar el valor de una variable que corresponde para un valor de la otra. En este análisis disponemos 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 llama dependiente.

Con el análisis de regresión lineal obtenemos una ecuación del tipo:

$$Y = \alpha + \beta X, \text{ donde } \alpha \text{ y } \beta \text{ son los coeficientes de regresión.}$$

El coeficiente α representa el punto en el que la línea corta el eje vertical (valor de Y para X = 0). El coeficiente β es la pendiente de la recta que muestra la cantidad que cambia Y por una unidad de cambio de X (MIR 97-98F 71).

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 a medida que disminuye una también lo hace la otra. Un coeficiente de correlación negativo indica que a medida que disminuye una variable aumenta la otra o viceversa.
- $r = 0$ indica que no hay correlación lineal.

(MIR 98-99, 206; MIR 98-99F, 222; MIR 97-98F, 82; MIR 96-97, 154; MIR 96-97F, 199).

TEMA 18. 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 y los condicionamientos que se está dispuesto a asumir. Un número insuficiente de participantes impedirá encontrar las diferencias buscadas, concluyendo erróneamente que no existen, y un número excesivo aumenta innecesariamente el coste. Manejar muestras es más barato, rápido, fácil y suficientemente preciso y exhaustivo. 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 vamos a desarrollar el cálculo del tamaño muestral:

- **Estimación de parámetros.** El tamaño dependerá de la variabilidad del parámetro que se desea estimar (ésta debe ser previamente conocida, o bien, aproximarse a partir de datos preexistentes o estudios pilotos), de la precisión con que se desee obtener la estimación (amplitud deseada del intervalo de confianza, teniendo en cuenta que a mayor precisión, mayor número de sujetos se necesitarán) y del nivel de confianza.
- **Contraste de hipótesis:** se utiliza para estudios comparativos y 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:
 - Objetivo del estudio y variable principal.
 - Magnitud del efecto que queremos detectar (delta).
 - Variabilidad de la variable principal.
 - Error tipo I o alfa asumido (0,05).
 - Error tipo II o beta asumido (0,2-0,1).
 - Poder estadístico (1-beta:80-90%).
 - Proporción de pacientes en los distintos grupos.
 - Proporción de pérdidas (d): deberemos multiplicar el tamaño calculado por: $1/(1-d)$.
 - Tipo de comparación (test unilateral o de una cola cuando se presupone que la diferencia entre los grupos será siempre a favor de uno de ellos; test bilateral o de dos colas cuando no se presupone la dirección de la diferencia; el test bilateral es más conservador y requiere de mayor tamaño muestral).

(MIR 05-06, 203; MIR 02-03, 32; MIR 02-03, 37; MIR 01-02, 218; MIR 00-01F, 239; MIR 00-01, 205; MIR 00-01, 207; MIR 99-00F, 199; MIR 98-99, 205; MIR 98-99, 197; MIR 98-99F, 220; MIR 96-97F, 218; MIR 94-95, 221).

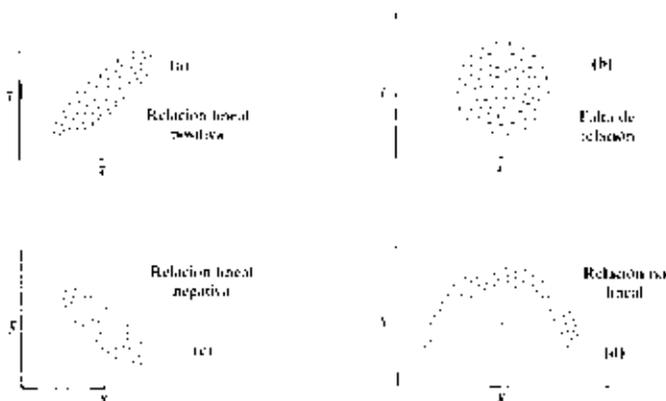


Figura 12. Distintos tipos de relación entre las variables.