

Curso OPE en Enfermería

INVESTIGACIÓN Y ENFERMERÍA BASADA EN LA EVIDENCIA



AMIR

**MANUAL AMIR OPE
INVESTIGACIÓN Y ENFERMERÍA BASADA
EN LA EVIDENCIA
2.ª edición**

ISBN
FALTA

DEPÓSITO LEGAL
M-0000-2023

ACADEMIA DE ESTUDIOS MIR, S. L.
www.academiamir.com
info@academiamir.com

DISEÑO E ILUSTRACIONES
Equipo de Diseño y Maquetación AMIR

MAQUETACIÓN E ILUSTRACIONES
Equipo de Diseño y Maquetación AMIR

COORDINACIÓN EDITORIAL:

Suarez Barrientos, Aida
López Frutos, Alicia
Orihuel Pérez-Klett, Álvaro
Solana Ramírez, Miguel
Hernanz Milanesi, Laura
Salmerón López, Susana

La protección de los derechos de autor se extiende tanto al contenido redaccional de la publicación como al diseño, ilustraciones y fotografías de la misma, por lo que queda prohibida su reproducción total o parcial sin el permiso del propietario de los derechos de autor.



Este manual ha sido impreso con papel ecológico, sostenible y libre de cloro, y ha sido certificado según los estándares del FSC (Forest Stewardship Council) y del PEFC (Programme for the Endorsement of Forest Certification).



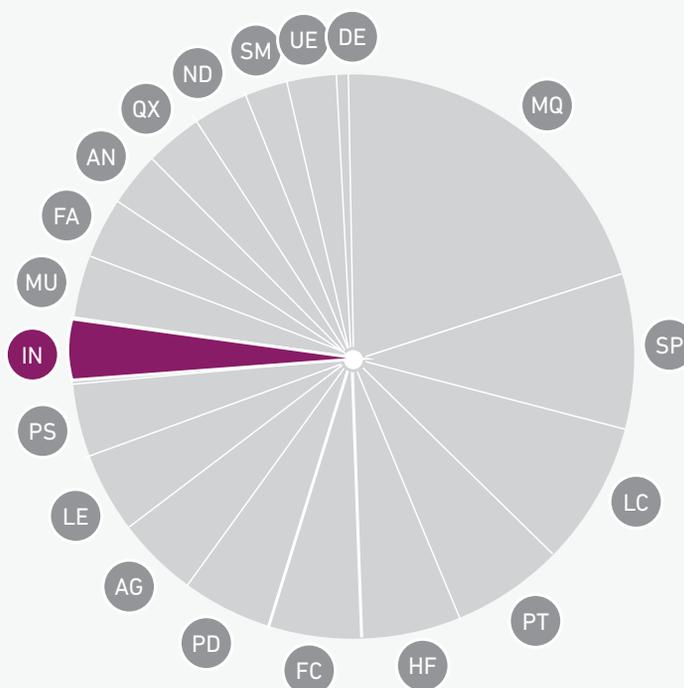
IN

AMIR

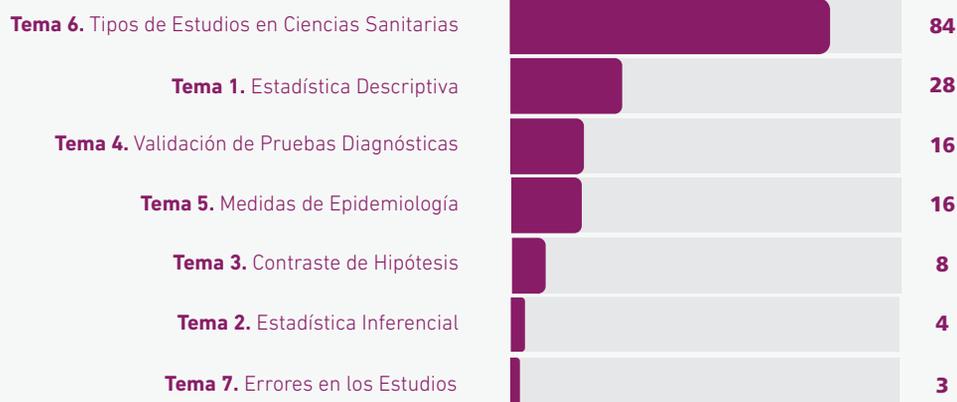
ORIENTACIÓN OPE

Investigación y Enfermería basada en la Evidencia es una asignatura que, a pesar de que en general resulta densa y pesada de estudiar, es agradecida de cara a los exámenes porque agrupa un buen número de preguntas, ocupando el décimo lugar en el orden de importancia. La gran mayoría de las preguntas son relativas al tema 6 de Tipos de Estudios en Ciencias de la Salud. En este tema veremos los diferentes diseños de las metodologías cuantitativas y cualitativas con sus aspectos clave, además del proceso de diseño de los estudios, lectura crítica, etc. Después le sigue el tema 1 de Estadística Descriptiva, en el que se encuentra el estudio de las variables, las distribuciones y las medidas de análisis (media, moda, mediana, desviación típica, varianza...). Compensa estudiar bien los temas 5 sobre Medidas en Epidemiología, donde aprenderás cuestiones sobre la prevalencia e incidencia; y el tema 4 de Validación de Pruebas Diagnósticas, especialmente los conceptos de sensibilidad, especificidad y las tasas de falsos y verdaderos positivos. El resto de temas son poco preguntados. En esta asignatura no encontramos preguntas específicas de ninguna Comunidad Autónoma.

IMPORTANCIA DE LAS ASIGNATURAS



IMPORTANCIA DE LOS TEMAS



INCIDENCIA DE LOS TEMAS POR COMUNIDADES

INVESTIGACIÓN Y ENFERMERÍA BASADA EN LA EVIDENCIA		Andalucía		Aragón		Canarias		Cantabria		C. y León		C. La Mancha		Cataluña		C. de Madrid		C.F. Navarra		C. Valenciana	
		2021	2019	2022	2019	2019	2016	2019	2015	2019	2016	2022	2018	2019	2019	2014	2022	2018	2022	2018	
Tema 1	Estadística Descriptiva	1	1			3		1		1					2	2	2			1	
Tema 2	Estadística Inferencia														1						
Tema 3	Contraste de Hipótesis	1			1			1		1							1				
Tema 4	Validación de Pruebas Diagnósticas			1									2			2	1		1		
Tema 5	Medidas en Epidemiología	1				1						1				1			1	1	
Tema 6	Tipos de Estudios en Ciencias Sanitarias	2	4	1	3	3	2	2	2	2	2	3	2		1	7	1	3		3	
Tema 7	Errores en los Estudios							1													

INVESTIGACIÓN Y ENFERMERÍA BASADA EN LA EVIDENCIA		Extremadura		Galicia		Balears		Ceuta - Melilla		La Rioja		País Vasco		Asturias		C. de Murcia		TOTAL	
		2022	2019	2019	2016	2019	2009	2021	2019	2021	2019	2022	2018	2019	2013	2019	2018	159	TOP 10
Tema 1	Estadística Descriptiva	1				1	1	4	2	2					1	2			28
Tema 2	Estadística Inferencia									1	1	1		1		1			4
Tema 3	Contraste de Hipótesis														1				8
Tema 4	Validación de Pruebas Diagnósticas	1	1			1	1			2	2		1						16
Tema 5	Medidas en Epidemiología	1	1		2	1				1	1	2			1				16
Tema 6	Tipos de Estudios en Ciencias Sanitarias	3	2	1		3	5	1	3	2	3	4	3	3	1	4	3		84
Tema 7	Errores en los Estudios				1			1											3



ÍNDICE

ESTADÍSTICA

TEMA 1. ESTADÍSTICA DESCRIPTIVA	9
1. Técnicas de muestreo	
2. Tipos de variables	
3. Medidas de análisis de los datos	
4. Principales distribuciones de probabilidad	
TEMA 2. ESTADÍSTICA INFERENCIAL	17
1. Estadística inferencial para variables cuantitativas	
2. Estadística inferencial para variables cualitativas	
3. Cálculo del tamaño muestral para estudio de inferencias	
TEMA 3. CONTRASTE DE HIPÓTESIS	19
1. Errores en contraste de hipótesis	
2. Test para contraste de hipótesis	

EPIDEMIOLOGÍA

TEMA 4. ESTUDIOS DE VALIDACIÓN DE UNA PRUEBA DIAGNÓSTICA	23
1. Parámetros de validez de una prueba diagnóstica	
2. Curvas ROC (de rendimiento diagnóstico)	
TEMA 5. MEDIDAS EN EPIDEMIOLOGÍA	27
1. Criterios de causalidad de Bradford Hill	
2. Medidas de frecuencia de una enfermedad	
3. Medidas de fuerza de asociación (medidas de efecto)	
4. Medidas de impacto	
TEMA 6. TIPOS DE ESTUDIOS EPIDEMIOLÓGICOS	33
1. Enfoques de investigación científica	
2. Estudios cualitativos	
3. Tipos de estudios epidemiológicos con enfoque cuantitativo en función de sus objetivos	
4. Estudios observacionales	
5. Estudios experimentales	
6. Niveles de evidencia científica	
7. Fases de desarrollo de un estudio cuantitativo	
8. Fases de desarrollo de un tratamiento (fases del ensayo clínico)	
9. Diseño paralelo y diseño cruzado en los estudios experimentales	
10. Propuestas de investigación	
Referencias bibliográficas	
TEMA 7. ERRORES EN LOS ESTUDIOS EPIDEMIOLÓGICOS	49
1. Errores aleatorios	
2. Errores sistemáticos (sesgos)	



Estadística descriptiva

Alba M^a Romero López, José Antonio Rueda Camino

1. Técnicas de muestreo.....	10
2. Tipos de variables	11
3. Medidas de análisis de los datos	13
4. Principales distribuciones de probabilidad.....	14

Enfoque OPE

Tema ampliamente preguntado. Lo más importante es el diagnóstico (tanto dar un diagnóstico a una gasometría en concreto, como asignar una gasometría a un diagnóstico), para ello debes seguir paso a paso el algoritmo del final del tema. Debes conocer también la clínica de cada trastorno y sus mecanismos compensadores.

El objetivo de la Estadística es el estudio de una o varias características (variables) en una o varias poblaciones diana. Habitualmente el estudio de todos los individuos de dichas poblaciones es imposible por problemas logísticos, así que se suele estudiar sólo a un grupo reducido de individuos de cada población (muestra).

La **estadística descriptiva** se encarga de estudiar las diferentes variables que nos interesan en la **muestra**. Como esta muestra es accesible y podemos estudiar a todos los individuos, todos los **datos** obtenidos serán **verídicos** para dicha muestra, y no tendremos que extrapolar resultados; es por esto que en estadística descriptiva **no existe probabilidad de cometer errores de inferencia**.

La **estadística inferencial** intenta **extrapolar** a partir de los resultados de la muestra cómo serían los resultados de la **población** si fuéramos capaces de estudiar a todos sus individuos. Así, los resultados estarán sujetos a una **probabilidad de error**, ya que al escoger una parte (muestra) del todo (población in-

vitiblemente se producirá un alejamiento(error) entre el valor de la población y la muestra.

Por último, el **contraste de hipótesis** compara los resultados de varias variables en una única población, o bien los resultados obtenidos para la misma variable en varias poblaciones. Al igual que en estadística inferencial, para obtener los datos poblacionales se parte de resultados de las muestras estudiadas, por lo que existe **probabilidad de error**.

(Ver figura 1)

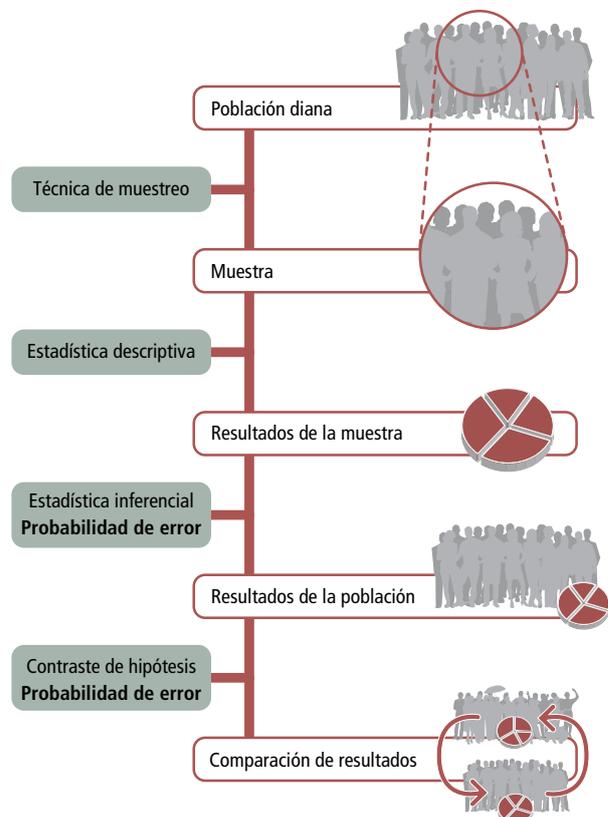


Figura 1. Esquema de realización de un estudio estadístico.



1. TÉCNICAS DE MUESTREO

El muestreo consiste en la selección de una muestra a partir de una población. El objetivo del muestreo es que la muestra escogida sea **representativa** de la población, para que los resultados obtenidos en la muestra sean extrapolables a la población.

Las mejores técnicas de muestreo son las **probabilísticas** (emplean el azar), ya que la utilización del azar para escoger la muestra (en lugar de cualquier criterio que defina el investigador) hace que existan más probabilidades de que la muestra sea representativa de la población. En las técnicas de muestreo aleatorias todos los individuos de la población tienen las mismas probabilidades de ser elegidos para formar parte de la muestra.

Además de la aleatorización, para poder extrapolar los datos a la población, es necesario que nuestra muestra sea homogénea (**IBSALUT 09, 23**).

Técnicas de muestreo probabilístico

Muestreo aleatorio simple

Primero asignamos un número a cada uno de los posibles participantes de la población, y posteriormente se escogen aleatoriamente tantos números como sujetos sean necesarios para la muestra.

Ejemplo: para obtener una muestra de 5 individuos en una población de 100 personas, se asigna a cada persona un número del 1 al 100. Se introducen en una urna 100 pelotas numeradas, y se sacan de la urna 5 pelotas.

Muestreo aleatorio sistemático

Esta técnica de muestreo combina el azar y una regla sistemática para seleccionar a los sujetos. Para emplearla, se selecciona por azar al primer sujeto (al igual que en la anterior los sujetos están numerados), y a ese primer número se le suma una constante o k para ir obteniendo el resto de sujetos.

Por ejemplo, estamos interesados en estudiar qué variables afectan a la calidad de vida de pacientes con una determinada enfermedad rara en España, habiendo 25 casos en toda España (la población, $N = 25$). No obstante, con los recursos económicos, asignados a nuestro proyecto de investigación, sólo podemos incluir a una muestra de 5 sujetos ($n = 5$) con dicha enfermedad. Para el cálculo de dicha constante se emplea la siguiente fórmula: $N / n = k$; por lo que k será igual a 5 ($25 / 5 = 5$) en nuestro ejemplo. La estrategia aleatoria se empleará para extraer al primer sujeto de entre esos k primeros pacientes. Así pues, siguiendo nuestro ejemplo, de entre esos 5 sujetos primeros, al aplicar el azar se extrae, por ejemplo, al paciente n.º 2, siendo éste el primer participante de los 5 que conformarán nuestra muestra. Los 4 participantes siguientes se elegirán sumando dicha k (5 en nuestro ejemplo) de forma consecutiva; de manera que el segundo

participante será el n.º 7 (2+5); el tercer participante será el n.º 12 (7+5); el cuarto participante será el n.º 17 (12+5) y el quinto y último participante será el n.º 22 (17+5).

Muestreo por conglomerados

Se recurre a las aglomeraciones naturales que hay en la población (ciudades, barrios, etc.) y se selecciona aleatoriamente uno o varios de estos conglomerados como muestra.

Muestreo por etapas o polietápico (**SMS 18, 51**)

El **muestreo por etapas o polietápico** es un subtipo de muestreo que obtiene la muestra final a partir de sucesivas etapas, seleccionado aleatoriamente subconglomerados, los cuales están contenidos en unidades superiores de agrupación de sujetos.

Estratificación (**INGESA 19, 29**)

La estratificación consiste en la división de la población en diferentes categorías o estratos de la variable a estudio, para aplicar un tipo de muestreo sobre cada estrato.

Independientemente de la técnica de muestreo elegida, la estratificación nos puede ayudar a controlar una variable que no queremos que influya en nuestros resultados (posible factor de confusión).

Ejemplo: nos interesa contrastar si la polución ambiental produce cáncer de pulmón, pero no queremos que el tabaco (posible factor de confusión) interfiera en nuestros resultados. Así, antes de escoger la muestra dividimos a la población en, por ejemplo, 3 estratos en función de la variable "tabaquismo" (fumadores, exfumadores, no fumadores), y posteriormente haremos el muestreo sobre cada uno de estos estratos.

Técnicas de muestreo no probabilístico

En este tipo de técnicas de muestreo no interviene el azar, y se van a utilizar fundamentalmente en la investigación cualitativa, ya que el interés de esta no radica en la representatividad de la muestra ni en la extrapolación de los datos recogidos, sino en la comprensión de un objeto de estudio para construir una teoría acerca del mismo.

Muestreo teórico

En el muestreo teórico la selección de sujetos a incorporar lo decide el investigador a partir del conocimiento teórico que se va construyendo con la información que ha emergido del estudio (teoría fundamentada), es decir, basándose en el análisis de los datos recogidos y en la pertinencia y saturación teórica (por más datos que recojamos, estos no aportarían nada nuevo).

Muestreo por conveniencia

Consiste en seleccionar a aquellos participantes que están disponibles y receptivos a participar en el estudio.

Muestreo a propósito

Cuando se selecciona específicamente a aquellos sujetos que reúnen una serie de características particulares que son relevantes para el objeto de estudio.

Muestreo por bola de nieve (SARS 19, 71)

Cuando tenemos difícil acceso a un perfil de participantes y recurrimos a los conocidos de los participantes que ya tenemos captados, para que sean estos quienes los impliquen en la investigación.

Ejemplo: estamos realizando un estudio sobre cuidados paliativos pediátricos y los cuidadores principales de los niños y, tras obtener a la primera familia, esta nos pone en contacto con otra, y así sucesivamente, como si fuera una bola de nieve que al rodar se hace más grande.



2. TIPOS DE VARIABLES (CANAR 19, 28)

VARIABLES CUALITATIVAS

Características que **no se expresan numéricamente**, como el color de pelo, la raza o el estado civil.

- ✓ **VARIABLES CUALITATIVAS ORDINALES:** cuando los distintos valores de una variable cualitativa siguen un orden, nos interesará asignar a cada valor un número arbitrario (que nos inventamos) en función del orden que ocupa cada categoría. Esto es así porque los test estadísticos que se utilizan para las variables que se expresan con números son más potentes que los test empleados para variables cualitativas "puras". Se distinguen de las variables cuantitativas en que los números asignados no cumplen propiedades matemáticas.

Ejemplo: escala del dolor: leve = 1, moderado = 2, intenso = 3. Tener un dolor "2" no significa tener el doble de dolor que un dolor "1".

- ✓ **VARIABLES CUALITATIVAS NOMINALES (SAS 21, 22; SCS 19, 39):** los valores no siguen un orden, por lo que los nombraremos con palabras y no números (por ejemplo, el color de pelo).

Si la variable cualitativa solo puede tener **dos valores** (por ejemplo, sexo: masculino o femenino) se denomina **dicotómica**. Si puede tener **más de dos valores** se denomina **policotómica**.

VARIABLES CUANTITATIVAS

Características que se expresan con **valores numéricos** que cumplen las propiedades matemáticas de los números (por ejemplo, tener 4 hijos es tener el doble de 2 hijos).

- ✓ **VARIABLES CUANTITATIVAS DISCRETAS (SERMAS 14, 22; INGESA 19, 9):** los valores numéricos no pueden adoptar cualquier valor (en general, sólo podrán ser números enteros).

Ejemplo: número de pacientes atendidos en un día en una consulta, número de hijos de una familia. Se pueden atender 23 o 24 pacientes, pero no 23,5 pacientes. ¡Ojo! Al trabajar con estas variables, por ejemplo al calcular la media, sí podríamos obtener decimales.

- ✓ **VARIABLES CUANTITATIVAS CONTINUAS (SMS 18, 80; OSASUN 22, 93):** los valores numéricos pueden adoptar cualquier valor, incluyendo decimales.
- Ejemplo: presión arterial: si tuviera un aparato lo suficientemente preciso podría indicar una PAS de 140,6 mmHg. ¡Ojo! Aunque habitualmente sólo utilizemos una variable con números enteros, debemos pensar si sería posible dar un valor con decimales de dicha variable.*

VARIABLES DEPENDIENTES E INDEPENDIENTES

Las **variables dependientes** o de **resultado** son aquellas variables que serán **objeto de interés** para nuestro estudio de investigación, ya que presentarán variaciones en función de otras, las denominadas independientes (**AVS 19, 58**).

Por otro lado, tenemos a las **variables independientes**, que son aquellas que van a determinar a la variable independiente, **ocasionando cambios** en la población estudiada.

ESCALAS DE MEDICIÓN

- ✓ **Escala cualitativa ordinal:** para variables ordinales.
- ✓ **Escala cualitativa nominal:** para variables nominales.
- ✓ **Escala cuantitativa de relación (de proporción):** para aquellas variables cuantitativas en las que el "0" representa ausencia de la variable (p. ej., la altura: medir "0" cm significa ausencia de altura).
- ✓ **Escala cuantitativa de intervalo:** en estas variables el "0" no representa ausencia, sino un punto de corte que se asigna por convenio. *Por ejemplo, 0° Celsius no implica ausencia de temperatura, sino la temperatura de fusión del hielo.*

CONCEPTO DE VARIABLE ALEATORIA

Una variable es **aleatoria** si los posibles valores que puede tomar son determinados por el azar; no sabemos qué valor va a tomar la variable antes de que

aparezca, pero podremos conocer las probabilidades de que se obtenga cada uno de los posibles valores.

Ejemplo: la variable "muerte hospitalaria" (dicotómica: sí / no) es una variable aleatoria; es el azar el que finalmente decide qué pacientes mueren y cuáles no, pero conociendo los factores de riesgo del paciente para morir podremos calcular las probabilidades de que el suceso ocurra. Otro ejemplo más claro sería la variable "tirar un dado de 6 caras" (cuantitativa discreta, del 1 al 6); no sabemos lo que va a salir antes de tirar el dado, pero tenemos 1/6 de probabilidades de obtener cada número.

Representación gráfica de las variables

Representación de variables **cuantitativas y cuantitativas discretas**:

- ✓ **Diagrama de barras:** representa rectángulos separados unos de otros, tantos como valores tenga la variable a estudio. La altura de estos rectángulos será proporcional a la frecuencia relativa (porcentaje) o absoluta (número de observaciones) (SMS 19, 96).

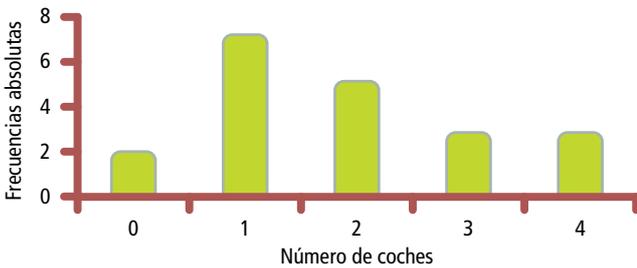


Figura 2. Diagrama de barras.

- ✓ **Pictograma:** exactamente igual que el diagrama de barras, pero los rectángulos se sustituyen por dibujos ilustrativos de cada categoría. (ver figura 3).

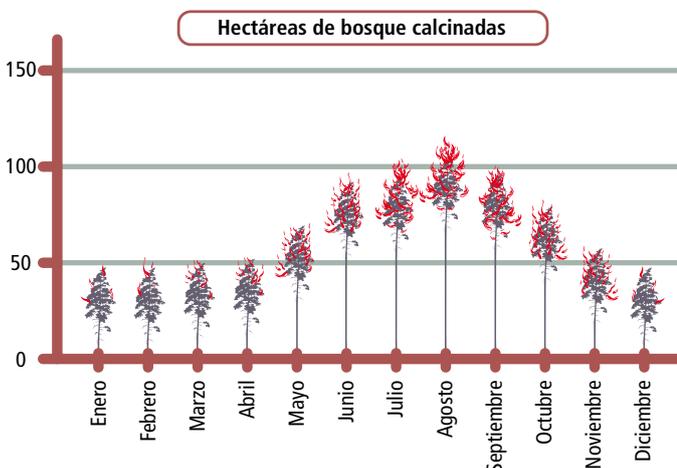


Figura 3. Pictograma.

- ✓ **Diagrama de sectores (SRS 19, 47):** se trata de un círculo dividido en porciones cuya superficie es proporcional a la frecuencia de cada valor de la variable (ver figura 4).

Pronóstico en intervención quirúrgica

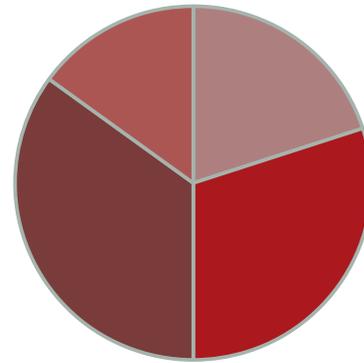


Figura 4. Diagrama de sectores.

Representación de variables **cuantitativas continuas**:

- ✓ **Histograma (de frecuencias no acumuladas):** a diferencia del diagrama de barras, en el histograma cada rectángulo representa un **intervalo de valores**, de manera que cada rectángulo estará unido al anterior. La altura de cada uno de los rectángulos será proporcional a la frecuencia absoluta de cada uno de tales intervalos. Hay histogramas donde cada intervalo de valores tiene la misma amplitud, frente a otros

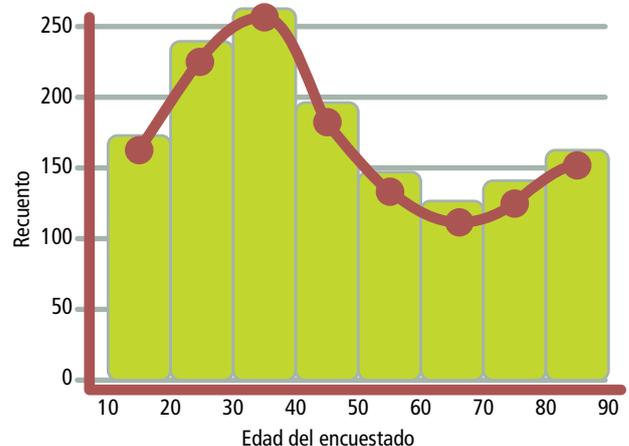


Figura 5. Histograma (verde) y polígono de frecuencias (marrón).

histogramas donde dicha amplitud del intervalo es distinta, siendo la anchura de cada rectángulo proporcional a la frecuencia absoluta del intervalo.

- ✓ **Polígono de frecuencias:** se obtiene de la unión del punto medio del vértice superior de cada barra de un histograma.
- ✓ **Gráficos de Kaplan-Meier:** en este gráfico la horizontal refleja el tiempo, y la vertical representa la probabilidad de permanecer en el estudio en un momento dado después de eliminar a los que han presentado el evento de interés (muerte, progresión de enfermedad, curación...) y aquellos que salen por otros motivos (sujetos censurados). Como resultado, obtendremos una línea escalonada, en la que cada escalón representa una salida de sujetos por evento de interés. Los sujetos censurados se reflejan mediante líneas verticales.

Por ejemplo: en un estudio se comparan los fármacos para el cáncer A y B. El suceso de interés es "progresión del cáncer". En el gráfico de Kaplan-Meier de la figura 6 veis las líneas correspondientes a cada tratamiento. Cada escalón hacia abajo significa que, al menos, un paciente ha salido del estudio en ese momento por progresar su cáncer. Cada línea vertical representa que un paciente ha salido del estudio por cualquier otro motivo (por ejemplo: se retira del estudio por voluntad propia, se muda a otro país...).

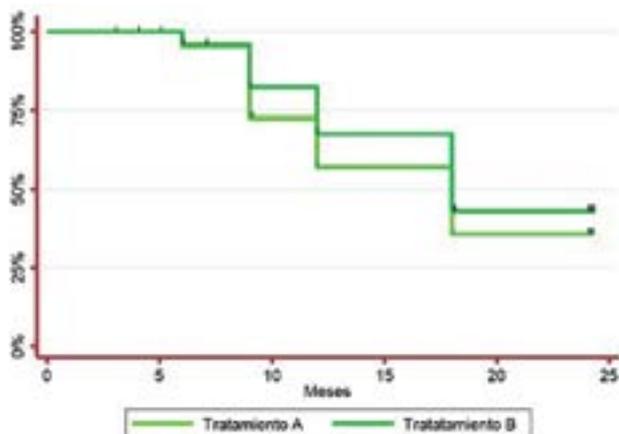


Figura 6. Gráfico de Kaplan-Meier comparando dos tratamientos.

3. MEDIDAS DE ANÁLISIS DE LOS DATOS

Las variables **cualitativas** se suelen expresar mediante **porcentajes** (indicando el porcentaje de observaciones que presenta cada categoría de la variable), y no tienen medidas de dispersión.

Sin embargo, las **variables cuantitativas** se deben expresar mediante una **medida de tendencia central** y una **medida de dispersión**. Además, existen **medidas**

de posición para indicarnos el lugar que ocupa cada observación dentro de la distribución.

Medidas de tendencia central (CANAR 19, 30)

Informan acerca de cómo se agrupan los distintos valores registrados de los individuos de la muestra, indicando dónde se encuentra el centro de la distribución.

- ✓ **Media aritmética (SRS 21, 61):** es la más utilizada y se debe emplear en **distribuciones simétricas**. Reflejaría el centro de gravedad del conjunto de valores, por lo que está muy influenciada por los valores extremos (tanto altos como bajos). Es por esto último que no debe utilizarse en distribuciones asimétricas.

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

- ✓ **Mediana:** es el valor de la variable que presenta el individuo que ocupa la posición central si ordenamos las observaciones de menor a mayor, esto es, divide al conjunto de observaciones en dos partes iguales (deja la mitad de las observaciones por encima y la mitad por debajo). Si la distribución de valores es simétrica, coincide con la media. Es la más indicada si los datos a analizar tienen una distribución **asimétrica** o presentan valores extremos.
- ✓ **Moda (SES 22, 24):** es el valor más repetido de todos los valores de la variable registrada. Puede haber una (**unimodal**), varias (**multimodal**) o ninguna. Es un valor que resulta útil cuando en la distribución hay varios picos de frecuencia, es decir, varias modas.

Ejemplo: sea la siguiente distribución: 5, 7, 7, 7, 8, 10, 12, 13, 15, 22. La media será = (5+7+7+7+8+10+12+13+15+22) / 10 = 10,6. La mediana será el valor que ocupa la posición central de la distribución; en este caso hay dos valores centrales (ya que el número de observaciones es par), 8 y 10 (dejan cuatro observaciones a su izquierda y cuatro a su derecha); la mediana será la media de dichos valores = (8+10) / 2 = 9. La moda es el valor que más se repite, el 7.

Medidas de dispersión (SAS 19, 27)

Cuando analizamos los resultados de una variable cuantitativa en una muestra de sujetos, no sólo nos interesa en torno a qué valor se agrupan los resultados obtenidos (medida de tendencia central), sino también si las observaciones se encuentran "cerca" o "lejos" del centro de la distribución. Este dato lo indican las medidas de dispersión.

Para las variables de distribución **simétrica** se utiliza la **desviación típica** (también llamada desviación estándar), y para las variables de distribución **asimétrica** el **rango y el intervalo intercuartílico**.

Ejemplo: la media de presión arterial sistólica de una muestra de pacientes puede ser de 130 mmHg porque la mitad tiene 129 mmHg y la otra mitad 131 mmHg (esta muestra tiene una PAS muy bien controlada), pero también puede ser 130 mmHg porque la mitad de pacientes tenga 90 mmHg y la otra mitad 170 mmHg (a pesar de tener la misma media, esta muestra es muy diferente de la otra, ya que los valores individuales están muy "alejados" del centro).

- ✓ **Desviación típica (desviación estándar, σ) (CANAR 19, 31; INGESA 19, 22; SERMAS 19, 46):** es la media de la diferencia que existe entre cada observación individual realizada y la media aritmética de la distribución. Se obtiene a partir de la raíz cuadrada de la **varianza (σ^2)**, que es la media del **cuadrado** de dichas diferencias.

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

Tanto la varianza como la desviación típica van a tener siempre **valores positivos**, puesto que la varianza incluye en su numerador la suma de números cuadrados, que siempre serán positivos.

Si todas las observaciones en una distribución de datos aumentan o disminuyen en "x" valores, la media aumentará en los mismos valores, pero la varianza seguirá constante (**OPE**).

La **covarianza** es una medida de la dispersión conjunta que tienen dos variables (x, y). Si existe correlación positiva entre ambas variables (cuando x aumenta, y aumenta) la covarianza es positiva. Si existe correlación negativa (cuando x aumenta, y disminuye) la covarianza es negativa. Si no existe correlación, la covarianza es = 0.

- ✓ **Rango (recorrido):** es la resta entre el valor máximo que toma la variable en nuestra distribución menos su valor mínimo.
- ✓ **Rango e intervalo intercuartílico:** el rango es la diferencia entre el cuartil 3 (C3) y el cuartil 1 (C1), es decir, el rango existente entre los sujetos que ocupan el 50% central de la distribución. El intervalo intercuartílico es la medida más adecuada para las distribuciones asimétricas, ya que indica qué valores son los que ocupan el p25 y el p75 sin restarlos.

Medidas de posición (localización) (SRS 19, 46)

Se basan en la ordenación de las observaciones de menor a mayor, y la posterior división de la distribu-

ción obtenida en grupos que contienen el mismo número de observaciones. A cada grupo se le asigna un número que indica el número de grupos situados a su "izquierda", esto es, que tienen valores de la variable **menores o iguales** a él. En general a estos grupos se les denomina "**centiles**", pero en función del número de grupos que se utilicen existen distintos nombres:

- ✓ **Cuartiles (INGESA 21, 2):** se divide a la distribución en 4 partes iguales.
- ✓ **Deciles:** se divide a la distribución en 10 partes iguales.
- ✓ **Percentiles (OSASUN 22, 89):** se divide a la distribución en 100 partes iguales.

La mediana, al ocupar la posición central de la distribución, se considera también una medida de posición. Equivaldría al cuartil 2 (C2), decil 5 (D5) o percentil 50 (p50).

Ejemplo: el percentil 75 (p75) será el valor de la variable obtenido por aquél individuo tal que el 75% de las observaciones hayan sido menores o iguales a dicho valor, y el 25% de las observaciones hayan sido mayores a dicho valor. El p75 equivale al C3 y al D7,5.

El **diagrama de cajas (SERMAS 14, 11; SRS 21, 62)** o "**Box plot**" es una representación gráfica de la distribución de probabilidades en la que las líneas de las cajas representan los **cuartiles C1** (o Q1), **C2** (o Q2) y **C3** (o Q3), siendo Q2 (la mediana) la línea central.

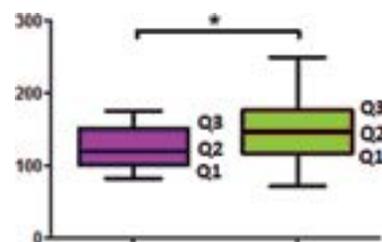


Figura 7. Diagrama de cajas o "Box plot" para representar distribución de probabilidades.

4. PRINCIPALES DISTRIBUCIONES DE PROBABILIDAD

Es una forma de denominar a la morfología de la representación gráfica de los resultados obtenidos. Siempre nos interesará que esta morfología sea similar a distribuciones conocidas que tienen propiedades matemáticas, para poder aplicarlas a nuestros resultados.

En el caso de **variables cuantitativas continuas**, nos va a interesar que sean similares a la **distribución normal (de Gauss)**.

Para las **variables cuantitativas discretas y cualitativas**, las más interesantes son la **distribución binomial** y la de **Poisson**.

Distribución normal (de Gauss)

La mayoría de las **variables biológicas** (presión arterial, temperatura, datos de laboratorio, peso, altura, etc.) se distribuyen con este patrón.

Se define por una función de probabilidad continua, cuyo **rango** va desde $-\infty$ hasta $+\infty$, en la cual los valores se agrupan en torno a un valor central con forma de campana.

- ✓ Es simétrica (**SERMAS 19, 95**).
- ✓ La media aritmética, mediana y moda coinciden.
- ✓ Es unimodal (tiene una única moda).
- ✓ El área bajo la curva de la distribución es igual a 1.

La distribución normal, aplicada a la estadística descriptiva, representa el porcentaje de observaciones que tiene cada valor posible, por lo que la suma de todos los porcentajes (área bajo la curva) será = 100% = 1.

El **grado de asimetría** de una distribución se denomina sesgo. Este sesgo será positivo si la asimetría hace que los valores se concentren a la izquierda de la media, habiendo por tanto una mayor dispersión a la derecha de la media. Por otro lado, si es negativo, los valores se concentran a la derecha de la media. Para las distribuciones totalmente simétricas, el sesgo es 0 (**Ver figura 8**).

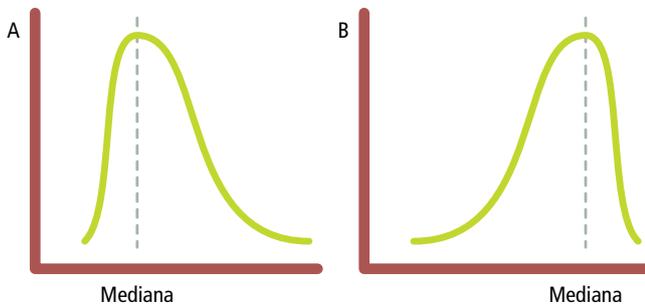


Figura 8. Distribuciones asimétrica positiva (A) y asimétrica negativa (B).

La utilidad matemática de la distribución normal radica en que permite definir **intervalos** que tienen un **área bajo la curva conocida**. Para la estadística

descriptiva implica el poder establecer intervalos que nos van a indicar **entre qué valores** se encuentra un **porcentaje de las observaciones conocido**:

- ✓ El intervalo $\bar{x} \pm \sigma$ comprende el 68% de los valores centrales u observaciones. Fuera de dicho intervalo queda el 32% de las observaciones (el 16% a cada lado).
- ✓ El intervalo $\bar{x} \pm 2\sigma$ comprende el 95% de los valores centrales u observaciones. Fuera de dicho intervalo queda el 5% de las observaciones (el 2,5% a cada lado).
- ✓ El intervalo $\bar{x} \pm 2,5\sigma$ comprende el 99% de los valores centrales u observaciones. Fuera de dicho intervalo queda el 1% de las observaciones (el 0,5% a cada lado).

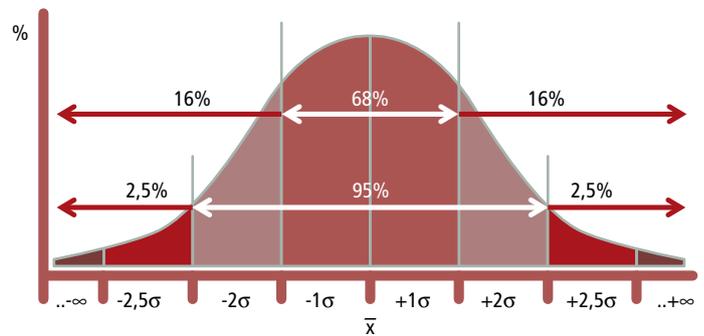


Figura 9. Distribución normal.

Distribución binomial

Se aplica a variables cuantitativas discretas o cualitativas, y consiste en convertir la variable en **dicotómica**, habiendo por tanto una probabilidad de “éxito” $p(A)$ y una probabilidad de fracaso: su probabilidad complementaria $p(1-A)$.

Distribución de Poisson

Es un caso particular de la distribución binomial que se utiliza para sucesos muy poco frecuentes: aquéllos en los que $p(A)$ ó $p(1-A) < 10\%$, y además hay < 5 individuos dentro de alguna categoría ($n \cdot p(A) < 5$ ó $n \cdot p(1-A) < 5$).

En la distribución de Poisson **la media coincide con la varianza**.

Estadística inferencial

Alba M^a Romero López, José Antonio Rueda Camino

1. Estadística inferencial para variables cuantitativas.....	17
2. Estadística inferencial para variables cualitativas	18
3. Cálculo del tamaño muestral para estudio de inferencias	18

Enfoque OPE

Tema poco importante, con escasas preguntas. Se ha preguntado por la definición de estadística inferencial y los estadísticos poblacionales.

Recuerda...

La Estadística inferencial estima cómo serían los resultados de la población objetivo si fuéramos capaces de estudiar a todos sus individuos (**SRS 19, 45**). Para ello, extrae conclusiones a partir de los resultados obtenidos en la muestra, por lo que existirá una probabilidad de error.



1. ESTADÍSTICA INFERENCIAL PARA VARIABLES CUANTITATIVAS

El objetivo va a ser estimar, con un determinado nivel de **confianza**, entre qué niveles se encontrará la **verdadera media poblacional** de la variable que hemos medido en nuestra muestra.

Para ello, de la distribución que obtenemos con los resultados de nuestra muestra, que refleja el porcentaje de observaciones de cada uno de los valores posibles, pasamos a una **distribución poblacional**, en la cual vamos a tener la **probabilidad** de que cada una de las posibles medias sea la **verdadera media poblacional**.

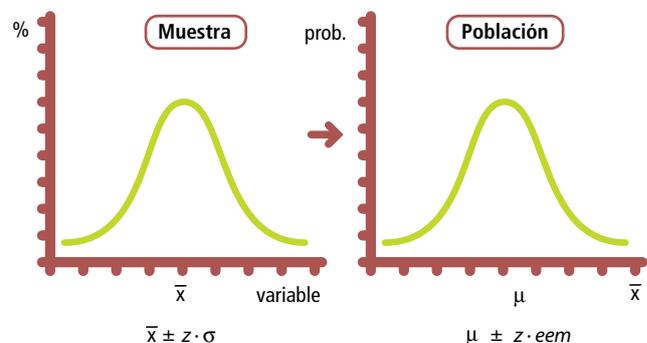


Figura 1. Estadística descriptiva (izquierda) y estadística inferencial (derecha).

Expresión de resultados de una inferencia de medias

Mientras que para los resultados de la muestra tenemos los intervalos que nos indican entre qué valores se encuentra un porcentaje determinado de observaciones, al **estimar** los resultados a la **población** utilizaremos los **intervalos de confianza (IC) (OPE)** que indican entre qué valores se encuentra con una elevada probabilidad, el **parámetro poblacional verdadero (SMS 18, 87; OSAKI 18, 87)**, como puede ser la media.

El intervalo de confianza puede tomar diferentes valores, aunque el más utilizado en ciencias de la salud es el intervalo de confianza al 95%, que se podría interpretar como que de cada 100 intervalos obtenidos en las muestras aleatorias de esa población, 95 contendrán el verdadero valor del parámetro poblacional.

- ✓ La medida de tendencia central (**media poblacional** = μ) se equipara a la media muestral ($\mu = \bar{x}$). *Si nuestra muestra es representativa de la población, la media muestral será el valor más probable que podrá tomar la media poblacional.*

- ✓ La medida de dispersión utilizada se denomina **error estándar de la media** (eem) o desviación típica poblacional, y se calcula a partir de la desviación típica muestral (**OPE**).

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

Para el cálculo de los **intervalos de confianza (IC)** se utilizan las propiedades matemáticas de la distribución normal:

- ✓ IC del 68% = $\mu \pm eem$
- ✓ IC del 95% = $\mu \pm 2 eem$
- ✓ IC del 99% = $\mu \pm 2,5 eem$



2. ESTADÍSTICA INFERENCIAL PARA VARIABLES CUALITATIVAS

Al igual que con las variables cuantitativas, el objetivo será **estimar** con un determinado nivel de confianza, los niveles entre los cuales se encuentra el **verdadero porcentaje poblacional** de cada una de las categorías de la variable medida.

Ejemplo: en nuestra muestra existe un 35% de individuos con el pelo moreno. Queremos estimar, con un 95% de confianza, qué porcentaje de individuos con el pelo moreno existirá en la población.

Para ello, empleamos intervalos de confianza análogos a los que se utilizan para variables cuantitativas, pero que utilizan el error estándar del porcentaje (**eep**) en lugar del error estándar de la media.

Como las variables cualitativas no tienen medidas de dispersión (no tienen desviación típica), el **error estándar del porcentaje (eep)** se calcula a partir del porcentaje muestral con la siguiente fórmula:

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

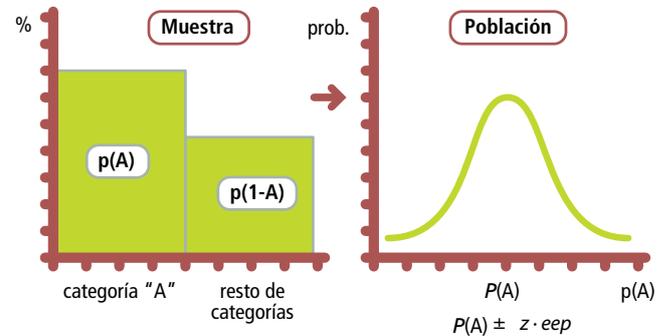


Figura 2. Estadística descriptiva (izquierda) y estadística inferencial (derecha).



3. CÁLCULO DEL TAMAÑO MUESTRAL PARA ESTUDIO DE INFERENCIAS

Para **calcular el tamaño muestral** mínimo necesario que nos permita obtener resultados precisos de medias o proporciones, tendremos que tener en cuenta tanto el **nivel de confianza** deseado, como la **variabilidad** del parámetro y la **precisión**.

Contraste de hipótesis

Alba M^a Romero López, José Antonio Rueda Camino

1. Errores en contraste de hipótesis.....	19
2. Test para contraste de hipótesis	20

Enfoque OPE

De este tema es importante conocer el concepto de error y sus tipos, así como la tabla con los distintos test de hipótesis y cuándo debe aplicarse cada uno.

Recuerda...

El **contraste de hipótesis** compara los resultados de varias poblaciones entre sí, para lo cual debe realizar inferencia poblacional a partir de muestras obtenidas de cada población. Por tanto, al igual que en Estadística inferencial, existe **probabilidad de error**.



1. ERRORES EN CONTRASTE DE HIPÓTESIS

Mediante el contraste de hipótesis, pretendemos evaluar si existen diferencias (comparación) o asociaciones (correlación) entre varias variables. El objetivo de este contraste es determinar si estas diferencias o asociaciones se deben al azar o, por el contrario, a un efecto real.

Para ello es necesario definir dos hipótesis junto con las probabilidades de que cada una de ellas se deba al azar (error alfa y error beta).

- ✓ **Hipótesis nula (H_0):** no existen diferencias o asociación entre las variables analizadas (**SARS 19, 84**).
- ✓ **Hipótesis alternativa (H_1):** existen diferencias o asociación entre las variables analizadas.

En la realidad sólo se podrá cumplir una de dichas hipótesis (o existen diferencias, o no existen), pero al realizar nuestro estudio podemos acertar o bien equivocarnos, viendo diferencias cuando no las hay (**error alfa**), o no viendo diferencias cuando las hay en la realidad (**error beta**). Así, existen cuatro posibilidades si contrastamos los resultados de la realidad con los obtenidos en nuestro estudio (**ver tabla 1**).

Tabla 1. Contraste de hipótesis.

		Realidad	
		Se cumple H_1 ($A \neq B$)	Se cumple H_0 ($A = B$)
Estudio test	Hallo diferencias (se acepta H_1)	Potencia $1 - \beta$	Error Tipo I α
	No hallo diferencias (no se rechaza H_0)	Error Tipo II β	Nivel de confianza $1 - \alpha$

Error tipo I (error alfa)

Error que cometemos cuando las diferencias observadas se deben al azar, siendo H_0 cierta, pero el investigador interpreta esta diferencias debidas a un efecto real (por lo que se acepta H_1 y se rechaza H_0). Por tanto, es la **probabilidad de rechazar la hipótesis nula siendo cierta (SMS 19, 120)**, estaríamos ante un "falso positivo".

La decisión acerca del **mantenimiento o rechazo de la hipótesis nula (H_0)** es el resultado de comparar el valor prefijado de α y el **nivel crítico o grado de significación (p) (SESPA 13, 47)**. Comúnmente se acepta un error alfa máximo de 0,05 (5%), que de forma muy

simplificada podríamos interpretar como que rechazaremos la hipótesis nula si la probabilidad de cometer el error alfa (es decir, que las diferencias se deban al azar y yo las esté dando por reales) es menor del 5%. Por tanto, se rechazará la hipótesis nula cuando el test de hipótesis dé una $p < 0,05$.

Error tipo II (error beta)

Es el error que se comete cuando las diferencias observadas son reales (en la realidad, H_1 es cierta), pero el investigador lo interpreta como debido al azar (en el estudio, no se acepta H_1 y no se rechaza H_0): es la **probabilidad de no rechazar la hipótesis nula siendo falsa**. Por lo tanto, es un resultado "falso negativo".

Cuando se realizan estudios epidemiológicos y se concluye que no existen diferencias, se suele requerir una probabilidad de haber cometido un error beta $< 0,20$ (menos del 20%).

Recuerda...

La hipótesis nula se rechaza o "no se rechaza".
La hipótesis Alternativa se Acepta o "no se acepta".

Potencia estadística (poder estadístico)

Es el mejor de los casos en un estudio: es la **probabilidad de detectar diferencias** (aceptamos H_1 y rechazamos H_0) **cuando en realidad existen** (en la realidad, H_1 es cierta). Por tanto, estamos hablando de la **probabilidad de rechazar la hipótesis nula**

siendo falsa. Estaríamos pues ante un "verdadero positivo".

La potencia estadística y el error beta son complementarios (potencia + $\beta = 1$). Por lo tanto:

$$\text{Potencia estadística} = 1 - \beta$$

$$\beta = 1 - \text{potencia estadística}$$

Así, de forma análoga al error beta, cuando se realizan estudios epidemiológicos y se concluye que no existen diferencias, se suele requerir que la potencia estadística fuera al menos de un 80%.

Recuerda...

Los errores alfa y beta son errores aleatorios, que ocurren debido al azar. La solución a estos errores es aumentar el tamaño muestral.

$\uparrow n \rightarrow \downarrow \alpha, \downarrow \beta, \uparrow$ potencia estadística



2. TEST PARA CONTRASTE DE HIPÓTESIS

(Ver figura 1)

Test para estudios de comparación de conjuntos de datos

Los principales test para comparación de conjuntos de datos se exponen en la **tabla 2**.

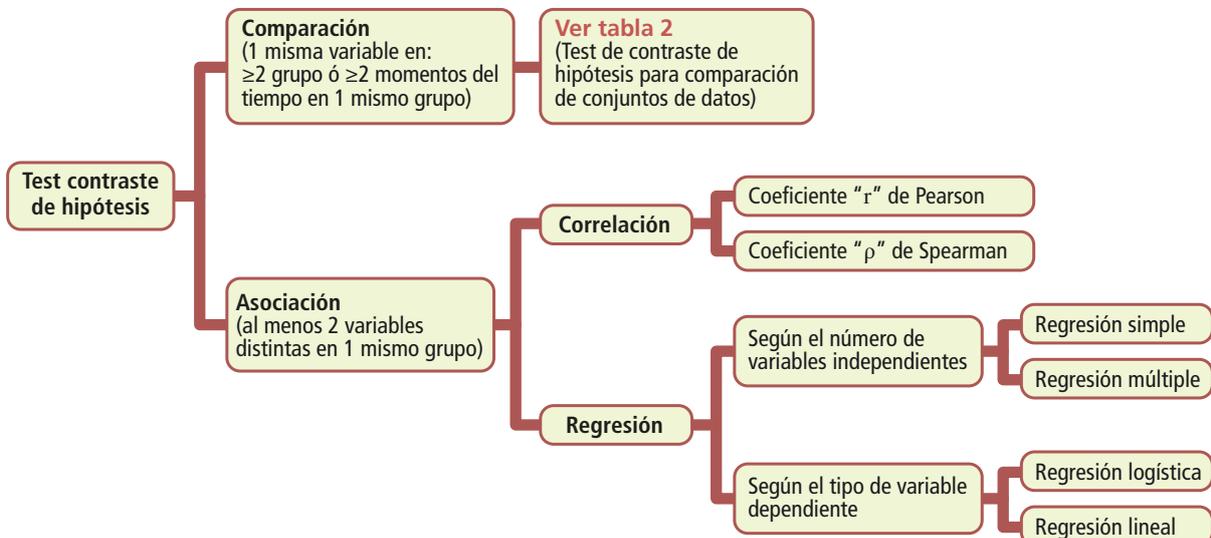


Figura 1. Test para contraste de hipótesis.

Tabla 2. Test de contraste de hipótesis para comparación de conjuntos de datos.

	Variable Cualitativa	Variable cuantitativa		
		2 GRUPOS 0 2 MOMENTOS DEL t	>2 GRUPOS 0 >2 MOMENTOS DEL t	
Datos independientes (VARIOS GRUPOS)	chi ² (χ ²) (SAS 21, 21; OSAKI 22, 81)*	Paramétrico	T Student (OSAKI 18, 86)**	ANOVA (F de Fisher-Snedecor) (SACYL 16, 13)
		No paramétrico	Mann-Whitney	Kruskal-Wallis
Datos Apareados (VARIOS MOMENTOS DEL t)	Mc Nemar	Paramétrico	T Student para datos apareados	ANOVA para datos apareados
		No paramétrico	Wilcoxon	Friedmann

(*) Cuando la frecuencia esperada es <5 en al menos un 25% de casillas de la tabla de contingencia, corresponde hacer el test exacto de Fisher. (**) Antes de aplicar este test hay que ver si las variancias son iguales (es decir, homocedásticas): para ello se aplica el test de Levene. Si éste arroja $p < 0.05$ las variancias son diferentes (heterocedásticas); si $p > 0.05$ se asume que son iguales (homocedásticas).

Para elegir el tipo de test que utilizar nos deberemos fijar en dos criterios:

- ✓ Si la **variable dependiente es cualitativa o cuantitativa**: cuando sea cuantitativa y tenga una distribución normal, utilizaremos los **test paramétricos (OPE)**. Para comprobar la bondad de ajuste a una distribución normal elegiremos el test de Shapiro-Wilk (también podremos utilizar el test de Kolmogorov-Smirnov, aunque siempre deberá ser la segunda opción). Estos test aportarán una mayor potencia estadística que los no paramétricos, permitiendo alcanzar en algunos casos la significación estadística. Si nuestra variable dependiente es cuantitativa, pero no sigue una distribución normal, o si tenemos una variable cualitativa ordinal, utilizaremos los **test no paramétricos (OPE)**.
- ✓ Si estamos comparando entre sí los resultados obtenidos en esa variable en **varios grupos de individuos (datos independientes)**, o bien en un único grupo de individuos pero en **varios momentos del tiempo (datos apareados)**.

Test para estudios de asociación entre variables

En este caso, lo que se pretende es demostrar si los cambios que se produzcan en una o varias variables

(variables independientes, x_i) van a influir sobre el valor que tome otra variable (variable dependiente, y); además, se pretende **cuantificar** dicha influencia. Todas las variables se recogen de una misma muestra.

Correlación

La **correlación** trata de expresar, mediante un **coeficiente de correlación**, lo “fuerte” que es el grado de asociación entre las variables evaluadas.

Los **test de correlación** más utilizados evalúan la correlación entre **dos variables cuantitativas**.

- ✓ **Coeficiente “r” de Pearson**: es un test paramétrico (se emplea cuando las dos variables siguen una distribución normal).
- ✓ **Coeficiente “ρ” de Spearman**: es un test no paramétrico (se emplea cuando alguna de las variables sigue una distribución no normal).

Los coeficientes de correlación pueden tomar **valores entre -1 y +1**. El **signo** del coeficiente de correlación indica si esta es positiva (cuando la variable x aumenta, también lo hace la variable y), o si es negativa (cuando la variable x aumenta, la variable y disminuye) (ver figura 2).

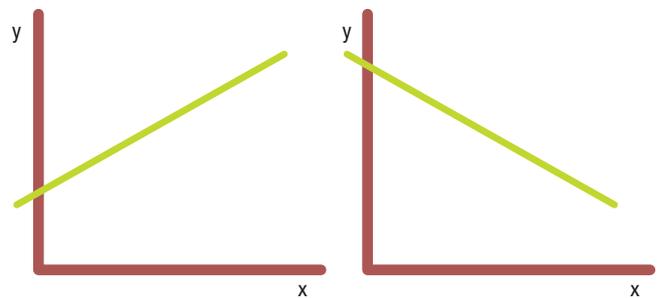


Figura 2. Correlación positiva (izquierda) y negativa (derecha).

El **valor absoluto** del coeficiente nos indica la magnitud de la asociación (es un índice de tamaño de efecto), de forma que tenemos:

- ✓ Valor absoluto **> 0.5**: Tamaño de efecto grande.
- ✓ Valor absoluto **entre 0.3 y 0.5**: Tamaño de efecto mediano.
- ✓ Valor absoluto **entre 0.1 y 0.3**: Tamaño de efecto pequeño.
- ✓ Valor absoluto **< 0.1**: Tamaño de efecto nulo.

Ausencia de correlación: si dos variables no están correlacionadas entre sí, los aumentos en “ x ” no producirán cambios de “ y ” (recta de regresión con pendiente = 0, esto es, horizontal), y los aumentos en “ y ” no producirán cambios en “ x ” (recta de regresión con pendiente = 0, esto es, vertical): las rectas de regre-

sión de "y" en función de "x" (y/x) y de "x" en función de "y" (x/y) se cortarán perpendicularmente.

La representación gráfica para ilustrar la relación entre dos variables cuantitativas es el gráfico de dispersión, tal y como se muestra en las **figura 3**.

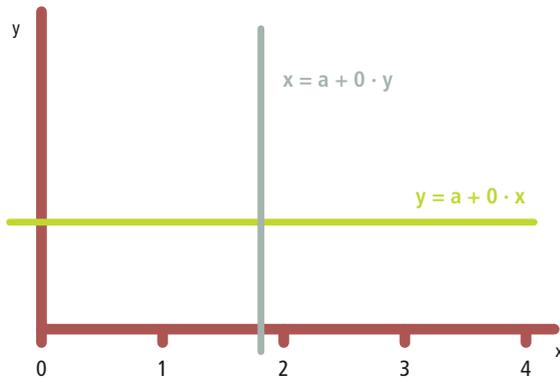


Figura 3. Ausencia de correlación lineal entre dos variables.

Regresión

La **regresión** trata de expresar mediante ecuaciones la asociación existente (mostrar mediante una fórmula matemática cómo varía la variable "y" con cada unidad de aumento de las variables "x_i").

Por ejemplo, en una muestra de individuos, analizar cuánto aumenta el colesterol (variable y) con cada kg que aumente el peso medio (variable x) en dicha muestra.

Si existe sólo una variable independiente (x_i) en la ecuación se habla de **regresión univariante o simple**. Si existen dos o más variables independientes (x_i) en la ecuación se habla de **regresión multivariante o múltiple**. La regresión multivariante sirve para evitar sesgos por factor de confusión, porque el efecto calculado para cada variable independiente es propio de sí misma, evitando la influencia de las demás.

El tipo de la variable dependiente (y) define el tipo de regresión:

- ✓ **Variable y es cualitativa:** utilizamos la **regresión logística**. Como medida de asociación utilizaremos la odds ratio, ya que esta medida nos habla de la probabilidad de enfermar o no (dicotómica) en función de la prevalencia de un factor de riesgo (o protector).
- ✓ **Variable y es cuantitativa:** utilizamos la **regresión lineal**, ya que es la fórmula matemática que define la ecuación de una recta.

$$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots + b_i \cdot x_i$$

El valor de la constante "a" indica el valor que toma la variable "y" (eje de ordenadas) cuando las variables independientes valen = 0. Se denomina ordenada en el origen.

*El valor de cada coeficiente "b" expresa cuantitativamente la asociación entre cada variable "xi" y la variable "y": indica cuánto aumenta la variable "y" con cada unidad de aumento de cada variable "xi". Se denomina **pendiente**.*

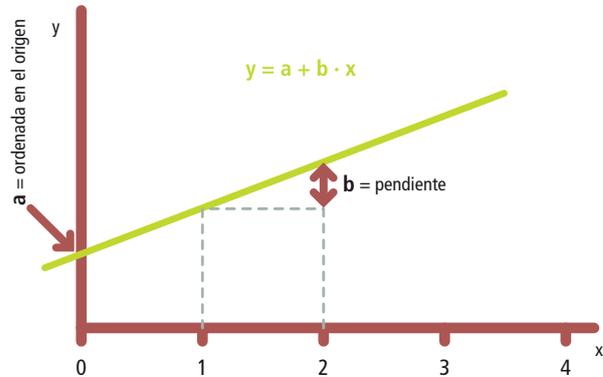


Figura 4. Regresión lineal simple.

Análisis de supervivencia

Es el método empleado cuando la variable de respuesta es el **tiempo** que transcurre entre un estado/evento inicial (fecha del diagnóstico de una enfermedad, fecha de inicio de un tratamiento) hasta que surge un **evento de interés** (tiempo transcurrido hasta la recaída de una enfermedad, tiempo durante el que se mantiene el efecto de un tratamiento, tiempo que pasa hasta la muerte de un paciente como resultado del evento inicial, tiempo transcurrido hasta el alta hospitalaria, etc.). Las variables presentan un componente cuantitativo (tiempo transcurrido) y un componente cualitativo (aparición o no de un evento).

En este análisis podemos emplear dos tipos de métodos para analizar los datos: los paramétricos (distribución exponencial, de Weibull y lognormal) y los no paramétricos (Kaplan-Meier, test de log-rank y regresión de Cox). Los más utilizados son los **no paramétricos**.

- ✓ **Kaplan-Meier:** es un estimador de la función de supervivencia, permitiendo el cálculo de las curvas de supervivencia (**ver tema 1, apartado "representación gráfica de las variables" y figura 6**).
- ✓ **Test de Log-Rank (logaritmo del rango):** es el test empleado para la comparación entre curvas de supervivencia, basa su cálculo en la chi cuadrado.
- ✓ **Modelo de regresión de Cox:** Es un modelo de regresión multivariante (**SERMAS 14, 13; SRS 19, 48**) que analiza la influencia de varios predictores sobre el tiempo que transcurre hasta el evento. *Por ejemplo, examina la influencia de la profesión, sexo y gramos de alcohol diario ingerido sobre el tiempo transcurrido hasta la aparición de hepatopatía.*

Estudios de validación de una prueba diagnóstica

Alba M^a Romero López, José Antonio Rueda Camino

1. Parámetros de validez de una prueba diagnóstica	23
2. Curvas ROC (de rendimiento diagnóstico)	24

Enfoque OPE

Es un tema básico de epidemiología en el que debes afianzar los conceptos de validez y fiabilidad, así como saber realizar problemas que utilicen parámetros tanto de validez interna como de validez externa.

Cuando se desea comercializar un nuevo test diagnóstico (p. ej., un nuevo modelo de esfingomanómetro), se debe llevar a cabo un **estudio de validación** mediante el cual se pretenden evaluar dos cualidades del test:

- ✓ **Validez (exactitud) (SARS 22, 86):** grado en el cual una medición representa el verdadero valor que se desea medir.
- ✓ **Reproducibilidad (fiabilidad, precisión):** capacidad del test de obtener el mismo resultado cuando la medición se repite bajo las mismas condiciones de medida.
- ✓ **Concordancia:** capacidad de obtener el mismo resultado cuando la medición es realizada en diferentes condiciones de medida.

El **índice kappa (κ)** se utiliza para evaluar la concordancia o reproducibilidad de los instrumentos de medida categóricos (2 o más categorías). Representa la proporción de acuerdos observados más allá del azar, respecto del máximo acuerdo posible.

- ✓ Resultado > 0.81: Grado de acuerdo casi perfecto

- ✓ Resultado entre 0.8 y 0.61: Grado de acuerdo sustancial
- ✓ Resultado entre 0.6 y 0.41: Grado de acuerdo moderado
- ✓ Resultado entre 0.4 y 0.21: Grado de acuerdo mediano
- ✓ Resultado entre 0 y 0.2: Grado de acuerdo insignificante.



1. PARÁMETROS DE VALIDEZ DE UNA PRUEBA DIAGNÓSTICA

Para evaluar la validez de una prueba diagnóstica, se realiza un estudio mediante el cual se comparan los resultados obtenidos por el test (que cataloga a los individuos en “+” o “-”) con los resultados obtenidos por el mejor método diagnóstico que esté disponible, llamado **gold standard** o **patrón oro** (que va a catalogar a los individuos del estudio en “enfermos” o “sanos”).



Tabla 1. Estudio de validación de una prueba diagnóstica.

	GOLD STANDARD		
	Enfermos	Sanos	
Test positivo	VP	FP	VP + FP total de positivos
Test negativo	FN	VN	FN+VN total de negativos
	VP+FN total de enfermos	FP+VN total de sanos	n

VP = verdaderos positivos; FP = falsos positivos; VN = verdaderos negativos; FN = falsos negativos.

Parámetros de validez interna

La **validez interna** es la capacidad que tiene el test para obtener resultados exactos en los sujetos de la **muestra** utilizada en el estudio. Por tanto, serán características intrínsecas del test, y **no se van a ver influenciadas por la prevalencia poblacional**.

- ✓ **Sensibilidad (S) (SES 19, 111; AVS 22, 64; SES 22, 115; OSASUN 22, 1):** capacidad del test de detectar a los sujetos enfermos. Es la probabilidad de que un sujeto enfermo (según el *gold standard*) saque "+" en el test. La probabilidad complementaria a la sensibilidad (esto es, la probabilidad de que un sujeto enfermo saque "-" en vez de "+" en el test) es la **tasa de falsos negativos (TFN)**.

$$S = VP / \text{total de enfermos}$$

$$TFN = FN / \text{total de enfermos}$$

$$S + TFN = 1 \rightarrow S = 1 - TFN; TFN = 1 - S$$

- ✓ **Especificidad (E) (SRS 19, 44; OSAKI 22, 21):** capacidad del test de detectar a los sujetos sanos. Es la probabilidad de que un sujeto sano (según el *gold standard*) saque "-" en el test. La probabilidad complementaria a la especificidad (esto es, la probabilidad de que un sujeto sano saque "+" en vez de "-" en el test) es la **tasa de falsos positivos (TFP)**.

$$E = VN / \text{total de sanos}$$

$$TFP = FP / \text{total de sanos}$$

$$E + TFP = 1 \rightarrow E = 1 - TFP; TFP = 1 - E$$

Parámetros de validez externa (SMS 18, 70; SRS 19, 42)

La **validez externa** es la capacidad que tiene el test para generalizar los resultados de la muestra a la población. Por tanto, la validez externa **sí se va a ver influida por la prevalencia poblacional**. La validez interna va a ser un requisito previo a la validez externa.

- ✓ **Valor predictivo positivo (VPP) (IBSALUT 09, 72; OSAKI 22, 18):** capacidad del test de predecir si un sujeto que ha sacado positivo en el test va a estar realmente enfermo. Es la probabilidad de que un sujeto "+" (según el test) sea enfermo según el *gold standard*.

$$VPP = VP / \text{total de positivos}$$

- ✓ **Valor predictivo negativo (VPN) (SESCAM 18, 22; SESPA 19, 37):** capacidad del test de predecir si un sujeto que ha sacado negativo en el

test va a estar realmente sano. Es la probabilidad de que un sujeto "-" (según el test) sea sano según el *gold standard*.

$$VPN = VN / \text{total de negativos}$$

Al ser los parámetros que miden la validez externa, los **valores predictivos dependen** de la **prevalencia** que haya en la población en la cual se apliquen (**SERMAS 14, 65**).

Así, si la prevalencia de una enfermedad es muy alta y un sujeto sale positivo en el test, será más probable que de verdad esté enfermo que si la prevalencia es muy baja (*simplemente por el hecho de que hay más individuos enfermos; el sujeto tiene más "papeletas" de ser enfermo*). Por el contrario, si la prevalencia de enfermedad es baja y un sujeto sale negativo en el test, será más probable que esté de verdad sano.

$$\uparrow \text{prevalencia} \rightarrow \uparrow \text{VPP}, \downarrow \text{VPN}$$

$$\downarrow \text{prevalencia} \rightarrow \downarrow \text{VPP}, \uparrow \text{VPN}$$

Los test que poseen alta sensibilidad catalogan correctamente a los enfermos, por lo que tienen muy pocos falsos negativos. Por lo tanto, serán útiles cuando sus resultados sean negativos, ya que en ese caso podremos descartar enfermedad (como hay muy pocos falsos negativos, casi todos los resultados negativos serán verdaderos: \uparrow VPN).

Sin embargo, los test que poseen alta especificidad catalogan correctamente a los sanos, por lo que tienen muy pocos falsos positivos. Por lo tanto, serán útiles cuando sus resultados sean positivos, ya que en ese caso podremos confirmar enfermedad (como hay muy pocos falsos positivos, casi todos los resultados positivos serán verdaderos: \uparrow VPP).

$$\uparrow S \rightarrow \downarrow TFN \rightarrow \uparrow VPN$$

$$\uparrow E \rightarrow \downarrow TFP \rightarrow \uparrow VPP$$

Por tanto, ante el diagnóstico de una enfermedad, primero emplearemos un test con alta sensibilidad (test de screening), para poder descartar a los sanos. Una vez los tengamos descartados, para confirmar la enfermedad utilizaremos un test con alta especificidad (diagnóstico de confirmación), así nos aseguraremos de que aquellos que habían dado positivo en screening, son verdaderos positivos.



2. CURVAS ROC (DE RENDIMIENTO DIAGNÓSTICO) (SERMAS 14, 38; SESSCAM 18, 103)

Las curvas ROC nos permiten obtener la mejor relación entre sensibilidad y especificidad de las

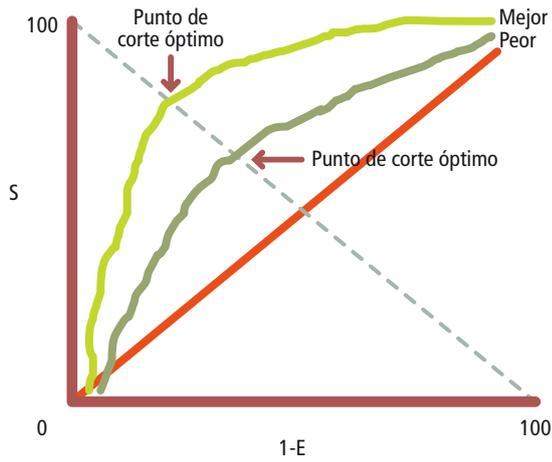


Figura 1. Comparación de la validez global de varios tests mediante sus curvas ROC.

pruebas diagnósticas. Gráficamente se representan poniendo en el eje de abscisas (X) la tasa de falsos positivos ($TFP=1-E$) y en el eje de ordenadas (Y) la sensibilidad, siendo el mejor punto de corte el que corta la bisectriz de la curva ROC.

El área bajo la curva representa el grado de validez global del test, es decir, la proporción de resultados válidos entre el conjunto de resultados ($VP+VN/n$).

La peor curva ROC posible es aquella que consista en la bisectriz del cuadrante formado por TFP y S ya que supondría una S y E máximas del 50%.

Medidas en epidemiología

Alba M^a Romero López, José Antonio Rueda Camino

1. Criterios de causalidad de Bradford Hill...	27
2. Medidas de frecuencia de una enfermedad.....	28
3. Medidas de fuerza de asociación (medidas de efecto).....	28
4. Medidas de impacto.....	29

Enfoque OPE

Tema bastante preguntado. Los conceptos más importantes son prevalencia e incidencia, tanto en definición como la diferencia entre ambos. Las medidas de impacto están cogiendo importancia, con lo cual debes conocer la diferencia entre cada una, además de su uso.



1. CRITERIOS DE CAUSALIDAD DE BRADFORD HILL (SMS 19, 84)

Para que se establezca una **relación de causalidad** se deben cumplir varios de los siguientes criterios (no hace falta que se cumplan todos):

Criterios de validez interna

- ✓ **Secuencia temporal:** la causa debe preceder al efecto. Es el único criterio de causalidad **imprescindible**.
- ✓ **Fuerza de asociación:** cuanto mayor sea la magnitud de la fuerza de asociación, mayor será la probabilidad de que exista relación causal.
- ✓ **Efecto dosis-respuesta (gradiente biológico):** a mayor dosis o tiempo de exposición al factor causal, mayor es el riesgo de enfermar.



Figura 1. Sir Austin Bradford Hill (1897-1991).

Criterios de coherencia científica

- ✓ **Consistencia:** los resultados deben ser reproducibles por otros investigadores y que estos obtengan resultados similares, es decir, que si el estudio se realiza por otro investigador tenga datos similares.
- ✓ **Coherencia:** los resultados de estudios que intenten establecer una relación causal entre factor-efecto deben ser similares entre sí, es decir, que diferentes estudios tengan conclusiones similares.
- ✓ **Plausibilidad biológica:** existencia de un mecanismo fisiopatológico conocido que explique la posible relación causal.

- ✓ **Especificidad de asociación:** si la posible causa conduce a un único efecto, y viceversa, el efecto sólo parece estar causado por un único factor; la verosimilitud de la relación causal aumenta. Este criterio también hace referencia a la **reversibilidad** de la asociación (que el cese de la causa conduzca al cese de la enfermedad).
- ✓ **Analogía:** factores causales similares al estudiado producen efectos similares.
- ✓ **Evidencia experimental:** es el criterio más potente, y nos habla de cuando ya se han hallado resultados sobre la asociación entre factor-efecto en estudios experimentales.



2. MEDIDAS DE FRECUENCIA DE UNA ENFERMEDAD (IBSALUT 19, 13)

Prevalencia

Es la proporción de individuos de una población que padecen una determinada enfermedad **en un momento dado** (SES 19, 6).

Es muy útil para valorar la extensión de **enfermedades crónicas** (CANAR 19, 34). Sin embargo, como sólo se evalúa un momento concreto y no un periodo de tiempo (OSAKI 18, 81), no es útil para el estudio de enfermedades agudas (*las enfermedades agudas aparecen y desaparecen, de modo que al estudiar un momento del tiempo concreto es probable no encontrar la enfermedad*).

Para estimarla, el diseño que resulta más eficiente es un **estudio transversal**, aunque con un estudio longitudinal también podríamos determinar la prevalencia de un momento dado.

La prevalencia de una enfermedad **aumenta en las siguientes circunstancias** (en las contrarias, por lo tanto, disminuirá):

- ✓ Aumento de la incidencia de la enfermedad (aumento de casos nuevos).
- ✓ Aumento de duración de la enfermedad (si disminuye su mortalidad).
- ✓ Descenso de la tasa de curación de la enfermedad.
- ✓ Mejora de los métodos diagnósticos de una enfermedad (se descubrirán más casos).
- ✓ Inmigración de casos enfermos o emigración de sujetos sanos.

Por tanto, todas aquellas medidas de **prevención primaria** van a disminuir la prevalencia al evitar la enfermedad, mientras que las medidas de **prevención secundaria y terciaria la aumentarán**, ya que se encaminan a la detección precoz de nuevos casos y la disminución de la mortalidad tras el diagnóstico.

Incidencia (incidencia acumulada) (IA)

(SAS 21, 20; SESCOAM 22, 28; SES 22, 27; OSAKI 22, 79)

Es la proporción de **casos nuevos** de una enfermedad que aparecen en una población en un determinado **periodo de tiempo**, con respecto al total de la población que es **susceptible** de enfermar. *Por ejemplo, si se desea medir la incidencia de una enfermedad crónica, habrá que calcular la proporción de casos nuevos respecto de los sujetos que no tengan dicha enfermedad (los que ya la tienen no son susceptibles de volver a enfermar).*

Indica la **probabilidad que tiene un sujeto sano de enfermar** a lo largo del periodo de tiempo que se haya tenido en cuenta para el cálculo de la incidencia (**riesgo individual de enfermar**).

Para hallarla, al verse modificada por el tiempo, necesitaremos realizar **estudios longitudinales prospectivos**.

Si queremos calcular la incidencia acumulada a partir del estudio de **varios periodos de tiempo** consecutivos, se dividirá el total de casos nuevos que hayan aparecido entre el **tamaño promedio** de la población susceptible que había en cada periodo.

Tanto la prevalencia como la incidencia son dos de los **indicadores** más importantes de la **morbilidad** (SERGAS 16, 44), ya que esta es un indicador epidemiológico que refleja el número de individuos enfermos en un momento dado. La **tasa de morbilidad** la hallamos al referir la morbilidad a un lugar y tiempo concretos, para poder determinar el impacto de una determinada enfermedad.

Densidad de incidencia (DI) (SERMAS 14, 31)

Es la **velocidad** con la que se propaga una enfermedad en una población, e indica el número de casos nuevos que aparecen por **unidad de tiempo**. El tiempo que se utiliza como unidad de medida es la suma del tiempo que ha estado expuesto a la enfermedad cada individuo hasta que la contrae: **suma de los tiempos de observación** (*en el momento que un individuo enferma finaliza su tiempo de observación, pero si un individuo no enferma a lo largo de todo el periodo de seguimiento, su tiempo de observación será lo que dure dicho periodo*).

Para calcularla también son necesarios **estudios longitudinales prospectivos** (ver tabla 1).



3. MEDIDAS DE FUERZA DE ASOCIACIÓN (MEDIDAS DE EFECTO) (AVS 18, 46)

Son razones que se calculan con un **cociente** entre el riesgo de sujetos expuestos a un factor y el riesgo de los sujetos no expuestos a ese factor.

Así, miden **cuántas veces** es más frecuente la enfermedad en el grupo expuesto respecto al no ex-

 **Tabla 1.** Prevalencia, incidencia y densidad de incidencia.

$\text{Prevalencia} = \frac{\text{n.º de casos en un momento puntual}}{\text{población}}$
$\text{IA} = \frac{\text{n.º de casos nuevos a lo largo de un periodo de tiempo}}{\text{población susceptible de enfermar al inicio del periodo}}$
$\text{DI} = \frac{\text{n.º de casos nuevos a lo largo de un periodo de tiempo}}{\sum t \text{ de observación de cada individuo susceptible de enfermar}}$

puesto, estimando pues la “**fuerza de asociación**” entre un factor causal y su efecto.

Su resultado oscila entre 0 e infinito (rango), y **no tienen unidades**:

- ✓ Si el resultado es <1: el factor estudiado es un **factor protector (SERGAS 16, 59)**.
- ✓ Si el resultado es >1: el factor estudiado es un **factor de riesgo**.
- ✓ Si el resultado es = 1: no existe relación causal entre el factor y la enfermedad (no es factor de riesgo ni de protección).

Al extrapolar estos resultados a la **población**, el intervalo de confianza (IC) que obtenemos informa sobre la **significación estadística** del mismo. Diferenciamos dos casos:

- ✓ El IC incluye el 1: no es estadísticamente significativo (SCS 19, 160)
- ✓ El IC no incluye el 1: es estadísticamente significativo.

En el caso de que el riesgo del grupo expuesto y no expuesto fuera el mismo, el cociente entre esos riesgos sería = 1, así que si el “1” está incluido en el intervalo de confianza, significará que existen probabilidades de que el factor estudiado sea tanto un factor protector (“parte” del intervalo de confianza que sea <1) como un factor de riesgo (“parte” del intervalo de confianza que sea >1).

Según qué medida de frecuencia de la enfermedad (del “riesgo” de enfermar) estemos utilizando, usaremos una u otra medida de asociación:

Riesgo relativo (RR) (AVS 22, 66)

Es la medida que se utiliza cuando disponemos de **incidencias acumuladas**. Como requiere del cálculo de incidencias, sólo se podrá calcular en estudios que partan de las exposiciones y vayan hacia las respuestas (como son los estudios de cohortes y los ensayos clínicos). Es la medida de efecto que **mejor estima** el riesgo real.

Odds ratio o razón de desventaja (OR)

Se utiliza en los estudios de **casos y controles**, que son aquellos en los que no calculamos incidencias, sino **odds** del factor de riesgo en el grupo enfermo y sano.

Es peor estimador del riesgo relativo, ya que tiende a **sobrestimar la fuerza de asociación**. Para intentar paliar esto, los casos y controles deben proceder de la misma población, y la incidencia de la enfermedad debe ser <10%, es decir, que en enfermedades poco frecuentes, se aproxima bastante al riesgo relativo.

(Ver tabla 2)

Hazard ratio

La hazard ratio es otro tipo de medida de fuerza de asociación, empleada en estudios de supervivencia, que informa acerca de la probabilidad de que ocurra un evento, p. ej., una enfermedad, en el subsiguiente intervalo temporal, al exponerse a un determinado factor o tratamiento.

El hecho de que exista una determinada fuerza de asociación entre un factor y una enfermedad NO implica necesariamente que dicho factor sea un factor causal de dicha enfermedad.



4. MEDIDAS DE IMPACTO

Quantifican cuál es el impacto de una medida preventiva al aplicarla en una población (al suprimir un determinado factor de riesgo, o al implementar un determinado factor protector). Utilizan **incidencias acumuladas**, por lo que se calculan **sólo en estudios con seguimiento prospectivo** (estudio de cohortes, ensayos clínicos, etc.).

Todas se calculan mediante una **resta** entre **incidencia en expuestos y no expuestos**, por lo que en las medidas de impacto, el **valor de no significación** (mismo riesgo en expuestos y no expuestos) será el 0. Al inferir los datos a la población, tendremos que:

- ✓ El IC incluye 0: no es estadísticamente significativo.
- ✓ El IC no incluye 0: estadísticamente significativo.

Medidas de impacto absolutas

Indican el beneficio absoluto (número de casos evitados por cada 100 personas en riesgo) que se obtiene al retirar un factor de riesgo o implementar un factor protector totalmente en una muestra o población en riesgo. Por lo tanto, son medidas útiles en **Salud Pú-**

 **Tabla 2.** Medidas de fuerza de asociación.

		enfermos	sanos
	Factor presente	a	b
	Factor Ausente	c	d

$$RR = \frac{IA \text{ expuestos (Ie) } a / (a + b)}{IA \text{ en no expuestos (Io) } c / (c + d)} = \frac{a / (a + b)}{c / (c + d)}$$

$$OR = \frac{\text{"Odds" del grupo enfermo}}{\text{"Odds" del grupo sano}} = \frac{\frac{\text{prevalencia del factor en enfermos}}{\text{prevalencia de no tener factor en enfermos}}}{\frac{\text{prevalencia del factor en sanos}}{\text{prevalencia de no tener factor en sanos}}} = \frac{\frac{a / (a + b)}{c / (a + c)}}{\frac{b / (b + d)}{d / (b + d)}} = \frac{a \cdot d}{b \cdot c}$$

blica (al permitirnos calcular, si conocemos el tamaño de una población, el número total de casos que evitaríamos en ella).

Riesgo atribuible, exceso de riesgo o diferencia de incidencias (RA, ER)

Medida de impacto absoluta utilizada para **factores de riesgo**. Indica el exceso de riesgo asociado a la exposición, y que podría evitarse si se eliminara ésta (número de casos evitados por cada 100 pacientes con el factor de riesgo a los que les quitas dicho factor).

$$RA = Ie - Io$$

Ejemplo: RA = 6% significa que por cada 100 expuestos hay 6 casos más de enfermedad que por cada 100 no expuestos. Si elimináramos el factor de riesgo en un grupo de expuestos, evitaríamos por cada 100 expuestos esos 6 casos de más.

El **riesgo atribuible poblacional (RAP)** nos indica el número de casos que se evitan por cada 100 personas de la población general si se lograra eliminar el factor de riesgo en esta población. Por tanto, dependerá de los valores de **prevalencia del factor de riesgo** en la población.

Es **más útil para Salud Pública** que el riesgo atribuible, ya que nos informa directamente del resultado de eliminar el FR deseado de la población diana.

$$RAP = I \text{ poblacional} - Io = RA \cdot \text{prevalencia del FR.}$$

Ejemplo: RAP = 6% significa que por cada 100 personas de la población existen seis casos más de enfermedad que por cada 100 no expuestos. Si elimináramos el factor de riesgo de la población, evitaríamos seis casos de enfermedad por cada 100 individuos.

El **número necesario de paciente a dañar (NNH)** es el número de pacientes que hay que "dañar" con un factor de riesgo para provocar un caso de enfermedad y se obtiene a partir de la inversa del RA:

$$NNH = 100 / RA \text{ (expresando el RA en \%)}$$

Reducción absoluta de riesgo (RAR)

Es la medida que se utiliza para **factores de protección**, y cuantifica la reducción de la incidencia de la enfermedad que podríamos conseguir al implementar ese factor protector en un grupo de riesgo:

$$RAR = Io - Ie$$

*Ejemplo: RAR = 5% significa que por cada 100 personas no protegidas hay cinco casos más de enfermedad que por cada 100 personas con el factor de protección. Si a las **personas no protegidas** les proporcionáramos el factor de protección, evitaríamos esos cinco casos que tienen de más.*

El **número necesario de pacientes a tratar (NNT)** es el número de pacientes que se debe tratar con un factor protector para prevenir un evento y se obtiene a partir del inverso de la RAR:

$$NNT = 100 / RAR \text{ (expresando la RAR en \%)}$$

Ejemplo: para una RAR = 5%, el NNT sería de 100 / 5 = 20, lo que significa que para evitar un caso de enfermedad, habría que proporcionar el factor protector a 20 personas.

Por tanto, mientras que la **RAR** representa los **casos evitados por cada 100** personas no protegidas si proporcionamos el factor protector, el **NNT** representará la cantidad de **pacientes** a los cuales tendre-

mos que **aplicar** el factor protector **para prevenir** un caso nuevo.

Medidas de impacto relativas

Indican el beneficio relativo (porcentaje de casos evitados del total de casos que padece una población en riesgo) que se obtiene al retirar un factor de riesgo o implementar un factor protector en una población en riesgo. Por lo tanto, son medidas útiles en **Epidemiología Clínica** (al sernos un indicador del porcentaje de riesgo de enfermar que evitamos en cada sujeto en riesgo).

Fracción atribuible o fracción etiológica de riesgo (FA, FER)

Medida de impacto relativa utilizada para **factores de riesgo (SRS 19, 50)**. Es la proporción de casos nuevos entre los expuestos que es atribuible a la exposición.

$$FA = (I_e - I_o) / I_e$$

Ejemplo: FA = 40% significa que de cada 100 casos de enfermedad que aparecen en un grupo de expuestos, 40 se deben a esa exposición (60 se deberán a otras causas). Así, si un individuo expuesto elimina su factor de riesgo (p. ej., deja de fumar) su riesgo de enfermar disminuirá un 40%.

La **fracción atribuible poblacional (FAP)** es el porcentaje de **casos eliminados** de la población si **eliminamos** de esta el **factor** de riesgo. Es una medida condicionada por la **prevalencia** de la exposición, y resulta más importante para Salud Pública que la fracción atribuible.

$$FAP = (I_{\text{poblacional}} - I_o) / I_{\text{poblacional}}$$

Ejemplo: FAP = 40% significa que de cada 100 casos de enfermedad que aparecen en la población, 40 se

deben a la exposición (60 se deberán a otras causas). Así, si lográramos eliminar el factor de riesgo en la población (p. ej., si las personas de la población que fuman dejan de fumar), el riesgo poblacional de enfermar disminuiría un 40%.

Reducción relativa de riesgo (RRR)

Medida de impacto relativa utilizada para **factores de protección**. Es la proporción de casos nuevos, entre los sujetos que no tienen el factor protector, que es atribuible a la ausencia de la protección que dicho factor confiere.

$$RRR = (I_o - I_e) / I_o = 1 - RR$$

Ejemplo: RRR = 35% significa que de cada 100 casos de enfermedad que aparecen en un grupo sin el factor protector (p. ej., no vacunados), 35 se deben a no tener el factor protector. Así, si un individuo adquiere el factor protector (p. ej., se vacuna) su riesgo de enfermar disminuirá un 35%.

 **Tabla 3.** Medidas de impacto.

	Factor de riesgo	Factor de protección
Medidas absolutas	Riesgo atribuible (RA) → Número necesario de pacientes a dañar (NNH)	Reducción absoluta de riesgo (RAR) → Número necesario de pacientes a tratar (NNT)
Medidas relativas	Fracción atribuible (FA)	Reducción relativa de riesgo (RRR)

Tipos de estudios epidemiológicos

Alba M^a Romero López, José Antonio Rueda Camino

1. Enfoques de investigación científica ...	33
2. Estudios cualitativos	34
3. Tipos de estudios epidemiológicos con enfoque cuantitativo en función de sus objetivos.....	36
4. Estudios observacionales.....	36
5. Estudios experimentales.....	38
6. Niveles de evidencia científica.....	39
7. Fases de desarrollo de un estudio cuantitativo.....	41
8. Fases de desarrollo de un tratamiento (fases del ensayo clínico)	43
9. Diseño paralelo y diseño cruzado en los estudios experimentales	44
10. Propuestas de investigación.....	45
Referencias bibliográficas	47

Enfoque OPE

Tema más importante de la asignatura con bastante diferencia. Preguntan sobre conceptos muy repetidos, por lo que resulta bastante rentable de estudiar. Presta especial atención a las diferencias entre investigación cualitativa y cuantitativa, así como a los objetivos de los diferentes tipos de estudios.



1. ENFOQUES DE INVESTIGACIÓN CIENTÍFICA

Existen dos grandes enfoques de investigación científica:

Enfoque cuantitativo (SRS 19, 41)

Características definitorias del enfoque cuantitativo:

- ✓ **Objetivo:** contraste de hipótesis mediante análisis estadístico, a partir de los datos recogidos sobre un objeto de estudio.
- ✓ **Método hipotético deductivo.**
- ✓ **Énfasis en la generalización:** buscamos reglas universales que expliquen el comportamiento humano.
- ✓ **Unicidad de la realidad,** por ello podemos observarla, estudiarla y medirla.
- ✓ **Imparcialidad y objetividad** del investigador.
- ✓ **Diseños** de investigación **muy estructurados y controlados (SMS 19, 47).**

Enfoque cualitativo (SMS 19, 11)

Características definitorias del enfoque cualitativo:

- ✓ **Objetivo:** comprender un objeto de estudio partiendo de datos recogidos, para construir una teoría acerca del mismo.
- ✓ **Método inductivo.**
- ✓ Tiene una **finalidad transformadora**, pues pretende producir cambios en el contexto estudiado.
- ✓ **Diferenciación clara entre realidad empírica** (objetiva; existe independientemente de tener conciencia de ella) y **epistémica** (realidad interpretada por el individuo; depende de quien la analiza u observa). **Énfasis en la realidad epistémica.**
- ✓ **Realidad múltiple que depende de variables** psicológicas, socioculturales e históricas que debemos contemplar (perspectiva holística), atendiendo a la diversidad y particularidad de cada una.
- ✓ **Subjetividad** del investigador: es utilizada como una **herramienta** que nos permite **generar conocimiento científico**. El investigador se integra en el contexto a estudiar (participativo), generando la mínima interferencia con el mismo.

- ✓ **Diseños de investigación flexibles (SAS 19, 91) y poco estructurados** que son **emergentes** (van modificándose a medida que avanza la investigación). Siempre debe ser **riguroso** y respetuoso con la **evidencia científica**, es decir, por ser flexible y abierto no es menos estricto.



2. ESTUDIOS CUALITATIVOS

Clasificación de estudios cualitativos (SAS 21, 24)

Algunos subtipos de investigación cualitativa son:

- ✓ **Fenomenología (SERMAS 14, 91; SACYL 16, 49; ICS 19, 41; OSASUN 18, 103; OSAKI 22, 80)**: es el estudio del significado subjetivo que las personas otorgan a situaciones/fenómenos sociales a través de su experiencia vivida.
- ✓ **Etnografía (CANAR 16, 65)**: pone el énfasis en las prácticas culturales de un determinado grupo social o una sociedad, haciendo especial hincapié en explicitar lo que subyace a estas prácticas sociales.
- ✓ **Etnometodología**: es el estudio de los procesos o métodos empleados por las personas para construir y dar sentido a sus prácticas sociales.
- ✓ **Investigación-acción**: es el estudio cualitativo que más énfasis pone en transformar la realidad objeto de estudio y no sólo en comprenderla. En este sentido, se trata de identificar un problema específico, requiriendo la participación de los afectados, los cuales serán protagonistas imprescindibles en el proceso de construcción del conocimiento que conduzca a mejorar la problemática identificada. Es un tipo de estudio que pretende hacer reflexionar y concienciar a los miembros de una determinada comunidad acerca de cuáles son los condicionantes de su realidad social, a fin de que tomen partido para poder mejorarla.
- ✓ **Estudio de casos**: estudio completo y pormenorizado de una cuestión y sus factores condicionantes tanto personales como contextuales. Pone especial atención en la singularidad del caso. res contextuales en los que está inmerso dicho individuo.

Técnicas de recogida de datos en investigación cualitativa

Las **técnicas de recogida de datos (CANAR 16, 82; SAS 19, 90; IBSALUT 19, 16)** utilizadas en investigación cualitativa son:

- ✓ **Grupo focal**: la importancia radica en el punto de vista particular de cada uno de los partici-

pantes que conforman el grupo. Existen modalidades que se centran más en estudiar la interacción entre participante y moderador, y otras centradas en la interacción entre participantes.

- ✓ **Grupo de discusión**: a diferencia del anterior, el grupo de discusión busca la producción de un discurso colectivo o compartido, grupal. A ello se llegará mediante la conversación, partiendo de las discrepancias hasta lograr el consenso.
- ✓ **Entrevistas en profundidad**: consisten en encuentros reiterados entre el investigador y el participante o un grupo de ellos. Estos encuentros están dirigidos hacia la comprensión de la perspectiva de los participantes sobre la realidad. Los encuentros son relativamente informales que evitan una secuencia estructurada de preguntas-respuestas, facilitando que el conocimiento se co-construya entre el investigador y el entrevistado.
- ✓ **Observación participante**: en esta técnica el investigador se integrará en el contexto social a observar, intentando no generar interferencias sobre la comunidad en cuestión. Previamente y en todo momento, el investigador debe ser consciente de que sus propias características van a influir en la forma que tiene de percibir e interpretar el objeto de estudio, por lo cual será relevante que adopte una actitud de aprendiz que le permita familiarizarse con la realidad. Es un proceso en el cual el investigador comienza realizando una mera observación descriptiva e indiscriminada, hasta que acaba siendo más selectiva sobre los aspectos de interés, a la par que va integrándose en la comunidad hasta ser un participante más, momento en el cual se retirará del campo, cesando la recogida de datos al perder la capacidad de discriminación de datos relevantes.
- ✓ **Diario de campo**: es un instrumento de recogida de datos especialmente utilizado en los estudios etnográficos, que consiste en un diario en el cual el investigador plasma sus percepciones sobre cómo se está implementando el estudio, limitaciones, oportunidades, interpretaciones teóricas, etc. Para ello, resulta imprescindible la rigurosidad de la información que concierne a la investigación, para que quede registrado en todo momento las decisiones que se van tomando.

Fases de la investigación cualitativa (SMS 19, 18)

1. Fase preparatoria: reflexión y planificación

Existe una primera fase en la que el investigador reflexiona y se plantea el problema a estudiar, para lo cual puede basarse en distintas fuentes de información: a partir de sus experiencias personales, de la

propia práctica profesional, así como de una revisión de la bibliografía; a fin de identificar un determinado tema de investigación, así como la finalidad del estudio. Esta pregunta de investigación y metas son las que guiarán las sucesivas fases del estudio. En esta fase inicial, dicha pregunta será general para ir posteriormente delimitándose y especificándose en mayor medida, a la par que se va aumentando la comprensión sobre el tema de interés.

Por otra parte, unido a esta pregunta de investigación, en esta fase se planifica el diseño del estudio que se va a realizar, el cual será también lo suficientemente flexible al inicio, para ir realizando luego las adaptaciones pertinentes a medida que se va adquiriendo una visión más profunda del objeto de estudio. Entre los aspectos que se han de planificar están: el **campo** o lugar donde se recogerán los datos; y el tipo de **técnicas de recogida de datos** que se han de emplear para responder al tema y objetivo de la investigación.

2. Trabajo de campo (CANAR 19, 32)

En esta fase, el investigador irá aproximándose al campo o comunidad a estudio, para recopiar toda la información necesaria, utilizando diferentes técnicas de recogida de datos (no siempre se empleará una única técnica). Durante esta fase, tanto el investigador como los participantes podrán ir adquiriendo roles más activos (o menos), tanto en la obtención de información como en la toma de decisiones.

Fases del trabajo de campo:

1. **Acceso al campo o fase de entrada:** el investigador entra en contacto con la realidad a estudiar. En esta fase, el investigador suele estar confuso al no estar familiarizado con el contexto objeto de estudio, de manera que debe aprender dónde y cómo encontrar la información relevante, así como ir acercándose a los participantes potenciales de la investigación. Esto puede hacerlo de una manera más informal ("vagabundeo") o más formal (ya sea mediante la confección de mapas espaciales sobre el lugar; a través de un registro de quién es quién en tal comunidad, etc.). Dentro de esta fase de trabajo de campo, cobra especial importancia la figura del **portero**. Este es el primer contacto del investigador que autoriza y facilita la implementación del estudio en una determinada comunidad, regulando así formal o informalmente el acceso al campo.
2. **Recogida productiva de datos:** esta fase requiere de gran sistematización y rigor por parte del investigador. Centraremos el problema y se recopilará la información que sea más importante, atendiendo a dos criterios fundamentales: adecuación y pertinencia. La **adecuación** de los datos decimos que se ha alcanzado cuando la cantidad de datos recogida

es suficiente para cumplir la **variación de datos** (los datos recogidos contemplan todas las variaciones y matices sobre el tema) y **saturación teórica** (por más información que recojamos no obtendremos nada nuevo, es decir, se solaparán los datos).

La **pertinencia** de los datos hace referencia a la **idoneidad** de la información recogida, e irá en función de las necesidades conceptuales de la propia investigación.

3. Análisis de los datos

En investigación cualitativa, la fase de análisis es **continua** y es una tarea que se inicia desde la misma fase de recogida de datos (y no exclusivamente después). Esta **conurrencia** entre **recogida y análisis** de datos es una **característica clave** en la metodología cualitativa.

Fases del análisis de datos:

1. **Reducción de datos:** es importante separar la información en unidades que sean útiles (especialmente cuando la información son textos o discusiones/entrevistas transcritas).
2. **Disposición y transformación de datos:** es el análisis propiamente dicho, en el que fundamentalmente se trata de clasificar y ordenar la gran cantidad de datos provenientes de diferentes fuentes de información: los derivados de técnicas observacionales, entrevistas, diarios de campo, discusiones grupales, otros documentos, etc. Para la ordenación de los datos se emplea la **codificación (SERMAS 14, 28; SMS 19, 110)** de la información, que conlleva crear etiquetas para determinadas frases, temas o conceptos que resuman el tema tratado. Esta codificación produce **categorías**, que es la forma de clasificar las unidades de información en función de las características compartidas, de forma que resulte más fácil ordenar los datos e identificar qué va encajando con las hipótesis y teorías que se van construyendo. Una vez codificada y categorizada la información, se empezarán a establecer relaciones entre tales categorías con la finalidad de ir tejiendo la teoría emergente de la investigación. Dado que este análisis de información se hace a medida que se van recogiendo los datos, esto permite tomar decisiones acerca de si continuar o no recabando mayor volumen de información (muestreo teórico).
3. **Obtención de resultados y conclusiones:** las conclusiones obtenidas deben verificarse, esto es, comprobar que son válidas. Llegada esta fase, el concepto de **triangulación** cobra un especial énfasis, el cual supone la combinación de datos provenientes de diferentes

técnicas de recogida de datos, observadores, experimentadores, etc.; con el fin de alcanzar una conclusión respecto del tema de investigación. La triangulación permite aportar evidencia de validez acerca de la información recogida, al ser contrastada mediante diferentes fuentes; a la par que nos permite construir una visión más completa y global del objeto de estudio. Hay cuatro tipos de triangulación (SARS 22, 23):

- **Triangulación metodológica:** al utilizar en la triangulación diferentes métodos se busca analizar el mismo fenómeno a través de diferentes acercamientos.
- **Triangulación de datos:** consiste en la verificación y comparación de la información obtenida en diferentes momentos mediante los distintos métodos.
- **Triangulación de investigadores:** el análisis de los datos es llevado a cabo por diferentes personas que, para dar una mayor fortaleza a los hallazgos, pueden ser de diferentes disciplinas.
- **Triangulación de teorías:** para esta triangulación partimos con diferentes teorías para observar un mismo fenómeno, con el fin de producir un entendimiento de cómo diferentes suposiciones y premisas van a afectar a los hallazgos e interpretaciones de los datos.

4. Fase informativa o de escritura del informe

Es la fase que termina con la redacción del artículo que recopila los hallazgos encontrados, y que presenta argumentos tanto a favor de los mismos como los que aportan interpretaciones alternativas.

Al igual que el diseño es más flexible y abierto que en investigación cuantitativa, la estructura del artículo también lo será. Además, en los resultados en lugar de incluir los datos hallados, se citan algunos comentarios de los participantes, siempre de forma anónima o cambiando el nombre.

Estándares de publicación:

Cuando se pretende enviar un artículo para su publicación en una revista científica, se han de cumplir una serie de criterios o estándares. La red EQUATOR (*Enhancing the QUALity and Transparency Of health Research*) propone diferentes guías de publicación. A continuación, se indican dos tipos de listas de comprobación que propone EQUATOR con los criterios que se han de verificar para los estudios cualitativos:

- ✓ **COREQ** (*Consolidated criteria for Reporting Qualitative Research*).
- ✓ **SRQR** (*Standards for Reporting Qualitative Research*).



3. TIPOS DE ESTUDIOS EPIDEMIOLÓGICOS CON ENFOQUE CUANTITATIVO EN FUNCIÓN DE SUS OBJETIVOS

En función de sus **objetivos** (SESCAM 22, 27), existen dos grandes grupos de estudios epidemiológicos:

- ✓ **Estudios descriptivos:** se centran en **describir** (OSASUN 18, 38) la naturaleza y magnitud del problema de salud. Pueden asemejarse a "tomar una foto" de la población a estudio. Se consideran estudios descriptivos:
 - Comunicación de un caso/serie de casos.
 - Estudios transversales y ecológicos (aunque pueden ser considerados analíticos, para ser descriptivos no pueden demostrar hipótesis).
- ✓ **Estudios analíticos** (AVS 18, 56): su objetivo es establecer la **relación** entre una determinada exposición y la aparición de un determinado problema de salud.

Se consideran analíticos los siguientes estudios epidemiológicos:

- Casos y controles.
- Cohortes.
- Estudios experimentales.
- Estudios cuasi-experimentales.



4. ESTUDIOS OBSERVACIONALES

En estos estudios **no hay intervención** (IBSALUT 09, 71) por parte del investigador sobre la población, solamente se limita a **observar** lo que ocurre habitualmente en la población. Pueden o no llevar **seguimiento**, y en función del mismo podemos diferenciar:

- ✓ **Sin seguimiento (estudios transversales):** estudio transversal y estudio ecológico.
- ✓ **Con seguimiento** (SESPA 13, 61; SES 19, 80):
 - Hablamos de seguimiento retrospectivo cuando el suceso de interés ha tenido lugar antes de diseñar el estudio. En cambio, si cuando nos sentamos a diseñar el estudio dicho suceso aún no ha ocurrido, hablamos de seguimiento prospectivo.

Comunicación de un caso / serie de casos (SACYL 19, 81)

Son estudios que describen las características o el manejo clínico realizado en un paciente o grupo de

pacientes con un diagnóstico similar. Por lo que no permiten el contraste de hipótesis al carecer de grupo control, siendo esta su principal limitación, pero sí que permiten **generar nuevas hipótesis** de trabajo.

Estudio transversal
(estudio de prevalencia (IBSALUT 09, 40; SAS 19, 28; ICS 19, 46; INGESA 19, 32), estudio de corte)

Es un estudio observacional de **base individual** (la unidad del estudio es el individuo) que **no presenta seguimiento** de los pacientes, esto es, que sólo estudia las características que tienen los pacientes en el **presente**: trata de describir o estudiar relaciones causales entre exposiciones y problemas de salud presentes **en un momento puntual**.

Ventajas

- ✓ **Rápido, barato y reproducible**, ya que carece de seguimiento.
- ✓ Es el diseño **más adecuado** para evaluar **validez de una prueba diagnóstica**, al igual que es el **más eficiente** para estimar la **prevalencia**.
- ✓ Resulta muy útil para el estudio de **enfermedades crónicas**, por lo que resulta útil para **planificación sanitaria y Salud Pública**.

Inconvenientes

- ✓ Muy sensible a los **sesgos**.
- ✓ No permite valorar la **secuencia temporal** (ya que estudia la presencia de la exposición y la enfermedad en el mismo momento). Por lo tanto, **no permite demostrar hipótesis etiológicas (causalidad)**, sino que sólo las **genera**.
- ✓ No es útil para el estudio de **enfermedades agudas** (ya que no permite medir incidencias por no tener seguimiento).

Estudio ecológico
(estudio de riesgo agregado)

Es un estudio transversal, pero que tiene una **base comunitaria**, lo que conlleva:

- ✓ Los datos recogidos pertenecen a **grupos de personas** formados por criterios geográficos.
- ✓ Los datos proceden de **registros**, por lo que son **datos indirectos**, al estar recogidos por otras personas. Por tanto, estos datos son de **peor calidad** y conllevan más riesgo de **sesgos**, al no poder controlar el método de recogida de datos.

Estudio de casos y controles (SERMAS 14, 15; SESPA 19, 28; SES 19, 41; OSAKI 22, 109)

Es un estudio observacional con una base individual, que tiene un **seguimiento retrospectivo**, ya que al iniciar el estudio el evento de interés ya ha ocurrido (partimos de la respuesta, enfermo o no enfermo, y vamos hacia las exposiciones). Por tanto, en este tipo de estudios tendremos siempre **dos grupos: grupo casos** (enfermos) y **grupo controles** (sanos). El **objetivo es evaluar la proporción de sujetos** de cada uno de los grupos que estaba **expuesto** al factor de riesgo o protector.

El seguimiento que se realiza en el estudio de casos y controles no es un seguimiento real que implique ir viendo a los pacientes varias veces en el tiempo, sino que el seguimiento es de forma virtual, es decir, se realiza utilizando la **memoria** del paciente o registros clínicos.

El seguimiento retrospectivo, por lo tanto, está sujeto a las limitaciones de la memoria humana y es peor (más sujeto a sesgos) que el seguimiento prospectivo.

- ✓ Utiliza como medida de fuerza de asociación la **odds ratio**, que **sobreestima** el riesgo con respecto al RR salvo en el estudio de enfermedades raras.

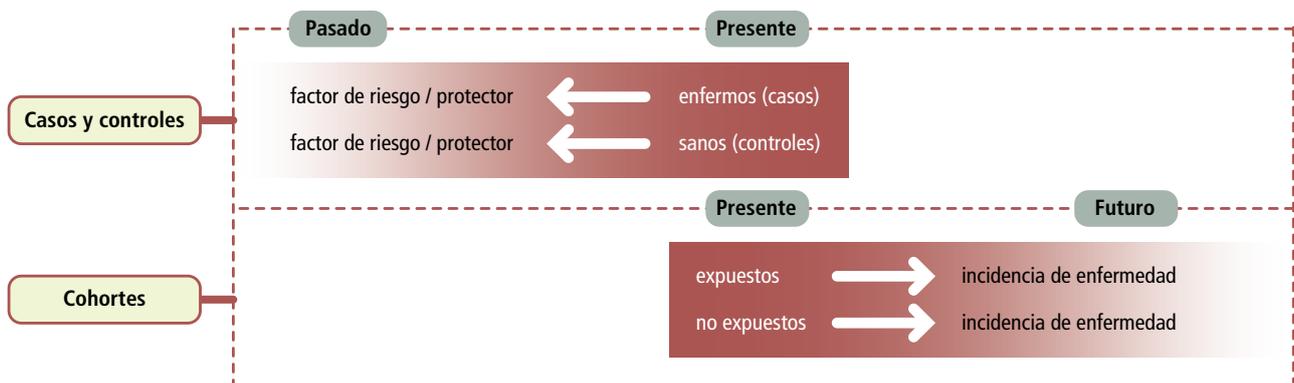


Figura 1. Estudio de casos y controles y estudio de cohortes.

- ✓ Sus ventajas e inconvenientes están recogidos en la **tabla 1**.

Estudio de cohortes (SMS 18, 6; SCS 19, 41; SRS 21, 54; SES 22, 25)

En un estudio de cohortes se sigue prospectivamente a **dos grupos de individuos sanos** con respecto al problema de salud que se quiere analizar: un grupo que está **expuesto** a un factor de riesgo o protector, y un grupo **no expuesto**. Se analiza la **incidencia** de enfermedad que aparece en cada uno de esos dos grupos a lo largo del periodo de seguimiento. Se trata, pues, de un estudio de **base individual**, con **seguimiento prospectivo** (si cogemos sujetos sanos hoy y los seguimos hacia el futuro a ver cuáles enferman según sus exposiciones) o **retrospectivo** (si cogemos sujetos sanos en el pasado, y “pasando las hojas de su historia clínica” llegamos hasta hoy viendo quién enferma según sus exposiciones). Los estudios de cohortes con seguimiento retrospectivo se llaman **estudios de cohortes históricas**.

Aunque el seguimiento pueda ser retrospectivo siempre partimos de las exposiciones y vamos hacia el resultado.

- ✓ Utiliza como medida de fuerza de asociación el **RR**, tanto si el seguimiento es prospectivo, como si es retrospectivo.
- ✓ Sus ventajas e inconvenientes están recogidos en la **tabla 1**.



5. ESTUDIOS EXPERIMENTALES

En estos estudios **sí que hay intervención** (SCS 15, 45; CANAR 19, 29) por parte del investigador, al introducir una nueva medida diagnóstica, terapéutica o preventiva en la muestra de sujetos. Por tanto, son los **mejores estudios** para **demostrar hipótesis**, ya que son **menos sensibles a sesgos**, al estar muy controlados.

En el otro lado tenemos que este tipo de estudios son **más caros** (SMS 18, 67), **lentos** y, por tanto, **poco reproducibles**, y están sujetos a **problemas éticos**.

Los tipos de estudios experimentales se diferencian según el **método de asignación de la intervención**:

- ✓ Estudios **experimentales**: la asignación de intervención es aleatoria
- ✓ Estudios **“cuasi-experimentales”**: la asignación de intervención es **no aleatoria**.

Estudios “experimentales” (IBSALUT 09, 20; IBSALUT 09, 32; IBSALUT 19, 92)

La asignación de intervención o no intervención a cada individuo se realiza de manera **aleatoria**, de



Tabla 1. Diferencias entre los estudios de casos y controles y los estudios de cohortes.

Casos y controles	Cohortes
Seguimiento retrospectivo	Seguimiento prospectivo
Calculan prevalencias de exposición	Calculan incidencias de enfermedad
OR : sobreestima la fuerza de asociación	RR : mejor estimador de la asociación
Baratos, rápidos, reproducibles	Caros, lentos, poco reproducibles
Peores para demostrar hipótesis	Mejores para demostrar hipótesis
Más sensible a sesgos	Menos sensible a sesgos
Útiles para estudiar enfermedades raras o con largo periodo de latencia	Útiles para estudiar exposiciones raras y para enfermedades agudas
Permiten estudiar multi-causalidad (varias causas de una enfermedad)	Permiten estudiar multi-efectividad (varios efectos de la misma exposición)
Dificultad : establecer el grupo control *	Dificultad : sensible a pérdidas **

* La principal dificultad de los estudios de casos y controles es seleccionar el grupo control, ya que óptimamente debe tener las mismas características que el grupo de casos para minimizar el riesgo de aparición de sesgos de selección o de sesgos por factor de confusión.

** Los estudios de cohortes (y cualquier estudio prospectivo) son sensibles a la aparición de pérdidas, ya que requieren ver a los pacientes en el futuro y eso puede no ser posible (fallecimientos, cambio de domicilio, incomparecencia del paciente...). En cambio, los estudios de casos y controles sólo requieren ver a cada paciente una vez, por lo que no pueden tener pérdidas.

modo que **es el azar el que forma los distintos grupos** (grupo intervención / grupo no intervención) que se van a comparar entre sí.

- ✓ **Ensayo de campo (ensayo preventivo)**: estudio experimental cuyo objetivo es la **prevención** de una enfermedad mediante la aplicación de una medida preventiva (p. ej., una vacuna) a un grupo de **sujetos sanos**, cuyos resultados se compararán con los de otro grupo de sujetos sanos a los que no se aplicó la medida preventiva.
- ✓ **Ensayo clínico** (SCS 15, 46; AVS 18, 60; OSAKI 18, 84; SARS 19, 106; SERGAS 19, 36; INGESA 19, 27): estudio experimental cuyo objetivo es el **tratamiento** de una enfermedad mediante la aplicación de una intervención (p. ej., un fármaco) a un grupo de **sujetos enfermos**, cuyos resultados se compararán con los de otro

grupo de enfermos a los que no se aplicó el tratamiento (**SESCAM 22, 21**).

Estudios “cuasi-experimentales”

En este tipo de estudios la asignación de la intervención se realiza siguiendo un **método no aleatorio**. En el caso de que solo haya **un grupo de pacientes** con intervención, **siempre** será **cuasi-experimental**, ya que la intervención se realizará sobre ese grupo, sin intervenir el azar.

*Por ejemplo, si los grupos que van a participar ya existen de manera natural (p. ej., dos localidades) y se elige cuál de los dos grupos recibirá la intervención mediante el azar (tirando una moneda al aire) no se está aleatorizando, ya que la aleatorización implica **formar los grupos** mediante el azar.*

- ✓ **Estudio de intervención comunitaria:** son estudios con **base comunitaria**, y en general para valorar medidas preventivas. Casi siempre son cuasi-experimentales (pero podrían ser aleatorizados) ya que, al utilizar una base comunitaria, los grupos de estudio suelen estar preformados y en esos casos es imposible formarlos mediante el azar.

*Los estudios de intervención comunitaria son habitualmente la alternativa a un **ensayo de campo** cuando es muy difícil utilizar una base individual. Por ejemplo, si se realiza un estudio para demostrar si la utilización de sal yodada previene problemas de hipotiroidismo en regiones endémicas, aleatorizar individuos y que unos tomen sal yodada y otros sal normal puede ser muy difícil (la sal se usa para cocinar y los sujetos de estudio tendrían que preparar su comida aparte de la del resto de miembros de su familia). En estos casos es más sencillo aplicar la intervención sobre comunidades enteras (p. ej., que a una determinada localidad se lleve sal yodada para la venta al público, y que a otra se lleve sal normal).*

- ✓ **Ensayo clínico no aleatorizado (estudio antes-después, pre-post):** es al igual que el ensayo clínico, con la diferencia de que la intervención no es aleatoria. La denominación de estudio pre-post (o antes-después), suele utilizarse cuando hay un único grupo de pacientes, en los cuales se compara la situación basal con la situación tras la administración del tratamiento.

El **nivel de evidencia científica** que genera un estudio, depende de la calidad del diseño del estudio. Hay numerosas escalas de cuantificación del nivel de evidencia, aunque la más utilizada es la que sigue el **esquema ABC**.

Niveles de evidencia científica ABC

- ✓ **Nivel de evidencia A:** son los estudios que aportan una **mayor evidencia**, y aquí se incluyen las recomendaciones basadas en varios estudios experimentales aleatorizados.
- ✓ **Nivel de evidencia B** (nivel de evidencia moderado):
 - Un único estudio experimental aleatorizado.
 - Estudios cuasi-experimentales.
 - Estudios de cohortes grandes.
- ✓ **Nivel de evidencia C** (menor nivel de evidencia):
 - Estudios observacionales (salvo cohortes).
 - Consenso de expertos.

Evaluación de la evidencia científica GRADE (SERMAS 19, 45; SES 22, 126)

El **grupo GRADE** (Grading of Recommendations, Assessment, Development and Evaluation) propone una clasificación del **grado de confianza** otorgado a un estimador de efecto para una recomendación, y resulta muy útil para la **elaboración de guías de práctica clínica**. Se divide en cuatro categorías:

1. **Alta:** existe confianza alta en que el estimador del efecto se encuentra muy cerca del efecto real.
2. **Moderada:** el estimador del efecto probablemente se encuentre cercano al efecto real, aunque las diferencias podrían ser sustanciales.
3. **Baja:** el estimador del efecto puede ser sustancialmente diferente al efecto real.
4. **Muy baja:** es muy probable que el estimador del efecto sea sustancialmente diferente al efecto real.

El grado de confianza de las recomendaciones no establece el reconocimiento de las decisiones y preferencias en la práctica clínica. Corresponde al personal clínico, basándose en evidencia y grado de recomendación, el seleccionar una u otra medida.



6. NIVELES DE EVIDENCIA CIENTÍFICA

Actualmente, las sociedades científicas médicas basan sus recomendaciones en los resultados de los estudios epidemiológicos disponibles sobre cada materia (**Enfermería Basada en la Evidencia**).

Índices de impacto (AVS 18, 59)

Son un instrumento para comparar y evaluar la importancia relativa de una revista determinada dentro

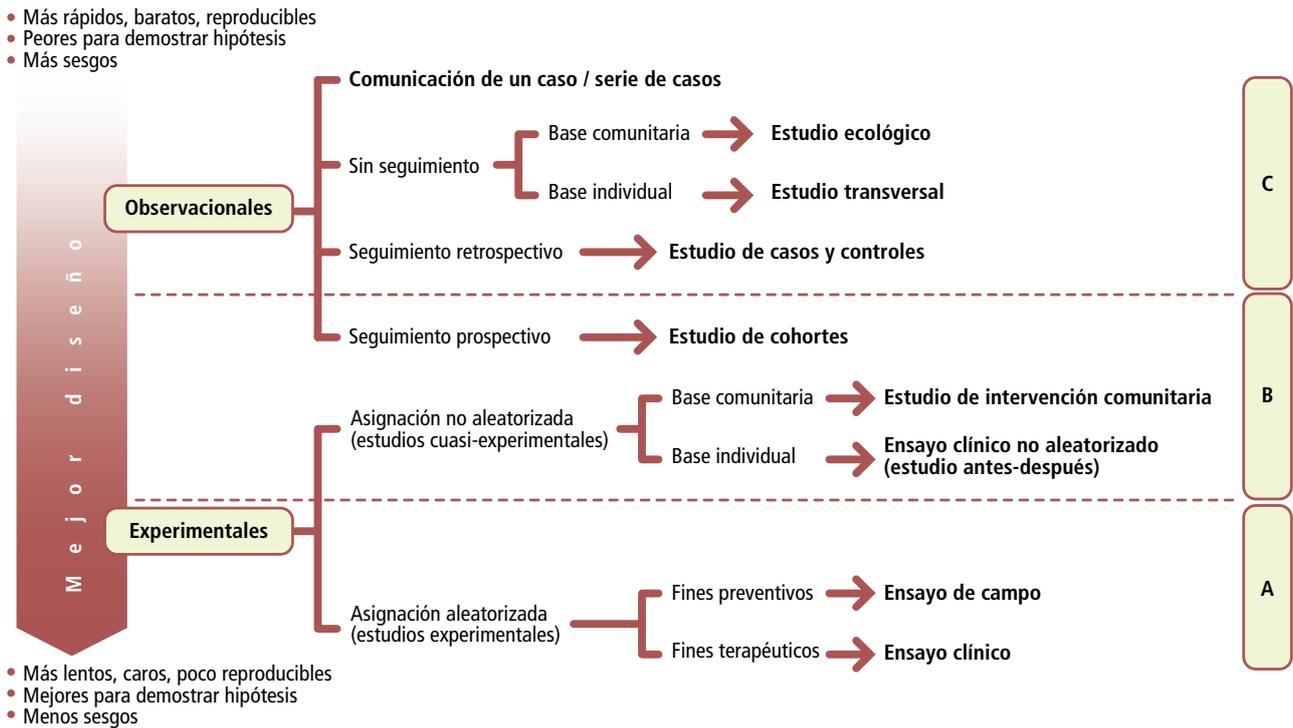


Figura 2. Tipos de estudios epidemiológicos y nivel de evidencia científica que generan.

de un mismo campo científico, en función del promedio de citas que reciben los artículos que en ella se publican durante un periodo de tiempo determinado. El más conocido es el **factor de impacto**, que puede ser consultado en el Journal Citation Report (JCR) o Scimago Journal Report (SJR).

Metaanálisis (IBSALUT 09, 53; INGESA 21, 94; SES 22, 121; OSASUN 22, 88)

El metaanálisis es una **revisión sistemática** de la literatura en la que **se combinan estadísticamente** los resultados de todos los estudios incluidos:

- ✓ **Revisión sistemática:** consiste en la realización de una búsqueda bibliográfica siguiendo unos **criterios de selección** concretos y definidos previamente, de forma que todo estudio que cumpla con esos criterios se incluirá en el meta-análisis. *Las revisiones narrativas por su parte, incluyen los artículos que elija el investigador libremente, sin atender a criterios preestablecidos de selección.*
- ✓ **Combinación estadística de resultados:** los resultados de los pacientes que participaron en cada estudio individual se tratan como si todos los pacientes hubieran participado en un único estudio. Así, se obtiene un **resultado combinado** que procede de un tamaño muestral inmenso (la suma de tamaños muestrales

de cada estudio individual), con lo que se consigue:

- **Mayor precisión:** intervalos de confianza más pequeños.
- **Mayor potencia estadística (SERMAS 14, 102):** mayor probabilidad de demostrar la existencia de diferencias de modo estadísticamente significativo, si realmente existen.

Los metaanálisis se realizan en situaciones en las que la **evidencia científica** disponible sobre un tema es **inconcluyente**, bien porque los estudios realizados tienen un pequeño tamaño muestral y no han demostrado diferencias significativas (quizá por falta potencia estadística), o bien porque los resultados de los distintos estudios arrojen conclusiones discordantes.

Clasificación por grado de procesamiento de la información

La **pirámide de Haynes** es un método para ordenar los recursos existentes según el grado de procesamiento de la información. Este método es una pirámide que consta de **6 escalones**. En la base de la misma encontraremos la información más en bruto (por ejemplo, metabuscadores como Pubmed), mientras que en la cúspide nos encontramos los sistemas de ayuda para la toma de decisiones clínicas (por ejemplo, UptoDate). (ver figura 3).

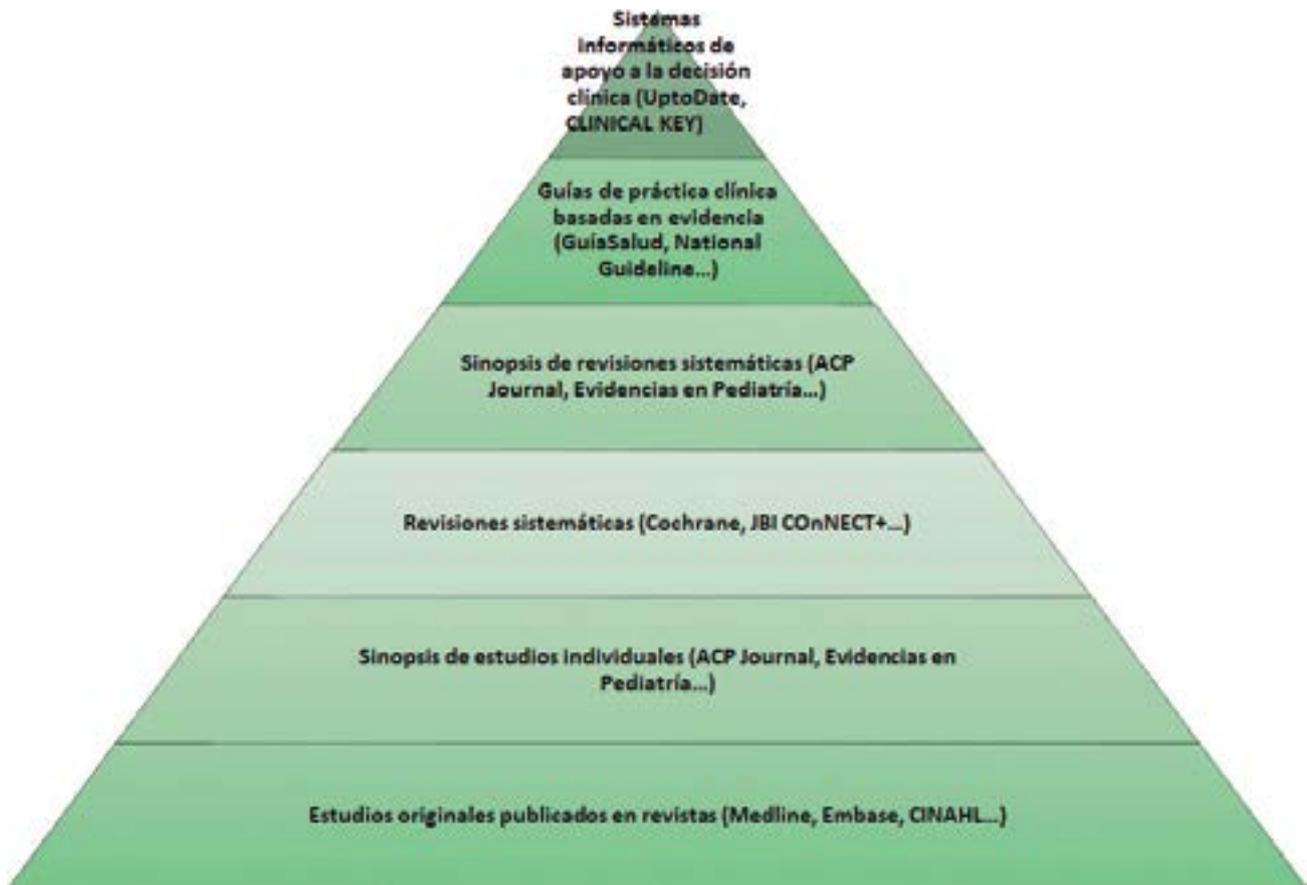


Figura 3. Escalones de pirámide de Haynes.



7. FASES DE DESARROLLO DE UN ESTUDIO CUANTITATIVO

1. Diseño

Se deben especificar *a priori* (antes de empezar el estudio) todas las características del estudio (aspectos metodológicos), con un cuidado especial en explicar todas las mediciones que se van a realizar y el método estadístico con el que se van a analizar los resultados.

Si se extraen conclusiones de un estudio epidemiológico *a posteriori* (estudios post hoc), dichas conclusiones no servirán para confirmar hipótesis, sino que sólo sirven para generarlas y se deberán confirmar con nuevos estudios específicamente diseñados para ello. Por ejemplo, realizar un análisis de subgrupos que no estaba planificado inicialmente y observar que un determinado subgrupo se beneficia del fármaco.

Por tanto, la fase de diseño se caracteriza por:

- ✓ **Identificación del problema.**
- ✓ Elaboración una **hipótesis** de trabajo.
- ✓ **Definición** de las **características metodológicas** del estudio.

2. Reclutamiento (inclusión de sujetos participantes)

Una vez definidas las características del estudio, se realiza la inclusión de los sujetos participantes. Para incluir a un paciente en un estudio epidemiológico, deben seguirse por orden los siguientes pasos:

2.1. Criterios de selección

Una vez decididos tales criterios de selección de los participantes, se verificará que cada potencial participante cumpla todos los **criterios de inclusión** y no reúna ningún **criterio de exclusión**.

2.2. Consentimiento informado

El participante debe expresar libremente y por escrito su consentimiento para participar en el estudio.

2.3. En estudios experimentales: asignación de intervención

Tras comprobar que los sujetos cumplen los criterios de selección y firman el consentimiento informado, se le asignará un grupo de estudio. En los estudios estrictamente experimentales en esta fase se realiza la **aleatorización**.

3. Monitorización

Fase de seguimiento de los participantes, en la que se realizan las mediciones previstas en el diseño y se obtienen los resultados.

4. Análisis de resultados y obtención de conclusiones (OSASUN 18, 37)

4.1. Verificar que el diseño sea correcto y no haya sesgos (validez interna)

Es el paso más importante del análisis de resultados, ya que la ausencia de validez interna no permitirá realizar inferencias causales. La **validez interna** se refiere a la calidad metodológica del diseño, de manera que los cambios sobre la variable dependiente se puedan atribuir exclusivamente a los cambios en la variable independiente. Para ello, se han de descartar todo tipo de sesgos (p. ej., sesgos de selección, de clasificación incorrecta, factores de confusión, atrición, etc.).

Recuerda...

La **lectura crítica (IBSALUT 19, 18)** de un artículo científico es el proceso sistemático por el que se evalúa el diseño y la calidad de un estudio en relación con su validez interna. Analiza fundamentalmente aspectos metodológicos del estudio (los encontramos en *Material y métodos*).

La **red CASPe (Critical Appraisal Skills Programme)** es una iniciativa que promueve la formación en lectura crítica para pacientes, investigadores y gestores.

4.2. Confirmar la existencia de significación estadística

4.3. Valorar la magnitud de las diferencias existentes y su relevancia clínica

La existencia de diferencias estadísticamente significativas entre dos intervenciones no implica que una de ellas sea más recomendable que otra. Sólo si la magnitud de dicha diferencia es suficiente (medida a través de índices de **tamaño de efecto**), y si implicará beneficios importantes desde un punto de vista clínico, podremos decir que una intervención es más relevante que otra y, por tanto, establecer una recomendación al respecto.

4.4. Determinar la validez externa

En esta fase se analizará si los resultados de la investigación se podrían generalizar a otros subgrupos poblacionales que no han formado parte de la muestra de nuestro estudio, pero que cumplen los criterios de selección delimitados en nuestra investigación.

Una de las estrategias metodológicas para aumentar la validez externa de cualquier diseño de investigación es la selección aleatoria de la muestra de sujetos.

5. Difusión de los resultados

Esta es una de las fases más importantes de cualquier investigación, pues si no se difunden los resultados es como si no se hubiera realizado. Tenemos diferentes formas de realizar esta difusión, bien sea mediante artículos científicos en revistas científicas, como capítulos de libros, o bien sea en congresos (formato de comunicaciones orales o póster).

5.1. Estándares de publicación

Para la publicación de un estudio o documento científico determinado se ha de seguir una normativa que variará en función del diseño de la investigación que ha sido desarrollada. A continuación, se enumeran los diferentes tipos de listas de comprobación que propone EQUATOR con los criterios que se han de verificar para los distintos tipos de estudios cuantitativos y otros documentos científicos:

- ✓ **CONSORT (CONsolidated Standards Of Reporting Trials)**: para publicar ensayos clínicos (aleatorizados y controlados) (**SERMAS 14, 100**).
- ✓ **STROBE (STrengthening the Reporting of OBservational studies in Epidemiology)**: para publicar estudios observacionales.
- ✓ **STARD (STAndards for Reporting Diagnostic accuracy studies)**: pruebas diagnósticas.
- ✓ **PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses)**: meta-análisis y revisiones sistemáticas.
- ✓ **QUOROM (QUality Of Reporting Of Meta-analyses)**: meta-análisis para ensayos clínicos aleatorizados.
- ✓ **TREND (Transparent Reporting of Evaluations with Nonrandomized Designs)**: para estudios cuasi-experimentales.
- ✓ **CARE (clinical CAse REporting guideline)**: casos clínicos.
- ✓ Otras listas de comprobación:
 - **AGREE (Appraisal of Guidelines, REsearch and Evaluation)**: se emplean para publicar guías de práctica clínica.

5.2. Normativas de publicación

Las citas bibliográficas que se incluyen en el texto deberán enumerarse en el apartado de referencias bibliográficas al final del texto. Para ello, existen numerosas normativas que establecen cómo debe ser el formato de estas referencias, qué debe incluir y de qué forma. Las más utilizadas en ciencias de la salud son:

✓ **Normativa APA:**

- **Artículos en revistas científicas.** Tanto el nombre como el número de la revista han de aparecer en cursiva. Por ejemplo:

- Aiken, L. H., Clarke, S. P., Sloane, D. M., Sochalski, J., & Silber, J. H. (2002). Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*, 288, 1987-1993.

- **Libro.** El nombre del título del libro ha de aparecer en cursiva. Por ejemplo:

- Alligood, M. R., & Tomey, A. M. (2018). *Modelos y teorías en enfermería*. Barcelona: Elsevier Health Sciences.

✓ **Normativa Vancouver:**

- **Artículos en revistas científicas.** Ningún elemento aparece en cursiva y las referencias bibliográficas aparecen secuenciadas por orden de aparición y no por orden alfabético. De ahí que aparezcan enumeradas en el listado de referencias bibliográficas. Por ejemplo:

- 1. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*. 2002;288(16):1987-93.

- **Libro.** Ningún elemento aparece en cursiva y el año de publicación aparece al final. Por ejemplo:

- 1. Alligood MR, Tomey AM. *Modelos y teorías en enfermería*. Barcelona: Elsevier Health Sciences; 2018.

✓ **Normativa UNE-ISO 690:**

- **Artículos en revistas científicas.** El apellido de los autores ha de aparecer completamente en mayúsculas y su nombre completo en minúsculas (salvo la letra inicial). El nombre de la revista ha de aparecer en cursiva y luego el año de publicación sin cursiva. Por ejemplo:

- AIKEN, Linda H., et al. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*, 2002, vol. 288, no 16, p. 1987-1993.

- **Libro.** El apellido de los autores ha de aparecer completamente en mayúsculas y su nombre completo en minúsculas (salvo la letra inicial). El nombre del título del libro ha de aparecer en cursiva. Por ejemplo:

- ALLIGOOD, Martha Raile; TOMEY, Ann Marriner. *Modelos y teorías en enferme-*

ría. Barcelona: Elsevier Health Sciences, 2018.



8. FASES DE DESARROLLO DE UN TRATAMIENTO (FASES DEL ENSAYO CLÍNICO)

En el desarrollo de un tratamiento tenemos dos fases claramente diferenciadas:

- ✓ **Fase preclínica:** engloba a la síntesis química del tratamiento, los estudios con material biológico in vitro, así como el primer testeo de eficacia y seguridad en animales.
- ✓ **Fase clínica:** esta fase comienza tras la anterior, y consiste en ensayos clínicos del tratamiento en humanos. A su vez, se divide en varias fases (ver tabla 2):



Tabla 2. Fases del ensayo clínico.

Fase	Población	Características	Objetivos
I	Voluntarios (sanos)	Diseño cuasi-experimental	<ul style="list-style-type: none"> • Características farmacocinéticas • Toxicidad preliminar
II	Enfermos	<ul style="list-style-type: none"> • Criterios de selección estrictos • ↓ n • Variables blandas 	<ul style="list-style-type: none"> • Ila: eficacia y seguridad • Iib: titular dosis
III	Enfermos	<ul style="list-style-type: none"> • Criterios de selección laxos • ↑ n • Variables resultado duros 	Eficacia y seguridad
IV	Práctica clínica habitual		<ul style="list-style-type: none"> • Efectividad • Nuevas indicaciones • Farmacovigilancia

Fase I

Es la primera fase dentro de la fase clínica, por lo que será la primera vez que utilizaremos el tratamiento en humanos.

En esta fase tendremos un único grupo de sujetos (**estudio cuasi-experimental**) que serán voluntarios, para poder estudiar las **propiedades farmacocinéticas** del nuevo tratamiento.

Como objetivo secundario, se estudia de forma preliminar la **toxicidad** del fármaco (si su administración produjo efectos adversos a los voluntarios).

Habitualmente se realiza sobre **voluntarios sanos** (que suelen recibir una remuneración económica por participar en el estudio), pero si el tratamiento se prevé que tenga muchos efectos adversos se realizará con voluntarios enfermos (*para que, al menos, el tratamiento tenga algún efecto beneficioso; p. ej., quimioterápicos*).

Fase II y Fase III

En estas fases sí que nos encontramos un **diseño de ensayo clínico**, ya que su principal objetivo es evaluar la **eficacia y seguridad** del tratamiento cuando se aplica a **sujetos enfermos**.

- ✓ **Fase II:** empleo del tratamiento en un **grupo reducido** de enfermos (en general <100) muy seleccionados, esto es, con **criterios de selección estrictos (estudios explicativos)**.

Además, suelen utilizar **variables resultado "blandas"** (determinaciones de laboratorio, pruebas de imagen...), que al ser en general cuantitativas proporcionan **mayor potencia** estadística pero **menor relevancia clínica** de los resultados.

Hay estudios que distinguen una **Fase IIb**, en la cual se prueban varias dosis diferentes para establecer la **relación dosis-respuesta** (titulación de dosis), para poder elegir la dosis más adecuada de cara a la Fase III.

- ✓ **Fase III:** empleo del tratamiento en un **grupo amplio** de enfermos (en general >100) con **criterios de selección laxos (estudios pragmáticos)**, lo que va a permitir que la muestra sea similar a los pacientes que se van a encontrar en la práctica clínica habitual.

Además, suelen utilizar **variables resultado "duras"** (variables clínicas: muerte, infarto, ictus...), que al ser en general cualitativas (*p. ej., muerte: sí/no*) proporcionan menor potencia estadística pero **mayor relevancia clínica** de los resultados.

De esta forma, la Fase III nos arroja evidencia de la utilidad del nuevo tratamiento para los pacientes, pero desde el punto de vista clínico. Es por ello que son los que permite que se **comercialice** el tratamiento.

Fase IV (fase postcomercialización)

Incluye aquellos estudios realizados con un fármaco tras su comercialización, que como hemos visto ocurre tras la publicación de los estudios en fase III.

Tiene fundamentalmente dos **objetivos**:

- ✓ **Búsqueda de nuevas indicaciones:** para ello, se deberán volver a realizar estudios con diseño análogo a la fase II y fase III.
- ✓ **Farmacovigilancia:** intenta detectar **reacciones adversas poco frecuentes** que pasaron inadvertidas en los ensayos realizados por su tamaño muestral, y que solamente pueden detectarse una vez el fármaco está comercializado y se administre a miles de personas. Para ello, hay un sistema de notificación espontánea de posibles reacciones adversas asociadas a un tratamiento, para el cual se debe rellenar la llamada **tarjeta amarilla**.

Si cualquier personal sanitario detecta una posible reacción adversa a un tratamiento no identificada previamente, debe notificarlo mediante la cumplimentación de una "tarjeta amarilla" que se envía a la Agencia Española del Medicamento.



9. DISEÑO PARALELO Y DISEÑO CRUZADO EN LOS ESTUDIOS EXPERIMENTALES

El **diseño paralelo** consiste en que un grupo de sujetos recibe un único tratamiento o intervención, y otro grupo recibe otro tratamiento distinto. Con la finalidad de que ambos grupos sean equivalentes y las diferencias individuales derivadas de los participantes sea una variable extraña (o factor de confusión) controlada en el diseño; los grupos de pacientes se forman al azar, a cuya técnica se le denomina **aleatorización**.

A diferencia del diseño paralelo, en el **diseño cruzado** todos los participantes recibirán todos los tratamientos o intervenciones sucesivamente. Al ser los mismos participantes, la **variable extraña** que tendremos que controlar es el **efecto de orden** de los tratamientos, ya que el efecto del último tratamiento podría verse magnificado o minimizado en función de cómo lleguen los participantes al mismo. Para ello, se podrían formar dos subgrupos de pacientes, y se les asignaría la secuencia de tratamiento de forma aleatoria. A esta técnica se le denomina **contrabalanceo** o **equiponderación**.

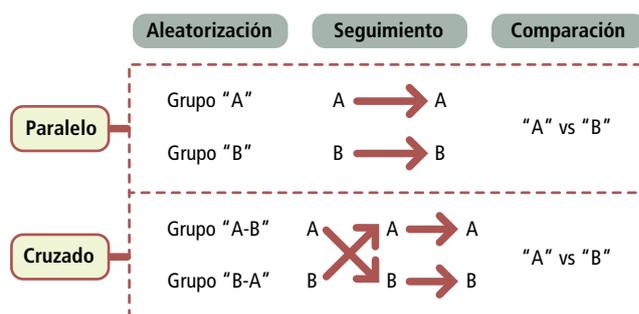


Figura 4. Diseño paralelo y diseño cruzado.

Ventajas del diseño cruzado respecto al paralelo

- ✓ **Requiere la mitad de tamaño muestral:** en un diseño cruzado, cada individuo recibe los dos tratamientos, por lo que sirve como su propio control y ahorra la inclusión de otro individuo para servir de control. *Si necesitamos 100 pacientes que tomen la medicación "A" y 100 pacientes que tomen la medicación "B", en un diseño paralelo tendremos que incluir 200 pacientes (100 por ramo de tratamiento), mientras que en un diseño cruzado sólo necesitaremos 100 pacientes (50 en el grupo A-B y 50 en el grupo B-A).*
- ✓ **Muestra más homogénea:** el grupo de individuos que toma el fármaco "A" es igual que el que toma el fármaco "B" (son las mismas personas); esta circunstancia hace surgir el concepto de "**variabilidad intraindividual o intragrupal**", que hace referencia al hecho de que en los diseños cruzados cada individuo es su propio control y por tanto la variabilidad que existe entre cada individuo y su control (o entre cada grupo y su control) es nula (son las mismas personas). Sin embargo, la "variabilidad interindividual o intergrupala" (entre dos sujetos o grupos diferentes del estudio) no varía respecto al estudio paralelo.

Inconvenientes del diseño cruzado respecto al paralelo

- ✓ **Mayor duración:** dura el doble de tiempo porque cada sujeto recibe los dos tratamientos.
- ✓ **Muy sensible a pérdidas:** al perder a un paciente, perderemos uno por cada tratamiento, con lo cual sería como perder dos pacientes.
- ✓ **Efecto periodo:** las características clínicas de la enfermedad pueden cambiar entre el primer tratamiento y el segundo. Por ello, la enfermedad debe ser crónica, no es útil para enfermedades agudas o con brotes (a menos que sean predecibles).
- ✓ **Efecto residual:** efecto que deja el primer tratamiento en el organismo. Para paliarlo, debemos dejar un tiempo entre el primer tratamiento y el segundo (**periodo de lavado**), para que se elimine el fármaco y su efecto residual. Este tiempo debe ser similar para ambos tratamientos. Por otro lado, el tratamiento **no puede ser curativo ni dejar efectos irreversibles**.

yecto que se quiere llevar a cabo y justifica éticamente su necesidad científica. Cuando se quiere realizar una investigación, se debe presentar dicha propuesta al comité ético de la institución donde se vaya a llevar a cabo (por ejemplo, en un hospital) para su aprobación. Si el proyecto no implica a seres humanos sino a animales, el proyecto deberá ser aprobado por el comité ético de protección animal de la institución.

Partes de una propuesta de investigación (SCS 19, 40; SESCOAM 22, 23)

Estos apartados son variables en función del organismo que solicita la memoria del proyecto de investigación, pero suelen incluir una serie de puntos comunes.

0. Resumen

Inicialmente se redacta un resumen, en el cual se ha de describir todo el proyecto brevemente, comenzando por una concisa justificación del mismo; continuar con la enumeración de sus objetivos; describir sus participantes destinatarios; así como indicar las características de la evaluación e intervención con los participantes/pacientes.

1. Problema de investigación (OSAKI 18, 82; OSAKI 22, 78)

Formulación del problema de investigación que debe ser relevante, innovadora, de interés profesional, clara y precisa (SRS 21, 53). Para ello se realiza una búsqueda bibliográfica del tema y se exponen los objetivos de la investigación. Se suele añadir una justificación ética de la necesidad de realizar la investigación en humanos o animales.

- ✓ **Fuentes de información científica:** las fuentes pueden ser **primarias** (contienen información original y se utilizan para la búsqueda directa de información) o **secundarias** (no tienen como objetivo ofrecer información, sino que facilitan la localización e identificación de documentos) (SESPA 19, 25). Para realizar la búsqueda bibliográfica recurriremos a las fuentes secundarias: metabuscadores y bases de datos:

- **Metabuscadores:** un metabuscador es una herramienta que permite hacer una búsqueda a partir de varias bases de datos simultáneamente, ya que no es un buscador que disponga de una base de datos propia. A continuación, enumeraremos los metabuscadores más empleados en Ciencias de la Salud:

- **PubMed Health (U.S. National Library of Medicine).** Es un amplio metabuscador que está centrado en investigación



10. PROPUESTAS DE INVESTIGACIÓN

Una propuesta de investigación es un documento redactado a modo de memoria que describe el pro-

sobre efectividad clínica, y ofrece recursos científicos de diferente naturaleza (revisiones sistemáticas, libros especializados, glosario técnico, etc.).

- **NHS Evidence (*National Institute for Health and Care Excellence, UK*)**. Ofrece acceso a evidencia científica en Ciencias de la Salud, Servicios Sociales y Salud Pública. Incluye información actualizada y gratuita sobre guías clínicas, revisiones sistemáticas, información a pacientes, resúmenes de hallazgos, etc.
- **TRIP database (*Turning Research Into Practice –TRIP–*)**. Permite realizar una búsqueda simple, avanzada y también siguiendo el **acrónimo PICO**. El acrónimo PICO (**SACYL 16, 1**) está formado a partir de las siguientes palabras en inglés: *population* (población diana, tipos de pacientes objeto de interés); *intervention*, (tratamiento o intervención diana); *comparison* (comparación de la intervención diana con otras intervenciones); y *outcome* (variable dependiente objeto de interés) (**SAS 21, 23**).
- **Epistemonikos**. Es la mayor fuente de revisiones sistemáticas de ciencias de la salud, así como una fuente de otro tipo de evidencia. Es un metabuscador colaborativo multilingüe.
- **SUMSearch 2**. Es un metabuscador médico gratuito, que busca simultáneamente artículos originales, revisiones sistemáticas y guías prácticas a partir de PubMed.
- **Bases de datos:**
 - **Cuiden**. Es una base de datos bibliográfica que contiene artículos de revistas científicas, libros, monografías, incluso materiales no publicados, especializados en el ámbito de la Enfermería y Cuidados de Salud.
 - **CINAHL**. Es una base de datos para documentos científicos (artículos de revistas, entre otros) del ámbito de la Enfermería y otras Ciencias de la Salud.
 - **Medline**. Es la base de datos que está asociada a PubMed, por tanto, se centra en Biomedicina y Ciencias de la Salud. Fue creada por la Biblioteca Nacional de EEUU.
 - **Cochrane**. Es una base de datos que contiene específicamente revisiones sistemáticas (**SERMAS 14, 18; IBSALUT 19, 17**).
 - **Scopus**. Es una base de datos para documentos, revisados por pares, de disciplinas pertenecientes al ámbito de la Medicina; Ciencias Sociales; Ciencia y Tecnología; así como Artes y Humanidades.
- **Otras fuentes de información científica:**
 - **SciELO**. Es una biblioteca electrónica que posibilita el acceso a ediciones completas de revistas científicas, la cual fue especialmente creada para potenciar la diseminación de resultados científicos en los países en desarrollo, y particularmente, de América Latina y el Caribe.
 - **Literatura gris** (literatura no convencional, literatura semipublicada o literatura invisible): cualquier tipo de documento que no se difunde por los canales de publicación comercial (**SARS 19, 76**), y que plantea problemas de acceso. Por ejemplo, aquí se incluyen tesis de pre y posgrado, actas de congresos, cuadernos de trabajo, etc.
 - **Tesauros (**SACYL 19, 26**)**: es una fuente de información científica especial que nos permite buscar la definición de conceptos y la palabra técnica concreta (descriptores, palabras clave) que ya han sido consensuadas para nombrar un concepto. Destacan dos:
 - El **DeCS** (Descriptores en Ciencias de la Salud): tesauro trilingüe que nos permite consultar en español, inglés y portugués. Fue elaborado a partir del MeSH.
 - El **MeSH** (Medical Subject Headings): tesauro exclusivamente en inglés que fue creado por la Biblioteca Nacional de Medicina de EEUU (está contenido dentro de PubMed)
 - ✓ **Estrategias de búsqueda**: en estos metabuscadores o bases de datos se pueden emplear diferentes estrategias de búsqueda, así como utilizar filtros que limiten dicha búsqueda bibliográfica, tal y como pasamos a exponer a continuación.
 - **Operadores booleanos (AND, OR, NOT) (**SERMAS 14, 50; OSAKI 18, 85; SESPA 19, 19; CANAR 19, 123; SRS 19, 43; OSAKI 22, 107**)**: son estrategias de búsqueda que diseña y produce el propio investigador, las cuales permiten encontrar referencias bibliográficas que combinen todas las ideas, términos o palabras clave especificadas en la búsqueda (usando AND entre todas las palabras clave); o permite encontrar referencias bibliográficas que presenten al menos uno de los términos especificados en la búsqueda (usando OR entre las palabras clave); así como permiten buscar referencias bibliográficas que contengan únicamente la primera palabra clave y no los siguientes tér-

minos especificados en la búsqueda (usando NOT entre el primer y los siguientes términos delimitados).

- **Truncamiento:** es una herramienta que nos facilita la búsqueda de referencias sobre un determinado concepto. Para ello, se emplea la raíz semántica de la palabra y se pone un asterisco detrás. Por ejemplo, si queremos hacer una búsqueda sobre espirometrías, podremos acortar la palabra para que quede: espiometr*, y que nos incluya referencias bibliográficas que contengan "espirometría", "espirómetro", "espirométrico". **Filtros:** los filtros son limitadores en la estrategia de búsqueda bibliográfica que ya aparecen por defecto en las distintas fuentes de información. Esto es, no son estrategias de búsqueda que ha de generar el investigador. A continuación, destacamos los dos siguientes filtros: el idioma y el periodo de tiempo.
 - **Idioma:** lengua en la que interesa que esté publicada la referencia bibliográfica, de manera que al seleccionar un determinado idioma, se limitará la búsqueda de referencias bibliográficas a aquellas que estén publicadas exclusivamente en dicho idioma.
 - **Periodo de tiempo:** intervalo de tiempo en el que interesa que hayan sido publicadas las referencias bibliográficas que queremos encontrar. Así pues, por ejemplo, podemos limitar la búsqueda a aquellas referencias publicadas en los 5 últimos años, a las últimas 4 semanas, así como personalizar un rango de años determinado.
- **Campos específicos:** son campos utilizados para diseñar una búsqueda más avanzada (Advanced search). Con esta herramienta podremos buscar a partir del apellido de un autor (campo Author), a partir del nombre de la revista (campo Journal), por ciertas palabras del título (campo Title) o en título y abstract de forma simultánea (campo Title/Abstract), entre otros.
- ✓ **Gestores de referencias bibliográficas:** un gestor de referencias bibliográficas es una herramienta que permite crear, mantener, organizar, compartir y dar forma a las referencias bibliográficas de artículos de revista, libros, capítulos de libros u otro tipo de documentos a partir de distintas fuentes de información (bases de datos, metabuscadores, revistas,

páginas web, etc.). Al realizarse de forma automática y utilizar formatos normalizados, simplifican el proceso de creación del listado de referencias bibliográficas, así como la citación de dichos documentos a lo largo del texto.

Entre los gestores de referencias bibliográficas más empleados se encuentran los siguientes: EndNote, Mendeley, RefWorks, Zotero, BibTeX, RefMan, etc.

- ✓ **Identificación de documentos científicos:** principalmente hay tres formas de identificar documentos: DOI, ISBN e ISBN-A.
 - **DOI** (Digital Object Identifier): es una forma de identificar digitalmente un objeto que suele ser de naturaleza gubernamental o académica. Es un sistema que permite acceder al documento al poner el DOI en la web y lo obtienes, al ser interoperable (los distintos sistemas reconocen el DOI y lo asocian únicamente a un material). También es permanente.
 - **ISBN** (International Standard Book Number): sistema numérico de identificación de materiales principalmente físicos.
 - **ISBN-A:** Al introducir el ISBN dentro del DOI, obtenemos el ISBN-A, que permite localizar el documento físico en las redes.

2. Plan de trabajo

Se describen las etapas en las que se organizará secuencialmente la investigación, las actividades que se realizarán en cada etapa (cronograma), y se detallan los aspectos metodológicos y estadísticos implicados en el estudio.

3. Recursos

Se exponen los recursos con los que se cuenta o que se necesitan (humanos, instalaciones, equipo...) y se elabora un presupuesto.

Finalmente, pueden añadirse **apéndices** con algunos de los documentos que se emplearán en el proyecto (la hoja de consentimiento informado, la hoja de información al paciente, las entrevistas y cuestionarios de evaluación que se emplearán, etc.).



REFERENCIAS BIBLIOGRÁFICAS

FONTES, S., GARCÍA-GALLEGO, C., QUINTANILLA, L., RODRÍGUEZ, R., RUBIO, P. Y SARRIÁ, E. (2010): *fundamentos de Investigación en Psicología*. UNED. Madrid.

Errores en los estudios epidemiológicos

Alba M^a Romero López, José Antonio Rueda Camino

1. Errores aleatorios	49
2. Errores sistemáticos (sesgos)	49

Enfoque OPE

Es un tema de muy baja importancia, con pocas preguntas, pero bastante agradecido. Entendiendo los conceptos más importantes de cada tipo de error, las preguntas suelen ser tener una dificultad baja.

Los errores los podemos clasificar en aleatorios y sistemáticos (sesgos). Mientras que los **errores aleatorios** van a afectar a la **validez externa** por no ser la muestra representativa; los **errores sistemáticos** van a afectar a la **validez interna** y, consecuentemente, a la validez externa, ya que no podremos extrapolar los datos, debido a que son erróneos ya en la muestra.



1. ERRORES ALEATORIOS

El hecho de que estudiemos muestras de individuos y no a la población completa implica que pueda ocurrir que nuestra **muestra no sea representativa de la población**. Igualmente si la **variabilidad** de nuestros datos es poca, los errores aleatorios **disminuirán (SCS 19,42)**.

- ✓ **Tipos de errores aleatorios:** error de tipo I (alfa) y error de tipo II (beta) (**ver tema 3.1. Errores en contraste de hipótesis**).
- ✓ **Consecuencia:** afectan a la **validez externa** del estudio (no afectan a la validez interna).
- ✓ **Solución:** al estar influidos por la cantidad de datos recogidos en la muestra, si **aumen-**

tamos el tamaño muestral (SERGAS 16, 67), tendremos más datos de la población y disminuirá el riesgo de cometer errores aleatorios, aumentando también la potencia estadística del estudio.



2. ERRORES SISTEMÁTICOS (SESGOS)

Son errores debidos a un **diseño inadecuado** del estudio.

- ✓ **Consecuencia:** afectan a la validez **interna** del estudio (por lo que, secundariamente, afectan también a la validez externa). La asociación que se establece será errónea, pudiendo **infraestimarse** o **sobreestimarse** dependiendo del tipo de sesgo:
 - **Infraestimación:** sesgos de selección, de clasificación incorrecta no diferencial y factor de confusión.
 - **Sobreestimación:** sesgos de clasificación incorrecta diferencial, atricción y factor de confusión.
- ✓ No se ven influidos por el tamaño muestral.

Sesgos de selección (INGESA 21, 1)

Estos sesgos se deben a la existencia de **diferencias en las características** de los distintos grupos a estudio (además de la característica estudiada), que acaban **influyendo en los resultados**. Por tanto, las características de los sujetos de la muestra van a diferir de las características de la población diana, y no serán representativos. Por ejemplo, un grupo es más anciano que el otro, lo que influye en que haya más muertes entre los sujetos de este grupo.

- ✓ **Solución: aleatorizar la formación de los grupos** disminuirá la probabilidad de que las características individuales de los pacientes se distribuyan de forma dishomogénea entre los grupos.

Ejemplos de sesgos de selección

- ✓ **Sesgo de autoselección (del voluntario):** cuando se recluta a pacientes voluntarios para participar en un estudio, suelen ser pacientes que no encuentran alivio con los tratamientos disponibles y buscan "a la desesperada" una solución. Dichos pacientes suelen por tanto estar **más graves** que la media, lo cual puede afectar a los resultados (cualquier fármaco tenderá a ser menos efectivo).
- ✓ **Sesgo del obrero sano:** si para estudiar una **enfermedad laboral** se acude al lugar de trabajo para seleccionar a los individuos, se infraestimarán la frecuencia de enfermedad, ya que aquellos sujetos enfermos no estarán trabajando sino de baja.
- ✓ **Sesgo diagnóstico (de Berkson):** sesgo que ocurre cuando se seleccionan los individuos de un estudio de entre **pacientes hospitalizados**, y el factor que se está estudiando es un factor de riesgo para hospitalizarse.

*Por ejemplo, se analiza la posible relación causal entre la tuberculosis y el cáncer de pulmón en pacientes hospitalizados mediante un estudio de casos (pacientes con cáncer) y controles (hospitalizados por otras causas), analizando la frecuencia de tuberculosis en cada grupo. Tanto la tuberculosis como el cáncer son por sí mismos factores de riesgo de ingresar en el hospital (si un individuo con tuberculosis o cáncer enferma por otras causas, es probable que le ingresen porque su enfermedad de base le convierte en un paciente de alto riesgo). Así, entre los pacientes "hospitalizados por otras causas" habrá un subgrupo de pacientes cuyo motivo de hospitalización será tuberculosis (mientras que el grupo de pacientes con cáncer ya tienen bastante con su tumor para ingresar, por lo que habrá menos proporción de pacientes con tuberculosis). Esto hará que **infraestimemos** la asociación entre cáncer y tuberculosis al encontrar más pacientes en el grupo control.*

- ✓ **Sesgo de incidencia/prevalencia (de Neyman):** este sesgo ocurre en los estudios de **casos y controles** cuando se estudian enfermedades con una fase aguda que tiene alta tasa de letalidad y una fase posterior crónica, puesto que solo podremos estudiar a los casos crónicos que han sobrevivido a la fase aguda (casos prevalentes), mientras que los casos agudos que fallecieron se nos pasarán desapercibidos (casos incidentes). Las características de los casos prevalentes pueden ser distintas a las de los casos incidentes, y ello podrá tener implica-

ciones en los resultados del estudio. *Por ejemplo estamos hablando de enfermedades como el IAM, ictus, o disección de aorta.*

Por ejemplo, al analizar la relación causal de la HTA con el ictus, infraestimaremos la asociación porque los ictus cuya causa es hipertensiva tienen una mayor letalidad que los ictus por otras causas. Así, los casos "prevalentes" de ictus tendrán una prevalencia de HTA menor que la del conjunto de pacientes con ictus (sumando los casos incidentes y los prevalentes), porque los pacientes hipertensos que sufrieron un ictus se murieron más que los no hipertensos.

Sesgos de clasificación (de información, de medida)

Aparece cuando **se clasifica erróneamente una variable** en estudio, pensando que los pacientes que presentan esa variable no la presentan o viceversa. *Por ejemplo, el esfigmomanómetro del estudio está estropeado y siempre marca PAS 150 mmHg; clasificaremos mal a los sujetos no hipertensos ya que pensamos que son hipertensos.*

Sesgo de clasificación incorrecta no diferencial

Este sesgo ocurre debido a **errores en los aparatos de medida**, y conllevan el mismo nivel de error en la clasificación de todos los grupos del estudio.

Estos errores tienden a diluir el efecto de la exposición o tratamiento que se estudia, por tanto, **infraestiman** la asociación, independientemente de si los resultados de la medición son por exceso o defecto.

La **solución** resulta evidente: **mejorar los aparatos de medida** (aumentar su sensibilidad y especificidad).

Sesgo de clasificación incorrecta diferencial

Aparece por errores **subjetivos** de los pacientes o del investigador a la hora de clasificar las variables resultado de los pacientes. Estos sesgos suelen aparecer **cuando el paciente o el investigador conoce a qué grupo pertenece el paciente**, de modo que los pacientes que reciban la intervención experimental o los investigadores pueden interpretar una falsa mejoría en **variables subjetivas** (bienestar, dolor, etc.) respecto de los pacientes del grupo control (que reciben o bien nada, o placebo, o un fármaco activo control): la clasificación de las variables es distinta en cada uno de los grupos.

Así, este sesgo **sobreestima** los resultados, ya que el grupo experimental verá sus variables resultado subjetivas artificialmente mejoradas.

Su solución consiste en emplear **técnicas de enmascaramiento (ciego)**:

- ✓ **Estudios abiertos:** los pacientes e investigadores conocen qué intervención reciben los pacientes.

✓ **Estudios ciegos (con enmascaramiento):**

- **Ciego simple:** los pacientes no saben qué intervención reciben (si la experimental o el control).
- **Doble ciego:** no lo saben ni los pacientes ni los investigadores.
- **Triple ciego (OPE):** no lo saben ni los pacientes, ni los investigadores ni los analistas de los datos (que suelen ser personas independientes de los investigadores).

Ejemplos de este sesgo son:

- ✓ **Sesgo de memoria:** en los estudios de casos y controles, el grupo de casos recuerdan más la exposición al factor de riesgo que los controles. Por ejemplo, si estudiamos teratogenia, recordarán mejor todo su embarazo aquellas mujeres con niños que la sufren (casos) que aquellas que no la han sufrido (controles).
- ✓ **Efecto Hawthorne:** los participantes de un estudio pueden actuar de modo distinto del habitual simplemente por sentirse observados.



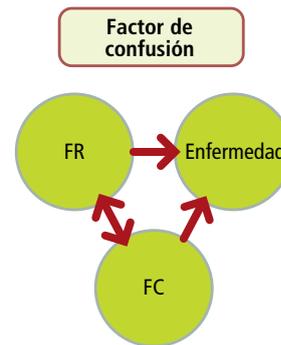
✚ **Figura 1.** Fábrica Hawthorne Works, de la compañía Western Electric. El efecto Hawthorne recibe su nombre por unos estudios sobre productividad industrial realizados en dicha fábrica de la localidad de Cicero (Illinois) entre 1924 y 1932. El estudio más famoso consistió en comparar la productividad industrial con iluminación ambiental más alta o más baja, y sus resultados fueron que la productividad aumentó tanto en el grupo sometido a alta iluminación, como en el sometido a baja iluminación; fue el hecho de saberse observados el que propició que los trabajadores aumentaran su productividad.

Sesgo por factor de confusión (FC)

Factor de confusión es todo aquel **factor (de riesgo o protector)** para la enfermedad a estudio, que tiene una **asociación estadística** a la presencia de la **exposición** cuya asociación queremos **estudiar**. Además, este factor de confusión actúa de forma independiente

a la exposición en el mecanismo de desarrollo de la enfermedad.

Como el FC y la exposición se asocian estadísticamente, los pacientes del grupo expuesto presentarán en mayor proporción el FC que los pacientes del grupo no expuesto, independientemente del método de selección de la muestra (aunque sea aleatorizada). Así, al suponer el FC un riesgo o protección para la aparición de enfermedad, **parte del efecto que atribuimos a la exposición se deberá realmente al FC (sobrestimando en algunas ocasiones o infraestimando en otras la verdadera asociación causal) (ver figura 2).**



✚ **Figura 2.** Relación entre el factor de riesgo, el factor de confusión y la enfermedad.

El sesgo por factor de confusión es el **único que puede eliminarse a posteriori:**

✓ **Soluciones a priori del sesgo por factor de confusión:**

- **Restricción:** el FC es un criterio de exclusión en el estudio. Por tanto, ningún paciente tendrá el FC presente, y no influirá en los resultados ni del grupo expuesto ni del no expuesto.
- **Apareamiento:** por cada paciente incluido en el grupo expuesto que posea el FC, incluiremos un paciente en el grupo no expuesto que lo posea. Así, el porcentaje de pacientes con el FC será el mismo en ambos grupos, y se eliminará la influencia del FC sobre la asociación exposición-enfermedad. Se utiliza sobre todo en el estudio de casos-controles.

✓ **Soluciones a posteriori del sesgo por factor de confusión:** se realiza mediante 3 tipos de análisis estadístico, que permiten calcular exactamente la parte del riesgo inicial que se le atribuye al FC, para eliminarlo. Tenemos: análisis de subgrupos, análisis estratificado y análisis multivariante.

Sesgo de atrición

Ocurre cuando existen **diferencias en el porcentaje de pérdidas** observadas en los distintos grupos

de un estudio prospectivo, y las pérdidas no se incluyen en el análisis estadístico de los resultados.

Habitualmente, si los dos grupos son homogéneos (no tienen características diferentes), la diferencia en las pérdidas se deberá a un efecto adverso del tratamiento experimental por el cual los sujetos de dicho grupo abandonan más el estudio que los del otro grupo.

Solución: estudiar los resultados utilizando un **análisis por intención de tratar**, en lugar del análisis por protocolo.

- ✓ **Análisis por protocolo:** sólo se estudian los resultados de los pacientes que finalizan el estudio. Si existe diferente proporción de pérdidas entre el grupo experimental y el control y

estas pérdidas se deben a efectos adversos del fármaco experimental, **sobreestimaremos** el beneficio del fármaco experimental al no tener en cuenta a los pacientes que tienen que dejar de tomarlo (y por tanto dejan de beneficiarse de él).

- ✓ **Análisis por intención de tratar:** se estudian los resultados de todos los pacientes **reclutados** en el estudio (todos los pacientes “**aleatorizados**”), incluyendo a los pacientes que cursen pérdida. Permite estudiar la causa de las pérdidas y el efecto global del fármaco teniendo en cuenta que un porcentaje de pacientes no se lo tomará (situación que simula a la que se observará en la **práctica clínica real** una vez se comercialice el fármaco).

Error	Consecuencia	Solución
Errores aleatorios	↓ Validez externa	↑ n
Errores sistemáticos	↓ Validez interna → ↓ Validez externa	Mejorar diseño
Selección	Habitualmente infraestima	Aleatorización
Clasificación incorrecta no diferencial	Infraestima la asociación	Mejorar aparatos de medida
Clasificación incorrecta diferencial	Sobreestima la asociación	Enmascaramiento
Factor de confusión	Sobreestima o infraestima la asociación	A priori: - restricción - apareamiento A posteriori: - análisis de subgrupos - análisis estratificado - análisis multivariante
Atrición	Sobreestima la asociación	Análisis por intención de tratar

Figura 3. Errores en los estudios epidemiológicos