

**Departamento Salud Pública y Materno-Infantil**

**Facultad de Medicina**

**Universidad Complutense de Madrid**



**Colección**

**“Guías para la iniciación a la investigación en Ciencias de la Salud”**

**Nº 1: *Fundamentos para la iniciación a la investigación en Ciencias de la Salud. Epidemiología***

**Editado por la Unidad Docente de Medicina Preventiva y Salud Pública**

**Madrid, 2018**

Editado por la Unidad Docente de Medicina Preventiva y Salud Pública  
Departamento de Salud Pública y Materno-infantil  
Facultad de Medicina  
Universidad Complutense de Madrid  
Madrid, 2018  
**ISSN 2605-5252**

## LISTADO DE AUTORES

- **ALBALADEJO VICENTE, R**
- **ASTASIO ARBIZA, P**
- **CALLE PURÓN, ME**
- **CEA SORIANO, L**
- **HOYOS MILLER, J**
- **MARTÍNEZ HERNÁNDEZ, D**
- **MARTÍNEZ SANTOS, A**
- **ORTEGA MOLINA, P**
- **PULIDO MANZANERO, J**
- **REGIDOR POYATOS, E**
- **SANTOS SANCHO, JM**
- **SORDO DEL CASTILLO, L**
- **VILLANUEVA ORBAIZ, R**

## ÍNDICE DE CAPÍTULOS

<b>CAP. 1.</b>	Secuencia de la Investigación en Ciencias de la salud.....	5
<b>CAP. 2.</b>	Hipótesis y objetivos.....	9
<b>CAP. 3.</b>	Cómo elaborar un protocolo de investigación.....	16
<b>CAP. 4.</b>	El ABC de la Epidemiología.....	21
<b>CAP. 5.</b>	Muestreo y tamaño muestral.....	41
<b>CAP. 6.</b>	Diseño y administración de instrumentos de recogida de datos.....	49
<b>CAP. 7.</b>	Datos Secundarios. Utilización de Bases de Datos.....	57
<b>CAP. 8.</b>	Revisiones sistemáticas y Meta-análisis. Conceptos básicos.....	63
<b>CAP. 9.</b>	Investigación de un brote epidémico.....	69
<b>CAP. 10.</b>	Farmacovigilancia. Farmacoepidemiología .....	76

# **CAPÍTULO 1**

## **Secuencia de la Investigación en Ciencias de la salud**

**Autores: Paloma Ortega Molina  
Paloma Astasio Arbiza  
Lucia Cea Soriano**

## Secuencia de la Investigación en Ciencias de la salud

Por regla general, la investigación nace de la curiosidad e inquietud personal o de la observación de un hecho para el que no encontramos explicación lógica o contradice las teorías aceptadas. Normalmente nos planteamos investigaciones relacionadas con la evaluación de intervenciones para mejorar la salud, prevenir la enfermedad o determinar sus efectos en un individuo o en una población. Por tanto, investigar es algo más que recoger, almacenar y analizar datos.

Si nos preguntamos, ¿qué es investigar?, podemos decir que: “resolver un problema o responder a una pregunta” aplicando el método científico de manera rigurosa y sistemática, permitiéndonos adquirir nuevos conocimientos. Por tanto, se trata de un proceso que se debe realizar de forma **sistemática, organizada y objetiva**.

El término **sistemático** significa que el método científico se aplica siempre siguiendo una secuencia establecida y de la misma manera. Partimos de la observación de un problema o situación de salud que nos llama la atención y de la revisión de los conocimientos existentes sobre el mismo, nos formulamos una hipótesis y unos objetivos concretos que trataremos de alcanzar mediante un diseño de estudio epidemiológico y un plan de trabajo preestablecido. Para ello, recogeremos la información que nos interesa en la población seleccionada, posteriormente analizaremos y obtendremos unos resultados que, tras su interpretación, generarán unas conclusiones que podrán mejorar los conocimientos existentes o generar nuevas hipótesis (Figura 1).

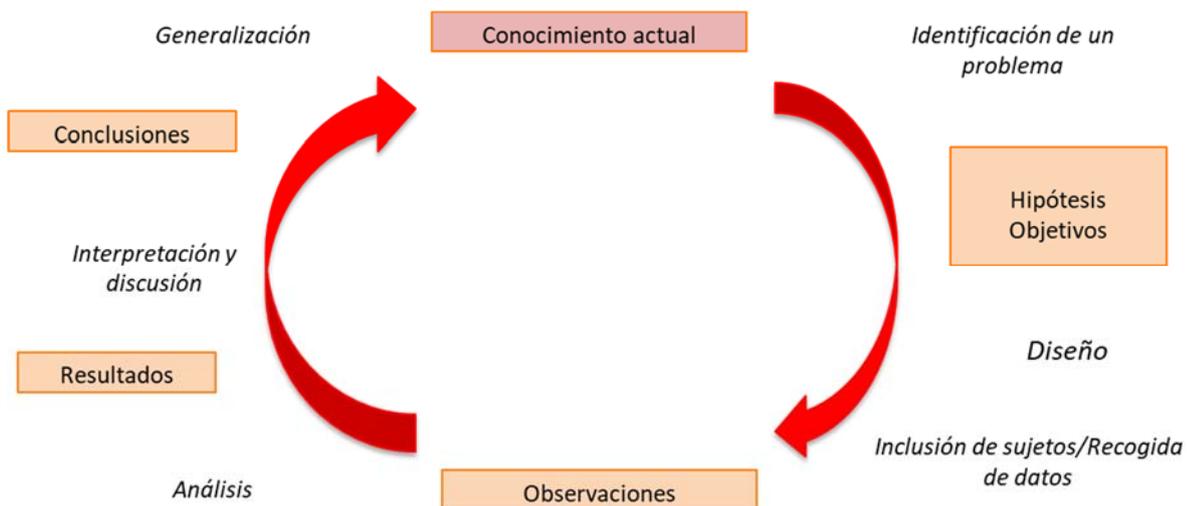


Figura 1. Adaptado de Metodología de investigación clínica y epidemiológica 4ª edición. 2013. Josep M<sup>a</sup> Argimon Pallás & Josep Jiménez Villa

Por **organizado** se entiende que todas las personas que participan en la investigación forman un equipo cohesionado y lo hacen siguiendo un protocolo de estudio, con unanimidad de criterios previamente establecidos y actuando de la misma manera ante situaciones similares. La cohesión de los miembros del

equipo de investigación y el protocolo de la misma son esenciales para que nuestro estudio llegue a buen puerto obteniendo conclusiones válidas.

El término **objetivo** indica que las conclusiones que obtenemos (extraemos) se basan en los resultados que hemos observado, medido y analizado. Las conclusiones no deben estar sujetas a interpretaciones subjetivas y debemos huir de interpretaciones que no han sido contrastadas con los resultados obtenidos.

Resumiendo, la respuesta a una pregunta de investigación va a seguir el esquema de la figura 2, siguiendo la dirección de la flecha. Plantearemos una investigación que abordaremos mediante un diseño de estudio, en el que mediremos distintas variables de interés en una población y posteriormente compararemos resultados y, por último, los interpretaremos.

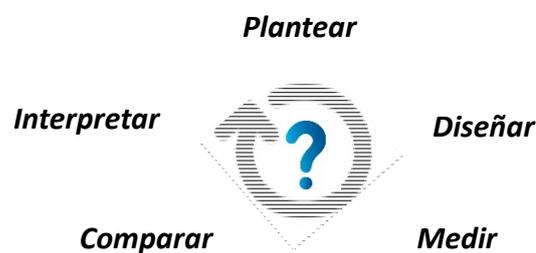


Figura 2. Secuencia a seguir ante una pregunta de investigación

### **Aspectos relevantes al inicio de la investigación**

El abordaje de un estudio de investigación requiere, por parte del investigador, de una actitud crítica y un pensamiento claro y organizado que le permitan responder a una pregunta concreta, valorando su pertinencia y viabilidad.

Toda investigación con seres humanos, investigación epidemiológica, tiene un antes y un después. El “antes” hace referencia al inicio de la misma, momento en el que se concibe el protocolo o proyecto de investigación. El “después”, una vez finalizado el estudio, consiste en dar difusión a nuestros resultados a través de una comunicación o una publicación, ya que estamos obligados a transmitirlo a la comunidad científica y a la sociedad en general.

Los aspectos más importantes en la elaboración del protocolo o proyecto de investigación se detallan en capítulos posteriores y se pueden resumir en:

1. Antecedentes y Justificación. Debemos dejar reflejados los conocimientos existentes hasta el momento sobre nuestra pregunta de investigación, justificando su pertinencia y relevancia a partir de la revisión de la literatura.

2. Hipótesis y Objetivos. Debemos elaborar una hipótesis clara, viable, concisa y contrastable para responder a nuestra pregunta de investigación. Que trataremos de confirmar o rechazar mediante la formulación de unos objetivos concretos y coherentes con la hipótesis.
3. Diseño y métodos de estudio (investigación). Debemos decidir el tipo de estudio epidemiológico que vamos a aplicar para satisfacer los objetivos planteados; cuál va a ser nuestra población de estudio, definiendo los criterios de inclusión y exclusión y calculando el tamaño muestral; seleccionando y definiendo las variables que queremos estudiar, elaborando el cuestionario de recogida de datos; y el procedimiento de análisis de datos que vamos a aplicar.
4. Reflexionar y controlar los aspectos éticos, limitaciones y posibles sesgos que pueden presentarse en el estudio.
5. Gestión del estudio: Duración, cronograma, seguimiento, recursos humanos y materiales.
6. Bibliografía relevante en la que nos hemos basado para realizar el protocolo.

### **Aspectos relevantes una vez realizada la investigación**

Como ya hemos comentado, una vez finalizado el estudio debemos dar difusión a nuestros resultados a través de una comunicación o una publicación, ya que estamos obligados a transmitirlo a la comunidad científica y a la sociedad en general.

La estructura general de una publicación científica puede esquematizarse en los siguientes apartados:

1. Título. Autores. Palabras clave
2. Resumen
3. Introducción
4. Material y Métodos
5. Resultados
6. Discusión
7. Bibliografía

Cada uno de estos apartados es tratado detalladamente en capítulos posteriores.

### **Bibliografía**

1. **Argimon J, Jiménez J.** Metodos de Investigacion Clinica y Epidemiologica (4ª Ed.) Elsevier. Barcelona 2013. Acceso electrónico: <https://www.elsevierlibrary.es/pdfreader/mtodos-de-investigacin-clnica-y-epidemiolgica52638>
2. **Greenberg RS.** Medical Epidemiology. 4ª ed. Lange. New York, 2013.
3. **Hulley S, Cummings S, Browner W, Grady D, Newman T.** Diseño de investigaciones clínicas (4ª Ed). Wolters Kluwer Lippincott Williams & Wilkins. Barcelona, 2014

## **CAPÍTULO 2**

### **Hipótesis y Objetivos**

**Autores:** Romana Albadalejo  
Rosa Villanueva  
M<sup>a</sup> Elisa Calle



## Objetivos

Al finalizar la lectura de este documento el alumno deberá:

1. Describir la hipótesis de un estudio de investigación
2. Determinar el objetivo general /es de un estudio
3. Seleccionar los objetivos secundarios, que permitirán delimitar la metodología con la que se alcanzará el objetivo general

## Introducción

Uno de los mayores logros de la civilización occidental, ha sido el desarrollo de un método científico de investigación, basado en la formulación de preguntas y el examen crítico de hipótesis. Las preguntas plantean los problemas y las hipótesis proponen unas respuestas que la investigación tratará de verificar. Tanto esta identificación del problema, como de la formulación de la hipótesis y de los objetivos a alcanzar, forman parte de la planificación del estudio, siendo el protocolo de investigación el documento escrito en el que se plasman estas cuestiones.

## Desarrollo

### *Características de la pregunta de investigación*

La pregunta de investigación debe cumplir con una serie de características entre las que se encuentran que sea relevante, concreta y viable.

#### 1.-Relevante.

Esta pregunta debe tener un impacto positivo para la sociedad y la salud: puede constituir un avance para la ciencia, influir en la política sanitaria y la práctica clínica, o ser el origen de posteriores investigaciones. Su potencial beneficio en enfermedades de alta prevalencia, graves, emergentes, de alto coste o que generen grave preocupación política y social, será siempre fundamental.

#### 2.-Concreta y explícita.

La pregunta deberá tener una respuesta explícita y determinada.

#### 3.-Viable

Es mejor conocer desde el inicio los límites prácticos y los problemas que pueden surgir, antes de invertir tiempo o dinero en líneas en las que no se podrá trabajar.

Por tanto, tendrá que ser factible, esto es susceptible de respuesta con el conocimiento actual, y su viabilidad abarcará también a la disponibilidad de sujetos, instalaciones, equipos y ética.

A continuación, vamos a ver un ejemplo práctico de una pregunta bien planteada

“¿Es evitable la biopsia hepática para la cuantificación de la concentración de hierro en hígado en el diagnóstico de hemocromatosis?”.

1.- ¿Es relevante?

Sí, puesto que se evitará la biopsia, que tiene riesgos para el paciente y un coste económico para el sistema sanitario.

2.- ¿Es concreta y explícita?

Sí, ya que se centra en el contexto clínico de un proceso patológico específico, la hemocromatosis. Por otro lado, es explícita, ya que de ello se deriva que se evite o no la biopsia.

3.- ¿Es viable?

Es viable, ya que están perfectamente determinadas las indicaciones de la biopsia hepática para la medición de la concentración de hierro en la hemocromatosis, así como la capacidad de la resonancia magnética para determinar esta concentración.

### ***Desarrollo de la hipótesis de investigación***

#### *Concepto de hipótesis*

Una vez que hemos resuelto la pregunta de investigación, tendremos que definir la hipótesis del trabajo, que será una “conjetura plausible”. Esto es, frente a la pregunta que ha surgido de un problema, la hipótesis será la respuesta más satisfactoria que se puede proponer, según el estado actual del conocimiento. Se situará, por lo tanto, al inicio del trabajo y se llamará “hipótesis a priori”, siendo como una apuesta previa a la obtención y análisis de los datos. El análisis de estos datos podrá replantear el formular otra hipótesis de trabajo o “hipótesis a posteriori”, que se contrastará en futuros estudios.

#### *Origen de la hipótesis*

Según el método científico, las hipótesis se pueden derivar de la aplicación de un razonamiento inductivo o deductivo.

El método de razonamiento inductivo se basa en la observación de un problema concreto, a través de unos datos, y proviene de la generalización de las relaciones observadas en la realidad. Parte de lo más específico y se llega a generalizaciones o teorías más amplias.

A su vez, el método deductivo emana de la teoría y sirve de punto de partida al investigador, para verificar su validez en la práctica. Funciona partiendo desde lo más general hacia lo más específico. Esta metodología hace que, si la hipótesis derivada de un marco conceptual no se confirma, se pueda cuestionar la validez de la teoría, pero también las limitaciones y validez del diseño del estudio realizado.

Finalmente, se podrán identificar problemas y formular hipótesis tanto de las comunicaciones y publicaciones científicas, como de la experiencia profesional, a lo que se podría añadir la curiosidad, la imaginación, la intuición y el escepticismo, que siempre contribuyen al progreso del saber.

Ejemplo de ambos razonamientos, aplicados a la formulación de las primeras hipótesis sobre el origen del SIDA:

1. Razonamiento inductivo: si dos o más personas, en las que se produce la enfermedad, tienen un factor común (hemofílicos, transfundidos, adictos a drogas por vía parenteral), mientras que dos o más de los que no enfermaron no presentan dicho factor, esa circunstancia que difiere en los dos grupos, es causa de la enfermedad.
2. Razonamiento deductivo: la observación de que la clínica de los enfermos del SIDA, era muy similar a la de un cuadro de inmunodeficiencia producido por un retrovirus en simios, llevó a postular que el VIH era un retrovirus.

#### *Tipos de Hipótesis.*

Las hipótesis se pueden clasificar según diferentes criterios, no siempre excluyentes, sino complementarios.

**Hipótesis conceptual.** Es la explicación teórica del problema, que intenta facilitar la solución práctica. Un ejemplo de esta será: “la resonancia magnética puede cuantificar la concentración de hierro en hígado de manera fiable, por lo que puede reemplazar a la biopsia hepática en el diagnóstico de hemocromatosis”.

**Hipótesis operativa.** Es una proposición que establece una relación entre una causa y un supuesto efecto, y que se puede verificar empíricamente. Plantea un modo concreto de verificación de la hipótesis conceptual. Se deberá redactar en una frase corta y afirmativa que se demostrará experimental o clínicamente. Debe ser evaluable y medible. Así, podríamos decir que “comparando las medidas de la resonancia magnética con la concentración de hierro en hígado, medida por biopsia hepática, se podrá hacer una calibración entre ambos parámetros, que permita en un futuro determinar la concentración de hierro mediante la resonancia magnética”.

**Hipótesis estadística.**

Para comprobar o contrastar una hipótesis, se deben aplicar las correspondientes pruebas de significación estadística y éstas requieren de una formulación conocida como *hipótesis nula* ( $H_0$ ). Según esta hipótesis “las diferencias observadas en el estudio son debidas únicamente al azar”. Niega, por lo tanto, que existan

diferencias entre los grupos, y, es la hipótesis que se pone a prueba directamente con el contraste estadístico. Se establece con el propósito de ser rechazada.

Por otro lado, se formula la *hipótesis alternativa* ( $H_1$ ) que establece que “existen diferencias reales entre los grupos comparable.

En la práctica, la comprobación para aceptar o rechazar estas hipótesis se basa en la aplicación de las pruebas de contraste de hipótesis, teniendo en cuenta que el nivel de significación que permite rechazar la  $H_0$  (o aceptar la  $H_1$ ), suele fijarse en el 0,05. Si se establece este nivel de significación estadística, la probabilidad de rechazar erróneamente una  $H_0$  verdadera es inferior o igual a 0,05 ( $p \leq 0,05$ ).

### **Formulación de objetivos**

Los objetivos de una investigación expresan los fines o los propósitos que se esperan alcanzar con el estudio del problema planteado. Por tal razón, se dice que los objetivos constituyen la finalidad de la investigación. Estos deben responder a la pregunta: ¿Qué se pretende alcanzar con la investigación? Por ello, es habitual que su redacción comience con un verbo en infinitivo y es conveniente definirlos después de haber elaborado el fundamento teórico y de tener clara la hipótesis.

Puede haber un objetivo primario o general, que debe responder explícitamente a la pregunta principal y tener una relación directa con el título del proyecto. Además, se podrán incluir uno o más objetivos secundarios, específicos, con un mayor nivel de exactitud y que permitirán delimitar la metodología adecuada.

Estas características de los objetivos determinan cómo deben ser redactados. La correcta exposición, redacción y determinación de los objetivos es clave en todo trabajo científico.

Los objetivos deben ser:

- Precisos: significa que se deben expresar de forma clara, con lenguaje sencillo, evitando ambigüedades.
- Concisos: se deben formular de la manera más resumida posible, sin rodeos, utilizando solo las palabras necesarias.
- Medibles: deben expresarse de modo tal que permitan medir las cualidades o características que caracterizan el objeto de investigación.
- Alcanzables: deben existir posibilidades reales de lograrlos.

Así, en nuestro ejemplo podríamos redactar como objetivo principal “Evaluar la capacidad de la resonancia magnética para cuantificar la concentración hepática de hierro en la hemocromatosis”.

Ejemplos de objetivos secundarios serían “elegir una población con hemocromatosis”, “identificar su concentración de hierro con biopsia hepática”, “aplicar la resonancia magnética a estos pacientes”, “comparar la concentración de hierro por ambas técnicas”, etc...

Redacción de los objetivos:

<b>Sí usar: verbo en infinitivo que indique una acción explícita</b>	
<i>Generales</i>	<i>Específicos</i>
Describir	Seleccionar
Determinar	Identificar
Demostrar	Evaluar
Comprobar	Separar
Evaluar	Aplicar
<b>No usar: verbos no concretos</b>	
Conocer	Comprender
Estudiar	Saber

### Resumen

La pregunta de investigación debe cumplir con una serie de características entre las que se encuentran que sea relevante, concreta y viable.

Una vez que hemos resuelto la pregunta de investigación, tendremos que definir la hipótesis del trabajo, que será una “conjetura plausible”, es decir, será la respuesta más satisfactoria que se pueda proponer en función del estado actual del conocimiento.

Según el método científico, las hipótesis se pueden derivar de la aplicación de un razonamiento inductivo, que parte de lo más específico y llega a generalizaciones más amplias, o del razonamiento deductivo, que funciona partiendo de lo más general hacia lo más específico.

Las hipótesis se pueden clasificar, según diferentes criterios, en conceptual, operativa y estadística (nula o alternativa).

Los objetivos de una investigación expresan los fines o los propósitos que se esperan alcanzar con el estudio del problema planteado. Puede haber un objetivo primario o general y uno o más objetivos secundarios. Estos objetivos deben ser precisos, concisos, medibles y alcanzables.

### Bibliografía

1. J.M. Alústiza Echeverría, E. Salvador Pardo y A. Castiella Eguzkiza. ¿Cómo establecer la hipótesis y los objetivos de un proyecto de investigación en Radiología? Radiología 2012: 54(1):3-8
2. S. Hulley, S. S. Cummings, W. Browner, D. Grady, T. Newman. Diseño de investigaciones clínicas, 4ª ed., Wolters Kluwer Lippincott Williams & Wilkins, Barcelona, 201
3. MT Icart Isern, J. Canela Soler. El uso de la hipótesis en la investigación científica. Atención Primaria 1998;21:172-178

## **CAPÍTULO 3**

### **Cómo elaborar un protocolo de investigación**

**Autores:** Luis Sordo del Castillo  
Romana Albaladejo  
Juan Hoyos Miller

**Objetivos:** Establecer las pautas básicas mínimas para la redacción de un protocolo de investigación

### **Introducción: Definición y uso**

Un protocolo de investigación es un documento en el que se describen los pasos a seguir para poder conducir el proceso de investigación. Es la hoja de ruta previa a iniciar una investigación que permite a los lectores tener conocimiento de la secuencia de actividades científicas que van a seguir los investigadores.

Un protocolo de investigación debe quedar redactado con suficiente claridad en todos sus apartados de tal forma que una persona que no sea de la misma área científica comprenda totalmente lo que se quiere investigar, de qué forma, en qué tiempo o circunstancias

### **1. Normas generales del protocolo**

Antes de ir paso a paso por las diferentes partes que componen un protocolo de investigación, se resaltarán tres características recomendadas en la escritura de los mismos:

1. Se redactará todo de forma impersonal (“se analizará, se evaluará” evitando la escritura en primera persona “nuestro estudio, nuestro centro”).
2. Los apartados deben escribirse evitando maneras “telegráficas”. Se trata de un documento de uso interno (equipo investigador) y externo (posibles evaluadores), por lo que debe ser completamente inteligible en cualquier momento del proceso.
3. Solo se recomienda el empleo de siglas o acrónimos cuando designen a términos que van a ser muy utilizados. En tal caso, se debe describir lo que significan la primera vez que se usan seguido, entre paréntesis, de las siglas que sustituirán en lo sucesivo a tal concepto. Son una excepción aquellos términos que ya están admitidos en forma de siglas (como puede ser SIDA)

### **2. Partes que componen un protocolo de un proyecto de investigación y sus características principales**

2.1. Título: Un buen título es aquel en el que se especifica la acción que se va a llevar a cabo (evaluación, comparación, determinación, variabilidad...), sobre qué tipo de pacientes se va a llevar a cabo y si es posible el ámbito (centro de Salud, hospital de tercer nivel, distrito...)

Así mismo, en el título deben figurar la mayor cantidad de palabras descriptoras de la temática del estudio, obviándose otras palabras como “estudio” o “informe” que no aportan nada. Un buen título debe ser corto, preciso y conciso, de ahí que se recomiende que su tamaño final no exceda de las 15 palabras.

2. 2. Investigador principal y colaboradores: Se deben listar las personas responsables de llevar a cabo la investigación, con datos de filiación al centro de trabajo y/o grupo de investigación, grado académico y contacto.

2. 3. Resumen: En pocas palabras ha de explicarse todo el protocolo. Deben estar redactados de manera que su lectura sea sencilla sin perder la rigurosidad científica y deben permitir ser entendidos incluso por un lector no especializado en el tema. Debe ser autoexplicativo. Que no haga referencias al texto en el que se basa, que pueda ser entendido sin recurrir a él. Los resúmenes constan de los mismos apartados que el protocolo, si bien a veces quedan reducidos a: Introducción/Antecedentes, Objetivo, Diseño, Ámbito de estudio, sujetos de estudio, Instrumentos y Limitaciones. Los apartados de este dependen de las normas que nos exijan para cada protocolo.

2. 4. Palabras clave: Los protocolos deben estar definidos por aquellas palabras que concretarán su posible búsqueda en un repositorio. La eclosión de las bases de datos de artículos, proyectos y protocolos hace que este apartado tenga mucha relevancia. Hemos de seleccionar aquellas palabras que mejor se ajustan a la investigación a llevar a cabo. Se recomienda que sean al menos 5, en inglés y que, de ser posible, hayan sido definidas como términos clave en el repositorio más importante en ciencias biomédicas que es PubMed.

2. 5. Antecedentes: Este apartado es el que establece el estado actual del tema a desarrollar o sobre el que trabajar. Se recomienda incluir cifras de prevalencia e incidencia de los estudios encontrados, que nos servirán para contextualizar el tema a tratar. Todo con las referencias bibliográficas necesarias. En este apartado deben esbozarse igualmente las carencias en el conocimiento que serán la base del siguiente apartado (Justificación)

2. 6. Justificación del estudio: El motivo que hace que se realice el estudio. Los motivos son muy variados, y es aquí donde debe especificarse lo que nos mueve al esfuerzo investigador.

Los antecedentes y la justificación han de tener una estructura tal que, tras leerlos, cada aspecto del estudio esté claro y el lector ya intuya los objetivos que va a leer

2. 7. Bibliografía: lo más actualizada posible y acorde con el tema que se quiere investigar. Se deberían incluir citas de 5-10 años como máximo de antigüedad. El formato de las referencias debe ser homogéneo (por ejemplo, siguiendo las normas de Vancouver)

2. 8. Hipótesis: Es la expresión de lo que se espera al realizar el estudio. Suelen redactarse en términos de una igualdad o afirmación: “el tratamiento A es más efectivo que el tratamiento B”, y es recomendable redactarlos en paralelo a el/los objetivo(s) principal(es)

2. 9. Objetivos:

- Objetivo general: El objetivo general debe ser muy parecido al título del proyecto, pero redactado en infinitivo. Se recomienda el uso de verbos como: determinar, comparar o analizar, evitando verbos como investigar o estudiar que son más imprecisos. Suele ponerse solo un objetivo general.

- Los objetivos específicos son un desarrollo más completo del objetivo general y comprenden acciones que no están descritas en éste, procurando no duplicar en los objetivos específicos lo ya expuesto en el objetivo general. Se suelen poner entre tres y seis y se ordenan del menos complejo al más complejo.

2. 10. Metodología: Es el apartado en el que se desglosan las actividades que van a desembocar en la consecución de los objetivos marcados. Está dividido en diferentes subapartados:

- A. Diseño del estudio: se debe describir de forma clara y concisa el tipo de diseño a realizar (experimental u observacional, longitudinal o transversal, prospectivo o retrospectivo y el diseño específico, caso-control, cohortes...)
- B. Población de referencia, diana y de estudio: La de referencia es la población global a la que se espera extrapolar los resultados; La población de diana es un subgrupo de la de referencia, y es la población que tiene que cumplir una serie de criterios de selección; y la población de estudio es aquella que estará formada por los individuos que se incluyen en el estudio.
- C. Criterios de inclusión y exclusión: se especificarán las características de los pacientes que entrarán en el estudio para que la selección se haga de la forma más homogénea posible y evitar sesgos.
- D. Tamaño muestral y procedimiento de muestreo: En el tamaño muestral se especifica el número de participantes que formarán parte del estudio. Debe estar descrito cómo se ha calculado éste. Si no hay posibilidad de calcularlo, se puede describir en base al número de pacientes que se pueden reclutar en el período de tiempo que dure el estudio. La técnica de muestreo describe no el número, sino la técnica que se utilizará para la recogida de la muestra (aleatorio simple, por conveniencia, estratificado...).
- E. Variables analizadas: Aquella información que va a recogerse en cada persona. En la descripción de las variables, debe haber varias independientes y una o pocas dependientes (la variable principal del estudio). Las variables, tanto dependientes como independientes, que sean categóricas, irán acompañadas de todas las categorías de medida posible, no dejando cabida a la ambigüedad.
- F. Recogida de datos y fuentes de información: especificar cómo se van a llevar a cabo las mediciones sobre los pacientes, quién las va a realizar, dónde, cuándo, qué instrumentos de medida o fuentes de información se usarán y qué hoja de recogida de datos (ésta debe incluirse en los anexos del protocolo).
- G. Análisis de datos: En este apartado se especifican los análisis estadísticos previstos.
- H. Dificultades y limitaciones del estudio: Las dificultades y limitaciones son una autocrítica a la realización del proyecto, donde se pueden especificar qué problemas podemos encontrarnos al

realizar el proyecto, tipos de sesgos. Tras describir los problemas, es conveniente explicar que estos no tienen por qué modificar lo que queremos medir, o resaltar que es inevitable que ocurran.

2. 11. Plan de trabajo: Se trata de un cronograma temporal donde se especificarán las distintas tareas a realizar en el proyecto indicando las fechas (trimestres, semanas, años...) y las personas que las realizarán.

2. 12. Experiencia del equipo investigador sobre el tema: Sin caer en glosar el currículum de cada investigador, sí es conveniente explicar su idoneidad para el proyecto presentado. Tanto en el conocimiento del área como en su capacidad para llevarlo a cabo.

2. 13. Aplicabilidad y utilidad práctica de los resultados: Aquí es donde se explica la utilidad del resultado de lo que se obtendrá en el proyecto y para qué servirá. Se debe redactar en términos prácticos y realistas

2. 14. Medios disponibles para la realización del proyecto: se especifican tanto los recursos humanos como los materiales de que se dispone (ordenadores, tensiómetros, personal de enfermería...) y para los que no se va a pedir financiación

2. 15. Anexos - En los anexos del protocolo se suelen incluir todo lo que puede ayudar a comprenderlo mejor como la hoja de recogida de datos, consentimiento informado...

## **Resumen**

Un protocolo de investigación es la hoja de ruta previa a iniciar una investigación que permite a los participantes en la misma y sus evaluadores tener conocimiento de la secuencia de actividades científicas que van a seguir los investigadores. Sus apartados son antecedentes, objetivos, hipótesis y metodología. Un buen protocolo es a la investigación lo que un plano a un edificio a construir.

## **Bibliografía recomendada**

1: Rodríguez del Aguila MM, Pérez Vicente S, Sordo del Castillo L, Fernández Sierra MA. How to elaborate a protocol on health research. *Med Clin (Barc)*. 2007. 8;129(8):299-302.

2: Dekkers OM, Cevallos M, Bühner J, Poncet A, Ackermann Rau S, Perneger TV, Egger M. Comparison of noninferiority margins reported in protocols and publications showed incomplete and inconsistent reporting. *J Clin Epidemiol*. 2015 May;68(5):510-7.

3: Rosenthal R, Schäfer J, Briel M, Bucher HC, Oertli D, Dell-Kuster S. How to write a surgical clinical research protocol: literature review and practical guide. *Am J Surg*. 2014;207(2):299-312.

4: Staessen JA, Bianchi G. Registration of trials and protocols. *Lancet*. 2003. 27;362(9389):1009-10.

## **CAPÍTULO 4**

### **El ABC de la Epidemiología**

**Autores:** Paloma Ortega Molina  
Paloma Astasio Arbiza  
Lucia Cea Soriano  
Enrique Regidor Poyatos

# 1. Concepto y Objetivos de la Epidemiología

## 1.1 Concepto

Etimológicamente: Epi = sobre  
Demos = pueblo  
Logos = tratado

“Ciencia que estudia los fenómenos de salud y enfermedad en las poblaciones”.

Se ha ido definiendo en base a la evolución e importancia de las causas de muerte, y en este sentido podemos comentar la definición de tres autores:

Stallybras

Ciencia que considera las *enfermedades infecciosas*, sus causas, propagación y prevención.

McMahon

Ciencia que se ocupa del estudio de las *enfermedades* y de los **factores que determinan su frecuencia y distribución** en las poblaciones

Rey Calero

Estudio de los *fenómenos de salud y enfermedad* en grupos sociales, así como los **factores que determinan su frecuencia y distribución**

La Epidemiología tiene, por tanto, **tres componentes** o pilares fundamentales:

1. Frecuencia: Cuantificación
2. Distribución: Quien enferma  
Cuándo enferma  
Dónde enferma
3. Determinantes de la enfermedad

Consecuentemente podemos definir la **Epidemiología** como la “Ciencia que estudia la frecuencia y distribución de las enfermedades, así como sus determinantes en las poblaciones”, y podemos hablar de 3 tipos de Epidemiología:

1. Epidemiología Descriptiva: Cuantificación y distribución de la enfermedad
2. Epidemiología Analítica: Estudio de los factores determinantes
3. Epidemiología Experimental: Establecemos medidas de control

## 1.2 Objetivos

1. Medir la frecuencia de la enfermedad
2. Describir los patrones de distribución
3. Investigar los brotes de una enfermedad
4. Identificar nuevos síndromes
5. Conocer la historia natural de la enfermedad
6. Determinar la validez de test diagnósticos y cribado (screening)
7. Identificar causas de enfermedad
8. Evaluar eficacia y seguridad de los tratamientos y acciones sanitarias
9. Predecir en términos probabilísticos el pronóstico de la enfermedad
10. Vigilancia epidemiológica. Farmacovigilancia.
11. Planificación sanitaria
12. Razonamiento y Metodología de investigación

## 2. Medidas de Frecuencia

Pretenden cuantificar la aparición de la enfermedad o evento en los grupos poblacionales

### 2.1 Prevalencia

Probabilidad de que un individuo esté enfermo en un tiempo determinado.  
Frecuencia de un evento existente ya que indica el nº de casos en una población.  
Prevalencia de punto: casos en un momento concreto  
Prevalencia de periodo: casos durante un periodo de tiempo determinado

$$P = \text{Nº de casos} / \text{Población total} \quad \text{casos} = \text{viejos} + \text{nuevos}$$

### 2.2 Incidencia

Cuantifica los casos nuevos de la enfermedad  
Frecuencia de un evento nuevo  
Dos tipos

- Incidencia Acumulada (IA) o Riesgo de Incidencia
- Densidad de Incidencia (DI) o Tasa de Incidencia

→ *Incidencia Acumulada (IA) o Riesgo de Incidencia (RI)*

Proporción de sujetos que desarrollan la enfermedad en un periodo de tiempo.

$$IA = \text{Nº casos nuevos en un tiempo} / \text{población total a riesgo}$$

Todos los miembros de la población, o cohorte, están libres de la enfermedad en el inicio de la observación.

“Riesgo de que un individuo desarrolle la enfermedad en un periodo de tiempo”

→ *Densidad de Incidencia (DI) o Tasa de Incidencia (TI)*

Refleja la ocurrencia de los casos nuevos de enfermedad y **mide la rapidez** con la que son diagnosticados nuevos pacientes en un tiempo.

Para estimar la DI, observamos una población, contamos el nº de nuevos casos de enfermedad en la población y medimos el tiempo neto, llamado persona-tiempo (PT).

$$DI = \text{nº de casos nuevos en un tiempo} / PT$$

PT = periodo de tiempo que cada persona permanece en observación

**Tabla 1. Diferencias entre Incidencia Acumulada, Densidad Incidencia y Prevalencia**

CARACTERISTICA	I. ACUMULADA	PREVALENCIA	D. INCIDENCIA
Que mide	Probabilidad de enfermar	% Población con enfermedad	Rapidez de aparición enfermedad
Unidad	Ninguna	Ninguna	Casos/persona-tiempo
Tiempo Diagnóstico Enfermedad	Diagnóstico nuevo	Existentes	Diagnóstico nuevo
Sinónimo	Riesgo de Incidencia	----	Tasa de Incidencia

## Ejemplo

Supongamos el estudio hipotético de 6 sujetos que se inició en 1982 y concluyó en 1991. El tiempo que cada sujeto permanece en estudio es distinto, pero todos estaban libres de la enfermedad al inicio del mismo. Todos los individuos fueron seguidos al menos durante 2 años.

- Paciente 1: Entró en 1982, fue diagnosticado de la enfermedad justo antes de 1984 y luego seguido hasta su muerte en 1989.  
Paciente 2: Se enroló en 1984, se siguió hasta 1986 sin desarrollar la enfermedad y luego dejó de participar en el estudio.  
Paciente 3: Entra en 1986 y se diagnosticó de la enfermedad justo antes de 1989, sobreviviendo hasta el final de la observación en 1991.  
Paciente 4: Entra en 1984 siguiéndose hasta el final sin desarrollar enfermedad.  
Paciente 5: Entra en 1989 y sigue hasta el final sin enfermar.  
Paciente 6: Entra en 1985 y sigue hasta el final sin enfermar.

La Incidencia Acumulada para el estudio será:

$$IA = 2/6 = 0,33 = 33\%$$
$$N^{\circ} \text{ casos nuevos} = 2 \text{ (pacientes 1 y 3)}$$
$$\text{Población estudiada} = 6$$

La Densidad de Incidencia será:

Para el cálculo de las personas-tiempo, el paciente 1 desarrolló la enfermedad 2 años después de entrar en el estudio. Dado que los sujetos contribuyen solamente persona-tiempo mientras que son susceptible para la enfermedad, la PT para el paciente 1 fue 2 años. Similarmente los pacientes 2, 3, 4, 5 y 6 contribuyeron con: 2,3,7,2 y 6 años respectivamente. Los pacientes 1 y 3 desarrollaron la enfermedad, por tanto:

$$PT = 2+2+3+7+2+6 = 22 \text{ personas-año}$$
$$DI = 2/22 = 0.09 \text{ casos/personas-año}$$

Esta tasa puede ser expresada como 9 casos/100pers-año al multiplicar por 100.

## 2.3 Relaciones entre Incidencia y Prevalencia

La Prevalencia depende de la Incidencia y de la Duración media de la enfermedad desde su comienzo hasta su resolución

Por ejemplo, la Prevalencia será pequeña cuando:

- Incidencia pequeña
- Alta letalidad
- Curación de la enfermedad alta

## 2.4 Supervivencia

Es la probabilidad de sobrevivir durante un período de tiempo específico.

Se estima  $TS = \text{Diagnosticados} - \text{Fallecidos} / \text{Diagnosticados}$

## 2.5 Letalidad

La tendencia de una enfermedad a producir la muerte de los pacientes afectados se conoce como casos fatales o letalidad.

La tasa de letalidad (TL) se estima mediante:  $TL = N^{\circ} \text{ Muertos} / N^{\circ} \text{ pacientes diagnosticados}$

Se puede considerar una proporción o multiplicar por 100 para convertirse en %.

### 3. Inferencia Causal

La inferencia causal consiste en extraer conclusiones a partir de los resultados de un estudio sobre la relación entre una posible causa y una enfermedad, aplicándolas posteriormente a poblaciones de mayor tamaño.

#### 3.1 Variable Epidemiológica

Característica/condición susceptible de ser medida, usando alguna escala de medición conocida y que puede adoptar diversos valores a los ojos del observador. Propiedad no constante

Podemos clasificar las variables en base a su:

##### 3.1.1 Naturaleza

- \* V. Cualitativa
  - \* V. Cuantitativa
- 

Discreta  
Continua

##### 3.1.2 Escala de medida

Criterios que utilizamos para definir las categorías en las que agrupamos las observaciones

###### *Nivel Nominal*

Categorías mutuamente excluyentes y colectivamente exhaustivas  
Sólo dos categorías: Variable dicotómica o binaria

###### *Nivel Ordinal*

Implícita una relación de orden

###### *Nivel Interval*

Establecemos intervalos o escalones

###### *Nivel de Razón*

Escala numérica continua

#### 3.2 Inter-relación de las variables

- a. V. Dependiente / V. Independiente
- b. V. Precedente / V. Subsiguiente
- c. V. Reversible / V. Irreversible
- d. Relación Determinante / Probabilística

En el modelo causal en el que  $A \longrightarrow B$  podemos decir que:

Variable A (Exposición) = V. Independiente, Precedente, Irreversible

Variable B (Efecto, enfermedad) = V. Dependiente, Subsiguiente, Reversible

### 3.2 Estimación de los Riesgos

Riesgo: Probabilidad de que suceda un determinado evento

*Factor de Riesgo:* Conjunto de fenómenos de los que depende la probabilidad de que suceda un evento

*Marcador de Riesgo:* Factor de riesgo no modificable

*Indicador de Riesgo:* Hecho que traduce la presencia de la enfermedad

*Signo de Riesgo:* Cualquier factor que está ligado estadísticamente con la enfermedad

#### 3.2.1 Medidas de Asociación

##### Riesgo Relativo

Estima cuantas veces tiene más probabilidades de enfermar las personas expuestas que las no expuestas

**Tabla 2. Tabla de 2X2 con la distribución según presencia/ausencia de enfermedad y exposición**

		<b>ENFERMEDAD</b>		
		+	-	
<b>EXPOSICIÓN</b>	+	<i>a</i>	<i>b</i>	<i>a+b</i>
	-	<i>c</i>	<i>d</i>	<i>c+d</i>
		<i>a+c</i>	<i>b+c</i>	<i>N</i>

$I_e = a / a+b$  Incidencia en Expuestos  
 $I_o = c / c+d$  Incidencia en No Expuestos  
 $I_t = a+c / N$  Incidencia Total  
 $P_o = a+b / N$  Proporción de expuestos  
 $P_e = c+d / N$  Proporción de No expuestos

Riesgo Relativo (Rr) =  $I_e / I_o$

$> 1$  Factor de Riesgo (FR)  
 $= 1$   
 $< 1$  Factor de Protección (FP)

Límites confianza =  $\text{Exp} \left\{ \ln (Rr) \pm 1.96 \sqrt{1/a + 1/b + 1/c + 1/d} \right\}$

**Medida de Asociación exclusiva de los estudios epidemiológicos en los que por diseño podemos conocer la Incidencia de la enfermedad**

##### Odds Ratio

Siguiendo con la Tabla 2 ahora conocemos:

$a/c$  Distribución (ratio) de la exposición entre los casos

$b/d$  Distribución (ratio) de la exposición entre los controles

$OR = (a/c) / (b/d) = a*d / b*c$

$>1$  FR  
 $= 1$   
 $< 1$  FP

Límites de confianza :  $\text{Exp} \left\{ \ln (OR) \pm 1.96 \sqrt{1/a + 1/b + 1/c + 1/d} \right\}$   
 Idéntica interpretación que el Riesgo Relativo

**Medida de Asociación de los estudios epidemiológicos en los que por diseño NO podemos conocer la incidencia de la enfermedad**

### 3.2.2 Medidas de Impacto

#### Riesgo Atribuible (Ra)

Contribución del FR a la enfermedad en el grupo de los expuestos:  $Ra = I_e - I_o$

#### Riesgo Atribuible Porcentual (Ra%)

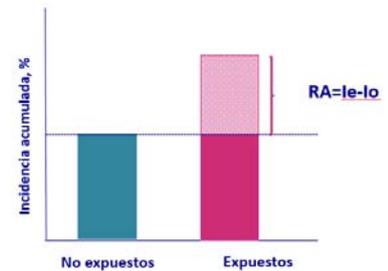
Porcentaje de la enfermedad que se debe al FR en el grupo de los expuestos

$$Ra\% = I_e - I_o / I_e * 100$$

$$Ra\% = Rr - 1 / Rr * 100$$

$$Ra\% = OR - 1 / OR * 100$$

La ventaja de la segunda fórmula, con la medida de asociación, es que se puede usar en estudios sin incidencia (Ej. C/C)



Cuándo es un FR será la **Fracción Etiológica de Riesgo**

Cuando es un FP será la **Fracción Prevenible =  $(1-Rr) * 100$**

#### Riesgo Atribuible a la Población (RAP)

Contribución del FR a la enfermedad en la población total:  $RAP = I_t - I_o$

#### Porcentaje de Riesgo Atribuible a la Población (RAP%)

En cuanto se beneficiaría la población si quitáramos el FR:  $RAP\% = I_t - I_o / I_t * 100$

#### Índice de Intervención / Número Mínimo que es Necesario Tratar

Es el inverso del Ra ( $1/Ra$ ). Número de expuestos en los que debería cesar la exposición para evitar un caso.

Número de pacientes que hay que tratar para evitar un caso

### 3.2.3 Asociación y Causalidad

Uno de los objetivos de la Epidemiología es conocer los factores determinantes de la enfermedad, para lo cual trata de establecer asociaciones causales entre distintas variables, para ello hemos de tener presente conceptos importantes:

**Causa:** Suceso o condición que inicia o permite, sola o en conjunto con otras causas, una secuencia de sucesos que resulta en un efecto

**Unicausalidad:** Una sola causa origina la enfermedad. Extremadamente ocasional y generalmente de las enfermedades infecciosas.

**Multicausalidad:** Varias causas originan la enfermedad

**Multiefectividad:** Una misma causa puede influir en varias enfermedades

La Asociación entre dos variables puede ser:

1. **Asociación no causal:** No hay una relación biológica lógica entre ambas variables, aunque si una relación estadísticamente significativa
2. **Causal directa:** La V. exposición está directamente asociada con la enfermedad
3. **Causal indirecta:** La V. exposición no es la causa última de la enfermedad
4. **Asociación espuria:** La variable exposición actúa de factor de confusión

#### Tipos de causas

Causa Suficiente: inevitablemente origina consecuencias

Causa Contribuyente: ella sola no es suficiente, pero contribuye a la aparición de la enfermedad

Causa Necesaria: debe estar presente para el desarrollo de la enfermedad

**Criterios de causalidad de Bradford Hill**

1. Que la asociación sea fuerte
2. Que la exposición preceda a la enfermedad (Temporalidad)
3. Que exista una relación dosis/respuesta
4. Que la asociación sea consistente
5. Que exista plausibilidad biológica

## 4. Clasificación de los Estudios Epidemiológicos

### Manipulación

- |                           |        |                        |
|---------------------------|--------|------------------------|
| 1. Existe manipulación    | —————> | <b>Experimentales</b>  |
| 2. No existe manipulación | —————> | <b>Observacionales</b> |

### Unidad de observación/análisis

1. Individual
2. Poblacional

### Seguimiento

- |                          |        |                                                             |
|--------------------------|--------|-------------------------------------------------------------|
| 1. Existe seguimiento    | —————> | <b>Longitudinales</b> (Cohortes, Experimentales)            |
| 2. No existe seguimiento | —————> | <b>No longitudinales</b> (Transversales, Casos y controles) |

### Sentido del análisis

1. Medimos la exposición antes de la enfermedad: **Anterógrados** (Cohortes)
2. Medimos la exposición después de la enfermedad: **Retrógrados** (Transversales, Casos y controles)

### Inicio del estudio

- |                    |        |                                                                                   |
|--------------------|--------|-----------------------------------------------------------------------------------|
| 1. Efecto ausente  | —————> | <b>Prospectivos</b> (Cohortes prospectivas, Experimentales)                       |
| 2. Efecto presente | —————> | <b>Retrospectivos</b> (Cohortes retrospectivas, Transversales, Casos y Controles) |

### 4.1 Epidemiología Descriptiva

La epidemiología descriptiva tiene como objetivo "DESCRIBIR" la caracterización y distribución de las enfermedades en un grupo poblacional, en un tiempo dado. Es decir, se limita a CONTAR.

La investigación descriptiva de una enfermedad está basada en 3 preguntas: ¿quién?, ¿dónde?, ¿cuándo? ocurre o aparece el evento, es decir va a responder a las 3 variables que caracterizan la distribución de dicho evento: PERSONA, LUGAR, TIEMPO.

#### Aplicaciones

1. Conocer la frecuencia de un evento, numerador de las medidas de frecuencia que ya conocemos.
2. Explicar el comportamiento de un evento en el área estudiada.
3. Describir la historia natural de la enfermedad.
4. Formular hipótesis o corroborarlas.
5. Permite a la Administración Sanitaria conocer las demandas y planificar.
6. Permite establecer bases para la investigación clínica, terapéutica y preventiva.

Son los estudios más sencillos, están muy limitados y por tanto tienen muchos problemas y son inexactos, pero a pesar de todo son el primer paso en la investigación epidemiológica.

#### Variables de Persona

Son rasgos, cualidades, propiedades de la persona.

Las principales variables de persona son: edad, sexo, etnia, estado civil, nivel socioeconómico, nivel educacional, historia familiar, profesión

#### Variables de Lugar

Variación de la frecuencia de la enfermedad en cuanto a lugar.

Podemos distinguir entre:

- a. **Lugar Epidemiológico**
- b. **Lugar Ecológico**
- c. **Lugar Geopolítico o administrativo**

Podemos comparar las tasas de una enfermedad en distintos lugares, pero **¡OJO!** deberemos de ajustar.

Al observar las diferencias podemos adivinar la influencia de factores ambientales, costumbres y factores genéticos, como es el estudio clásico de los inmigrantes japoneses

### **Variable Tiempo**

Podemos preguntarnos ¿Cuándo aparecen los aumentos o disminuciones de la enfermedad?, ¿Se presentan con la misma frecuencia esta enfermedad ahora que en el pasado?

Al observar cómo se presentan los casos veremos si existen aglutinaciones en el tiempo "Cluster", en el lugar o en ambas variables.

Si estudiamos la enfermedad en un punto (tiempo): E. Transversal (fotografía de la enfermedad, Prevalencia).

Si por el contrario realizamos estudios en puntos sucesivos: Series temporales (Película.)

La variable tiempo nos permite establecer las **TENDENCIAS**.

### **Relaciones Tiempo-Lugar**

1. Distinto tiempo y el mismo lugar  
En una comunidad se producen variaciones en distintos tiempos, por:
  - Modificaciones en nacimientos y defunciones
  - Modificaciones en criterios diagnósticos, terapéuticos y preventivos (vacunas)
2. Mismo tiempo y distinto lugar  
¡OJO! al comparar tasas brutas.

### **Tipos de Estudios Descriptivos**

1. Caso Índice y Serie de Casos
2. Estudios Transversales
3. Estudios de Series Temporales (Vigilancia)
4. Estudios Ecológicos

### **Formulación de Hipótesis a partir de estudios descriptivos**

Hemos comentado que estos estudios son el primer paso para la investigación de cualquier evento y nos pueden permitir formular hipótesis para seguir investigando. La formulación de estas hipótesis ha de ser siempre **razonable** y su metodología deriva de los 5 items de J. Mill Stuart (1862) que resumimos en 3:

1. Método de las diferencias
2. Método de las concordancias
3. Método de las variaciones concomitantes

## 4.2 Epidemiología Analítica

A diferencia de la descriptiva los estudios analíticos nos permiten establecer relaciones causales entre exposiciones y enfermedades, es decir realizar un análisis mediante la comparación de dos grupos de sujetos.

Se realiza así la segunda función de la Epidemiología, COMPARAR.

### *Estudios de Cohorte*

- \* **Cohorte:** grupo de sujetos con una característica común
- \* Son estudios **Observacionales**
- \* Se comparan dos cohortes
  - Expuestos
  - No expuestos
- \* Cuando se inician todos los sujetos están sanos (t=0 población sana)  
Vamos de Causa → Efecto → Anterógrado
- \* Se realiza SEGUIMIENTO de los sujetos  
Detectamos “casos nuevos” → Incidencia

### **Objetivos**

Describir la enfermedad y medir su frecuencia durante el seguimiento (Contar)

Analizar las posibles relaciones entre exposición (FR) y enfermedad (Comparar)

### **Tipos**

- Cohortes Prospectivas
- Cohortes Retrospectivas

### **Cohortes Prospectivas**

#### **Van del presente hacia el futuro**

#### **Pasos**

1. Seleccionar dos cohortes de la población
2. A t=0 descartar enfermedad y medir exposición para determinar cohorte de expuestos y cohorte de no expuestos
3. Realizar seguimiento
  - Detectar aparición enfermedad (Incidencia)
  - Valorar los posibles cambios en la exposición

#### **Medidas Epidemiológicas**

Frecuencia: Incidencia

Asociación: Riesgo relativo

Impacto: Riesgo atribuible porcentual y riesgo atribuible a la población porcentual

#### **Ventajas**

1. Podemos describir las dos cohortes
2. Conocemos la **Historia Natural** de la enfermedad. Describir enfermedad
3. Vamos de CAUSA-EFECTO. Aseguran el principio de causalidad Tiempo-Respuesta
4. Inferencia fuerte
5. Requieren una planificación detallada por ello menor probabilidad de error aleatorio y SESGOS o errores sistemáticos
6. Permiten conocer multiefectividad

#### **Inconvenientes**

1. Su principal problema es las PERDIDAS y ABANDONOS y los CAMBIOS EN LA EXPOSICIÓN durante el seguimiento
2. Son estudios de larga duración.
3. Necesitan personal durante todo el proceso. CAROS

4. No aptos para enfermedades poco frecuentes, “raras”

Ejemplo: Estudio de Framingham

## Cohortes Retrospectivas

### Van del Pasado al Presente

Son semejantes a las anteriores, pero:

- 1.- El efecto ya está presente al iniciar el estudio (presente)
- 2.- La exposición se midió antes de aparecer el efecto (pasado)

Sigue siendo un estudio **anterógrado**

### Pasos

- 1.- Identificamos una cohorte reunida en el pasado
- 2.- Datos de la exposición en el pasado, cuando todos estaban sanos
- 3.- Recogemos los datos referentes al efecto (en el pasado o presente)

### Medidas Epidemiológicas

Frecuencia: Incidencia

Asociación: Riesgo relativo

Impacto: Riesgo atribuible porcentual y riesgo atribuible a la población porcentual

### Ventajas

1. La exposición (FR) precede a la enfermedad: ANTERÓGRADO
2. Inferencia causal sigue siendo fuerte
3. Son más cortos.
4. Más económicos que los de cohortes prospectivas

### Inconvenientes

1. No tenemos seguridad absoluta sobre la medida de exposición
  - . puede ser inexacta
  - . puede estar incompleta
2. Pueden existir cambios en los criterios diagnósticos
3. La posibilidad de cometer sesgos es mayor que en las prospectivas

Ej: Estudios de Epidemiología ocupacional.

## Estudios No Longitudinales

Se caracterizan por ser estudios

- Observacionales Analíticos
- Retrospectivos → Efecto presente al inicio del estudio
- Retrógrados → Según sentido del análisis vamos de Efecto → Causa
- No hay seguimiento

## Estudios Transversales

- \* También llamados de Prevalencia o de sección transversa
- \* Medimos al mismo tiempo exposición y enfermedad ( $t=0$ )
- \* La frecuencia de la enfermedad vendrá dada por su Prevalencia
- \* No hay seguimiento
- \* Describen la distribución de la enfermedad y la de las exposiciones

## Medidas Epidemiológicas

Frecuencia: Prevalencia  
Asociación: Odds Ratio  
Impacto: Riesgo atribuible porcentual

### Ventajas

1. Rápidos y económicos
2. No hay seguimiento, no hay pérdidas
3. Útiles para conocer entramados causales

### Inconvenientes

1. No para enfermedades poco frecuentes
2. No proporcionan  $H^a$  natural de la enfermedad
3. Al medir causa y efecto al mismo tiempo puede ser difícil establecer quien es quien

## Estudios de Casos y Controles

Seleccionar: un grupo de personas Enfermas: **Casos** y  
un grupo de personas No Enfermas: **Controles**

Medir las Exposiciones en ambos grupos y tratar de determinar la posible asociación con la Enfermedad

### Selección de los Casos

- A. Definición concreta de "caso"
- B. Podemos tomarlos de:
  1. Hospital
  2. Consultas Externas
  3. Población General de un área geográfica concreta
  4. Casos detectados mediante sistemas de vigilancia
- C. Podemos utilizar:
  - Casos Incidentes
  - Casos Prevalentes

### Selección de los Controles

1. Criterio Principal: "No estar enfermo y tener, o haber tenido, la misma probabilidad de exposición que los casos"
2. Proceder de la misma población diana que los casos y ser HOMOGENEOS a ellos
3. Ideal que la selección fuera aleatorizada

### Medidas Epidemiológicas

Frecuencia: Ninguna  
Asociación: Odds Ratio  
Impacto: Riesgo atribuible porcentual

### Ventajas

1. Útiles para estudiar enfermedades poco prevalentes o con periodo de latencia largo
2. Permiten generar nuevas hipótesis
3. Relativamente económicos y corta duración
4. Conocemos multicausalidad de la enfermedad

### Inconvenientes

1. Dificultad en la obtención de los Casos / Controles homogéneos
2. Dificultad en asegurar relación tiempo / respuesta
3. No permiten estudiar Multiefectividad
4. No conocemos ni Incidencia ni Prevalencia
5. Muy susceptibles a sesgos de todo tipo:
6. No permiten conocer Historia Natural de la enfermedad

Ej. Talidomida y síndrome focomélico. Intoxicación por consumo de aceite de colza desnaturalizado

## Casos y Controles Equiparados

- \* Se basan en seleccionar los casos y controles en función de la variable que queremos controlar y que podría plantear problemas en la investigación por comportarse como FC
- \* Variables a controlar más habitualmente: Edad, Etnia, Sexo, Nivel socioeconómico, ...
- \* Por cada "caso" con la variable a equiparar tomaremos 1 / 2 "controles" con la misma variable
- \* Podemos eliminar influencias del FC equiparado
- \* No permiten estudiar si la variable por la que equiparamos influye en la enfermedad
- \* Requiere un análisis específico, analizamos la razón de pares discordantes

## Casos y Controles "Anidados" en una Cohorte

Estudios **Retrospectivos** que están "a caballo" entre Cohortes y Casos/Controles

### Pasos

- 1.- Identificar una cohorte de reunida en el pasado. Obtener permiso
- 2.- A  $t=0$  se midió "exposición", normalmente una muestra y se conservan
- 3.- Definir "Caso" e identificar en la cohorte
- 4.- Seleccionar una muestra de "no casos" de los restantes
- 5.- Localizar las muestras tomadas a  $t=0$  de los casos y no casos
- 6.- Análisis

### Medidas Epidemiológicas

Frecuencia: Incidencia Total  
Asociación: Odds ratio  
Impacto: Riesgo atribuible porcentual

### Ventajas

1. Muy útil si se han de hacer determinaciones costosas
2. Conserva el carácter anterógrado de los estudios de cohortes.

### Inconvenientes

1. Sólo para variables predictivas "raras" que se puedan conservar
2. Se pueden deteriorar o perder durante el seguimiento

3. Pueden no estar recogidas correctamente

Ejemplo: estudio sobre la relación entre carotenos y cáncer

### 4.3 Epidemiología Experimental. Estudios Experimentales

#### **Concepto**

Proceder metodológico en el que un grupo de individuos ha sido dividido de forma **aleatoria**, en dos grupos: grupo control y grupo experimental o de estudio, analizándose con respecto a alguna variable o característica que se pretende evaluar o estudiar.

#### **Aplicaciones**

1. Evaluar eficacia de nuevos procedimientos: Terapéuticos, Preventivos, Diagnósticos
2. Revisar procedimientos tradicionales
3. Evaluación y control de programas de Salud Comunitaria y ayuda a la toma de decisiones administrativas en Salud.
4. Establecer relaciones de causalidad de una enfermedad. Es el diseño de estudio epidemiológico más potente para medir la relación entre dos variables

#### **Tipos de Estudios Experimentales**

**Ensayo Clínico** Se trata de una experiencia en la que un tratamiento u otra medida se administra a sujetos individuales sanos o enfermos con intenciones diagnósticas, terapéuticas o de prevención primaria. Por tanto, vamos a **evaluar la eficacia** en **sujetos "individuales"**. Nuestra unidad de intervención y análisis es el individuo.

**Ensayo en la Comunidad** Se emplea un "**grupo**" de individuos como "**un todo**" para determinar o evaluar la eficacia del proceder a investigar. Por tanto, nuestra unidad de intervención son comunidades, grupos o poblaciones considerados de forma completa.

En este caso hablamos de una Comunidad Control y una Comunidad de intervención o experimental. Ejemplo, suplemento de flúor en el agua de bebida.

#### **Ensayo Clínico**

Podemos definir al Ensayo Clínico como: "Una experiencia científica en la que un tratamiento u otra medida se administra a sujetos individuales, sanos o enfermos con intención diagnóstica, terapéutica o de prevención, así como conocer su farmacodinamia, farmacocinética y reacciones adversas".

#### **Desarrollo de un nuevo Fármaco**

El desarrollo y puesta en marcha de un nuevo medicamento es una empresa compleja y de larga duración, conlleva varias etapas que deben demostrar una relación causal entre el tratamiento y la curación (o la mejora del estado de salud)

Lógicamente la investigación en humanos no se iniciará en ninguna circunstancia hasta que la experimentación animal demuestre su inocuidad para el hombre.

En el desarrollo de un nuevo fármaco se pueden distinguir dos fases principales: Investigación preclínica e investigación clínica. La investigación clínica se divide en Preautorización y Post-autorización.

La investigación clínica pre-autorización comprende los ensayos clínicos siguientes:

**FASE I:** En ella se estudia la Farmacología clínica y la toxicidad.

Objetivo: Evaluar seguridad del fármaco en humanos. Farmacocinética, farmacodinamia

Sujetos: Voluntarios sanos

Diseño: No controlados

Constituye el primer paso en la investigación de una sustancia o medicamento nuevo en el hombre, y va dirigido a conocer la "seguridad" de la droga y no su eficacia.

**FASE II:** Se efectúa investigación clínica inicial del efecto terapéutico.

Objetivo: Evaluar eficacia/seguridad, indicaciones, ajustar dosis

Diseño: Controlados, aleatorizados, doble ciego



## Consideraciones Éticas

Dado que estamos investigando con nuevas drogas o procedimientos en seres humanos se plantean diferentes problemas éticos, motivo por el que este tipo de estudios se encuentran sometidos a unas reglas de ética médica.

Los ensayos clínicos y en general los estudios experimentales en humanos, han de efectuarse en condiciones de máximo respeto a los derechos fundamentales de la persona, siguiendo las Normas de Buena Práctica Clínica, así como los postulados éticos contenidos en la **Declaración de Helsinki**; todo ello se encuentra contemplado en el título 3º de la ley del Medicamento.

Es obvio que no se pueda iniciar ningún ensayo clínico en tanto no se disponga de la suficiente información en animales de experimentación: ensayos farmacológicos y toxicológicos que garanticen "los riesgos mínimos admisibles" para el ser humano.

El Código de Helsinki, elaborado en 1964 en el seno de la Asociación Médica Mundial, teniendo como antecedente el Código de Núremberg (1947), considerado como la primera declaración de la comunidad científica Internacional, recoge los postulados básicos de la investigación clínica.

La Declaración de Helsinki de 1964, se ha revisado en distintos momentos, siendo la más reciente la del año 2002.

Podemos decir que la realización de un estudio experimental, y en general de cualquier investigación en seres humanos, debe tener siempre presente cuatro **principios éticos**:

**Principio de no Maleficencia**, ha de ser mayor el beneficio que el riesgo, siendo éste mínimo y contar con un diseño correcto

**Principio de Justicia**, seleccionar equitativamente la muestra, siendo representativa de la población diana a la que pretendemos inferir nuestros resultados, compensación de los daños y seguros de riesgo (recogido en nuestra legislación).

**Principio de Autonomía**, consentimiento informado, mantenimiento de la intimidad y la confidencialidad.

**Principio de Beneficencia**, teniendo presente las decisiones de sustitución, y la protección de los grupos vulnerables (niños, ancianos, embarazadas).

## 5. Sesgos. Concepto. Tipos. Su Problemática y Control

Los estudios epidemiológicos, como el resto de los de investigación, pueden ser considerados un ejercicio de medición. Por tanto, lo que se pretende es que dichas medidas no sean erróneas.

Por ello a la hora de explicar una asociación, que al fin y al cabo nos indica una diferencia entre los sujetos expuestos y no expuestos, nos hemos de preguntar si ésta se podría explicar por otras razones tales como:

- \* Que sea consecuencia del azar: VARIABILIDAD ALEATORIA
- \* Que sea consecuencia de diferencias sistemáticas: SESGOS

Por otra parte, habitualmente no se estudian todos los sujetos sino una fracción de los mismos, y es a partir de aquí desde donde realizamos la inferencia.

Estas fracciones poblacionales pueden ser extraídas como: MUESTRA o POBLACIÓN

### TIPOS DE POBLACIÓN

Población de estudio	Sujetos que estudiamos
Población real o accesible	Sujetos a los que tendríamos acceso
Población diana	Sujetos en los que se aplicarán las conclusiones
Población externa, (Universo)	

### VALIDEZ Y PRECISIÓN. TIPOS DE ERROR

Hay que diferenciar convenientemente los dos grandes tipos de error que se pueden dar en un estudio los cuales van a estar relacionados con la Validez y la Precisión del mismo.

#### 1.- PRECISIÓN. ERROR ALEATORIO.

Debido a variaciones fruto del azar que se denominan **Variación aleatoria**. Puede modificar los resultados en cualquier sentido. Se denomina **PRECISIÓN** o reproducibilidad de la medida al grado de semejanza entre distintos resultados obtenidos en idénticas circunstancias.

Tipos de error aleatorio:	1.- Error de muestreo
	2.- Error en la medida
	3.- Errores $\alpha$ y $\beta$

#### 2.- VALIDEZ. ERROR SISTEMÁTICO. SESGOS.

El resultado que obtenemos se aparta siempre en el mismo sentido, **sistemáticamente**, del valor verdadero. Se denomina **VALIDEZ** a la capacidad que tiene un estudio para reflejar lo que realmente sucede, mide lo que intenta medir.

Los conceptos de Validez y Precisión están muy relacionados, pero no necesariamente discurren paralelos, ya que una medida válida no tiene por qué ser precisa y viceversa.

Un error sistemático o sesgo es el resultado de: **un proceso, en cualquier etapa de la inferencia que tiende a producir resultados que se apartan sistemáticamente del valor verdadero.**

Se define sesgo relativo como:

$$\text{SESGO (OR}_1, \text{OR)} = (\text{OR}_0 - \text{OR})/\text{OR}$$

Esta ecuación nos va a resultar muy útil puesto que nos permite conocer:

- magnitud (cuantificar)
- dirección

$\text{OR}_0 > \text{OR}$  efecto más fuerte que el real, POSITIVO, SOBRESTIMACIÓN, DIFERENTE DEL NULO

$\text{OR}_0 < \text{OR}$  efecto más débil que el real, NEGATIVO, SUBESTIMACIÓN, HACIA EL NULO

### CLASIFICACIÓN DE LOS SESGOS

- SESGO DE SELECCIÓN
- SESGO DE INFORMACIÓN
- SESGO DE CONFUSIÓN, FACTORES DE CONFUSIÓN

### 1.- SESGO DE SELECCIÓN

La distorsión del efecto verdadero se produce como consecuencia del proceso de selección de los sujetos.

#### Causas del Sesgo de Selección

- \* Elección del marco de muestreo: Sesgo de Berckson o Paradoja de Berckson
- \* Pérdida del seguimiento
- \* Vigilancia diagnóstica desigual en los estudios de casos y controles
- \* Voluntarios

### 2.- SESGO DE INFORMACIÓN

Se produce porque existe una diferencia sistemática al obtener la información de los sujetos o una clasificación incorrecta de los mismos en una o más variables por lo que otros sinónimos son sesgo de OBSERVACIÓN, MEDICIÓN y MALACLASIFICACIÓN.

#### Fuentes de Sesgo de Malaclasificación

- \* Sesgo del entrevistador
- \* Procedimiento para diagnosticar la enfermedad inexacto, criterios diagnósticos incorrectos o diferentes para los distintos grupos. Estado de enfermedad mal clasificado.
- \* Procedimiento para determinar la exposición inexacto o incorrectos o diferentes para los distintos grupos. Estado de exposición mal clasificado.
- \* Recuerdo selectivo
- \* Vigilancia diagnóstica desigual en los estudios longitudinales.

### 3.- FACTORES DE CONFUSIÓN

Este error sistemático se produce cuando existe un tercer factor extraño que puede sobre o subestimar la asociación dependiendo de su relación con exposición y enfermedad, pudiendo incluso modificar la asociación.



Es el único el que "a priori" se puede corregir en la fase de análisis.

Las condiciones que debe cumplir un factor extraño para ser un factor de confusión serán:

- a.- **Factor de riesgo para la enfermedad entre los no expuestos.**
- b.- **Asociada a la exposición en la población.**
- c.- **No ser un paso intermedio en la relación causal.**

#### MÉTODOS PARA EL CONTROL DE LOS FACTORES DE CONFUSIÓN

1. Aleatorización
2. Equiparación/Matching
3. Estratificación
4. Restricción
5. Análisis de regresión

## 6. INTERPRETACION DE LOS RESULTADOS DIAGNOSTICOS

### Sensibilidad, Especificidad, Valores Predictivos

1. Puesto que la variabilidad es una ley de la biología, la probabilidad va a ser una constante de nuestro quehacer diario. Por tanto, nunca podremos decir si A siempre B, tendremos que decir, si A *probable y posiblemente* B.
2. Si tenemos en cuenta que la probabilidad de estar enfermo de un determinado paciente se puede ir desplazando en una escala que va de 0 a 1 a medida que tenemos más información, un test será todo aquello que me permita desplazar la probabilidad.
3. Sabemos que una prueba diagnóstica concreta nos permite clasificar a la población en sanos y enfermos, mientras que un test usado en screening, cribado, nos permite clasificar a la población en alto y bajo riesgo.

#### VALIDACIÓN DE UN TEST

		ENFERMEDAD			
		+	-		
RESULTADO DEL TEST	+	<i>a</i>	<i>b</i>	<i>a+b</i>	<i>a</i> = verdaderos +
	-	<i>c</i>	<i>d</i>	<i>c+d</i>	<i>b</i> = falsos +      FP
		<i>a+c</i>	<i>b+c</i>	<i>N</i>	<i>c</i> = falsos -      FN
					<i>d</i> = verdaderos -

#### SENSIBILIDAD – ESPECIFICIDAD

Tenemos que conocer la capacidad del test para CLASIFICAR correctamente a la población.

Enfermos → correctamente → SENSIBILIDAD  
No Enfermos → correctamente → ESPECIFICIDAD

#### SENSIBILIDAD

Es la "capacidad para detectar a los **casos**",  $S = a/a+c$ . Luego podemos definir S como: Probabilidad de que estando enfermo el test de positivo.

$S + \text{tasa de FN} = 1$ . Por tanto, son Complementarios.

#### ESPECIFICIDAD

"Capacidad para detectar correctamente a los **no casos**",  $E = d/b+d$ . Luego se puede definir E como: Probabilidad de que estando sano tenga el test negativo.

$E + \text{FP} = 1$ . Son complementarios

No existe ningún test con S y E del 100%, aunque así se diga en algunos artículos.

#### CALCULO DE S y E DE UN NUEVO TEST

- \* "ESTANDAR DE ORO" y comparar.
- \* El límite entre positivo y negativo del nuevo test lo pone el investigador (CUTOFF), dependiendo de dónde lo sitúe variarán la S y E del test.
- \* Para elegir la S y E más adecuadas se usan las denominadas curvas ROC (COR = Características Operativas del Receptor)

#### PROBABILIDADES POST-TEST -- VALORES PREDICTIVOS

- \* Probabilidades post-test, + ó -, de estar enfermo o sano respectivamente y sus complementarios.

### VALOR PREDICTIVO POSITIVO

"Probabilidad post-test positivo de estar enfermo", "Probabilidad de que teniendo el test (+) tenga la enfermedad".

$VP^+ = a/a+b$ ;  $VP^+ + \text{tasa de FP} = 1$

tasa de FP = Probabilidad post-test positivo de estar sano

### VALOR PREDICTIVO NEGATIVO

"Probabilidad post-test negativo de no estar enfermo (estar sano)", "Probabilidad de que teniendo el test (-) no tenga la enfermedad".  $VP^- = d/c+d$ ;  $VP^- + \text{tasa de FN} = 1$ .

tasa de FN = Probabilidad post-test negativo de estar enfermo

### Variación de las probabilidades post-test y prevalencia

Cuando la probabilidad pretest, la prevalencia (P), varía se modifican las probabilidades post-test

### Razón de Probabilidad

Hemos visto como la probabilidad post-test se encuentra influenciado por la probabilidad pre-test.

Existe una forma de marginar la variación de los VPs calculando las razones de probabilidad. Se trata de calcular un odds

$$RP^+ = \frac{\text{Probabilidad de que teniendo la enfermedad tenga el test (+)}}{\text{Probabilidad de no teniendo la enfermedad tenga el test (+)}} = S / 1-E$$

Es decir Aciertos test + / Fallos test +

$$RP^- = \frac{\text{Probabilidad de que teniendo la enfermedad tenga el test (-)}}{\text{Probabilidad de no teniendo la enfermedad tenga el test (-)}} = 1-S / E$$

Es decir Fallos test - / Aciertos test -

Hemos marginado la PREVALENCIA

### REQUISITOS PARA UTILIZAR UN TEST EN UN PROGRAMA DE SCREENING

- Validez de la prueba (S y E)
- Criterios de predicción de la enfermedad (Valores Predictivos)
- Reproducibilidad (Indice Kappa)
- Aceptado por la población
- Sencillo y fácil de realizar
- Económico

### BIBLIOGRAFÍA

4. **Argimon J, Jiménez J.** Metodos de Investigacion Clinica y Epidemiologica (4ª Ed.) Elsevier. Barcelona 2013. Acceso electrónico: <https://www.elsevierlibrary.es/pdfreader/mtodos-de-investigacin-clnica-y-epidemiologica52638>
5. **Gordis L.** Epidemiología. 5ª Edición. Elsevier. Madrid 2014. Acceso electrónico: <https://www.elsevierlibrary.es/pdfreader/epidemiologa>
6. **Greenberg RS.** Medical Epidemiology. 4ª ed. Lange. New York, 2013.
7. **Hulley S, Cummings S, Browner W, Grady D, Newman T.** Diseño de investigaciones clínicas (4ª Ed). Wolters Kluwer Lippincott Williams & Wilkins. Barcelona, 2014

8. **Martinez González MA.** Conceptos de Salud Pública y Estrategias Preventivas. 2ª Ed. Elsevier. Barcelona 2018.
9. **Piédrola Gil y cols.** Medicina Preventiva y Salud Pública. 12ª edición. Elsevier. Madrid 2015.

## **CAPÍTULO 5**

### **Muestreo y tamaño muestral**

**Autores:** Ana Martínez Santos  
David Martínez Hernández  
Juana María Santos

## Objetivos

Esta actividad pretende introducir al estudiante en el concepto de muestreo, sus tipos y cómo predeterminarlo. Se incluye el enlace a una hoja de cálculo que permite dicho cálculo.

## Desarrollo

El muestro, pretende determinar qué casos de la población elegible debemos extraer para que sea representativa; y por tanto permita inferir los resultados obtenidos en nuestro estudio a la población de referencia. Siempre que utilicemos muestras y no la población general, cometeremos un error al generalizar los resultados, es el error aleatorio; ya que solo vamos a analizar una parte de los pacientes. Pretendemos obtener un conjunto de datos que contenga una imagen fiable a escala de los datos de la población general.

## Tipos de muestreo

### Muestreo Probabilístico

Es una técnica de muestreo en la que se recogen las muestras de forma que todos los individuos de la población tienen la misma probabilidad de ser seleccionados.

Hay varios tipos de muestreo probabilístico:

- Muestreo aleatorio simple.
- Muestreo sistemático.
- Muestreo aleatorio estratificado.
- Muestreo por conglomerados (clusters)

### Muestreo aleatorio

Este tipo de muestreo es la forma más fácil de muestreo probabilístico. El investigador debe asegurarse de que todos los miembros de la población ( $N$ ) estén incluidos en la lista y posteriormente escoger al azar el número deseado de sujetos ( $n$ ). La probabilidad de estar incluido será:  $p=n/N$ .

Ventajas: Es un método fácil de realizar y comprender, que permite cálculos fáciles que se pueden realizar a través de paquetes informáticos.

Desventajas: Hay que disponer de una relación exhaustiva de los casos de la población, y si el tamaño ( $n$ ) es pequeño puede no ser representativa.

### Muestreo sistemático

Este tipo de muestreo lo podemos comparar con una progresión aritmética en donde la diferencia entre dos números consecutivos es la misma. Por ejemplo, supongamos que tenemos  $N$  pacientes diabéticos y queremos obtener una muestra con  $n$  casos. El intervalo de salto ( $k$ ) lo obtendremos:  $k = N / n$

Si en nuestro ejemplo  $N = 100$  y  $n = 20$  obtendremos  $k = 100/20 = 5$

Escogeremos ahora un número aleatorio (caso inicial) a partir del cual iremos dando los sucesivos saltos para seleccionar los casos de la muestra. Por ejemplo (3).

Con lo que nuestra muestra ( $n$ ) la formarán los pacientes: 3, 8, 13, 18, 23... y así hasta llegar a los 20 casos que necesitamos.

Ventajas: Es muy fácil de utilizar, y no siempre es imprescindible disponer de la relación de casos de la población de referencia.

Desventajas: En el caso de que la dimensión del intervalo de salto tenga relación con el fenómeno estudiado, podemos sufrir sesgos, que en algunos casos pueden ser importantes.

### **Muestreo aleatorio estratificado**

Aquí los sujetos se agrupan en diferentes categorías, (edad, sexo, nivel socioeconómico, etc.,) Para obtener nuestra muestra ( $n$ ) se seleccionan los sujetos de forma aleatoria de las distintas categorías. Las diferentes categorías no deben superponerse. Se utiliza cuando se desea estudiar un determinado subgrupo de la población.

Cuanto más homogéneos sean los estratos, mejor representados están los casos de población.

Ventajas: Asegurar que la muestra represente mejor a la población de estudio, en función de las variables por las que se estratifican.

Desventajas: Hay que saber la distribución de las variables por las que se quiera estratificar en la población de referencia. Puede requerir la ponderación de cada caso de la muestra (asignar pesos a cada elemento).

### **Muestreo por conglomerados**

Este muestreo lo realizamos si no podemos realizar un muestreo aleatorio simple debido al tamaño de la población. Por ejemplo, si queremos realizar un muestreo aleatorio simple cuando la población en cuestión es toda la población de África.

Aquí debemos seleccionar de forma aleatoria una serie de áreas (países), teniendo en cuenta que las diferentes áreas tengan la misma probabilidad de ser seleccionadas. Y podremos incluir todos los individuos de cada área seleccionada o bien, escoger de forma aleatoria a los sujetos de cada área.

Ventajas: Es muy eficiente cuando la población es muy grande y dispersa. Reduce costes. Exige disponer de listados sólo de las posibles unidades primarias de muestreo.

Desventajas: La variabilidad (error típico) es mayor que en otros tipos de muestreo (aleatorio simple o estratificado) y su cálculo complejo.

### **Cálculo del tamaño muestral**

Cada estudio tiene un tamaño muestral ( $n$ ) ideal para obtener una muestra representativa de la población de referencia, con una probabilidad de error conocido. Un tamaño insuficiente no es representativo del problema que queremos estudiar y un tamaño excesivo es ineficiente.

Cada estudio tiene un procedimiento y una fórmula de cálculo específica. La base de cálculo es el error típico<sup>1</sup>, en el que se basan los intervalos de confianza que utilizamos con los diferentes estadísticos que calculamos. A mayor tamaño muestral ( $n$ ), menor error; a menor tamaño muestral ( $n$ ), mayor error; la fórmula del error típico tiene  $n$  en el denominador, por ello a mayor  $n$ , menor error típico y viceversa. Si tomamos como muestra toda la población, el error típico desaparece (con toda la población  $n=\infty$ ).

### **Cálculo del tamaño muestral en un estudio**

En la fase inicial de cualquier estudio es imprescindible realizar la predeterminación del tamaño muestral. No realizarlo puede tener importantes consecuencias. Por un lado, un tamaño muestral menor del necesario nos dará NO significativo, cuando en realidad el fenómeno si lo es. En el caso contrario, estudiaríamos un número excesivo de casos, con la consiguiente pérdida de tiempo, coste innecesario y en muchas ocasiones, y en el mejor de los casos, molestias para los pacientes.

Para calcular el tamaño muestral de un estudio, debemos considerar dos posibles situaciones.

- a) Estudios para determinar parámetros. Pretendemos estimar un parámetro (suelen ser medias y proporciones<sup>2</sup>) en la población a partir de los resultados observados en una muestra.
- b) Estudios para contraste de hipótesis. Pretendemos realizar una comparación de medias o proporciones en las muestras para ver si son diferentes.

### **Estimación de parámetros**

En la primera situación (*a*) esta estimación de parámetros nos permitirá extrapolar (estimar, inferir) nuestros resultados en medias y proporciones, a partir de una muestra, a la población de referencia.

#### **a) Estimación de proporciones**

Si deseamos conocer el tamaño muestral para inferir el valor de una *proporción* de una variable a partir de una muestra a una población de referencia, debemos saber:

A: El nivel de confianza ( $1-\alpha$ ) que en nuestro caso suele ser del 95%. Una seguridad del 95 se corresponde a la región bajo la curva de una distribución normal comprendida entre la *media  $\pm 1,96 Dt$*  por ello el valor de  $Z$  será igual al 1.96. Una seguridad del 99% se corresponde a la región bajo la curva de una distribución normal comprendida entre la *media  $\pm 2,58 Dt$*  por ello el valor de  $Z$  en este caso será igual al 2,58.

B: La precisión, es decir el error máximo admisible pongamos un 2,5%. La amplitud del intervalo de confianza es dos veces el valor de la precisión.

---

<sup>1</sup> El término en castellano es *error típico*, en inglés es *error estándar*. Ambos términos son sinónimos.

<sup>2</sup> Es un tipo especial de razón en la cual los elementos del numerador están incluidos en el denominador. Se expresa en tanto por uno. Si se expresa en tanto por ciento se llama porcentaje.

C: Realizar una estimación del valor de la proporción que deberíamos encontrar. Esta la podemos obtener por estudios previos de este tema, por resultados de otros autores, por consulta de la literatura o mediante la realización de un pequeño estudio piloto que nos aproxime al valor probable que deberíamos obtener. Si no disponemos de ninguno de estos datos, utilizaremos el valor que traduce el resultado más desfavorable, es decir una probabilidad del 50% (0,5). Por tanto, el valor complementario  $q$  es igual a  $1-p$ , es decir  $0,5$ . Si la proporción esperada fuera de 35%,  $p = 0,35$  y  $q=1-p$ , es decir 0,65.

Veamos un ejemplo:

¿Cuál es el tamaño muestral de un estudio de prevalencia en una ciudad de 200.000 de habitantes?

El nivel de confianza normalmente utilizado es del 95% por lo que el valor  $Z=1,96$ , asumamos un nivel de error máximo asumible de 3% y una proporción esperada (según estudios previos) de un 20%

Utilizando la [calculadora Excel de Fisterra](#), obtenemos un tamaño muestral de 681 pacientes.

### **b) Estimar una media**

Si deseamos conocer el tamaño muestral para inferir el valor de la *media* de una variable a partir de una muestra a una población de referencia, debemos saber:

A: El nivel de confianza ( $1-\alpha$ ) que en nuestro caso suele ser del 95%. Igual que en el apartado anterior de estimación de proporciones.

B: La precisión, es decir el error máximo admisible, pongamos un 2,5%. La amplitud del intervalo de confianza es dos veces el valor de la precisión.

C: Realizar una estimación de la varianza de la distribución de la variable que hay en la población de referencia. Igual que en el caso anterior con proporciones, ésta la podemos conocer por estudios previos, por resultados de otros autores, por consulta de la literatura o mediante la realización de un pequeño estudio piloto que nos aproxime al valor probable que deberíamos obtener.

Veamos un ejemplo:

Si deseamos conocer la media de la glucemia basal de una población de 200.000 habitantes, con una seguridad del 95 % y una precisión de  $\pm 3$  mg/dl y tenemos información por un estudio piloto o revisión bibliográfica que la varianza es de 250.

Utilizando la [calculadora Excel de Fisterra](#), obtenemos un tamaño muestral de 107 pacientes.

### **Estimación del tamaño muestral para el contraste de hipótesis**

En la segunda situación (b), los test de contraste de hipótesis nos permitirán saber si las diferencias observadas entre las medias o las proporciones, estudiadas en las variables de dos grupos de pacientes, son significativas o se deben al azar.

### **a) Estimar una diferencia de medias**

En este caso queremos saber si existen diferencias significativas entre las medias observadas en dos grupos diferentes de pacientes. Para ello necesitamos conocer:

A: El nivel de confianza ( $1-\alpha$ ) que en nuestro caso suele ser del 95%. Una seguridad del 95% se corresponde a la región bajo la curva de una distribución normal comprendida entre la media  $\pm 1,96 Dt$  por ello el valor de Z será igual al 1.96. Una seguridad del 99% se corresponde a la región bajo la curva de una distribución normal comprendida entre la media  $\pm 2,58 Dt$  por ello el valor de Z en este caso será igual al 2,58.

B: La precisión, es decir el error máximo admisible, pongamos un 2,5%. La amplitud del intervalo de confianza es dos veces el valor de la precisión.

C: Realizar una estimación de la varianza de la distribución de la variable que hay en la población de referencia. Esto lo inferimos por estudios previos nuestros de este tema, por resultados de otros autores, por consulta de la literatura o mediante la realización de un pequeño estudio piloto que nos aproxime al valor probable que deberíamos obtener.

D: El poder estadístico deseado<sup>3</sup>, también conocido como potencia estadística, cuyo valor usual es del 90%.

E: Si el contraste va a ser unilateral o bilateral<sup>4</sup>.

Los cálculos pueden realizarse con la [calculadora Excel de Fisterra](#).

## **b) Estimar una comparación de dos proporciones**

En este caso queremos saber si existen diferencias significativas entre las proporciones observadas en dos grupos diferentes de pacientes. Para ello necesitamos conocer:

A: El nivel de confianza ( $1-\alpha$ ) que en nuestro caso suele ser del 95%. Igual que en el apartado anterior de estimación de diferencia de medias.

B: Las dos proporciones observadas de la variable que hay en la población de referencia. Esto lo podemos conocer por estudios previos nuestros de este tema, por resultados de otros autores, por consulta de la literatura o mediante la realización de un pequeño estudio piloto que nos aproxime al valor probable que deberíamos obtener.

C: El poder estadístico deseado, también conocido como potencia estadística, cuyo valor usual es del 80%.

D: Si el contraste va a ser unilateral o bilateral.

Los cálculos pueden realizarse con la calculadora Excel de Fisterra

---

<sup>3</sup> La potencia de una prueba estadística o poder estadístico es la probabilidad de que la hipótesis nula sea rechazada cuando la hipótesis alternativa es verdadera (es decir, la probabilidad de no cometer un error del tipo II). Por lo tanto, la potencia es igual a  $1 - \beta$ , que también se conoce como la sensibilidad

<sup>4</sup> Un contraste es bilateral cuando cualquiera de los dos parámetros a comparar (medias o proporciones) puede ser mayor o menor que el otro. No se establece dirección. Un contraste es unilateral cuando se considera que uno de los parámetros debe ser mayor que el otro, indicando por tanto una dirección de las diferencias. La hipótesis bilateral es una hipótesis más conservadora y disminuye el riesgo de cometer un error de tipo I; rechazar la  $H_0$  cuando en realidad es verdadera.

## Tamaño muestral ajustado a pérdidas

En todos los estudios se producen una serie de pérdidas de casos por diferentes circunstancias, como historias no encontradas, incumplimiento de criterios de inclusión y exclusión, negativa a colaborar del paciente, ausencia de registros de calidad, etc. Por ello el tamaño muestral debe ser mayor, aumentándolo en un cierto porcentaje. La cifra es variable según el tipo de estudio. A nivel orientativo, se puede utilizar un porcentaje de pérdidas del 15% de pérdidas.

Este tamaño muestral considerando las pérdidas se puede calcular:

$$\text{Muestra ajustada a las pérdidas} = n (1 / 1-R)$$

n = número de sujetos sin pérdidas

R = proporción esperada de pérdidas

Ejemplo: si en un estudio con 48 pacientes esperamos tener un 15% de pérdidas, el tamaño muestral necesario sería:  $48 (1 / 1-0.15) = 56$  pacientes en cada grupo.

## Glosario

*Población:* Conjunto de individuos de los que se quiere obtener una información ya los que queremos trasladar los resultados.

*Muestra:* Conjunto de elementos de análisis obtenidos de la población.

## Bibliografía

Argimón Pallas J.M., Jiménez Villa J. Métodos de investigación aplicados a la atención primaria de salud. 2ª ed. Barcelona: Mosby-Doyma; 1994.

Calculadora del tamaño muestral. <https://goo.gl/hxieCR>

Carrasco JL. El método estadístico en la investigación médica. 5ª ed. Madrid. Editorial Ciencia.

Contandriopoulos AP, Champagne F, Potvin L, Denis JL, Boyle P. Preparar un proyecto de investigación. Barcelona: SG Editores; 1991.

Cook TD., Campbell DT. Quasi-Experimentation. Design & Analysis Issues for Field Settings. Boston: Houghton Mifflin Company; 1979.

Dawson-Saunders B, Trapp RG. Bioestadística Médica. 2ª ed. México: Editorial el Manual Moderno; 1996.

Domenech JM. Métodos Estadísticos en Ciencias de la Salud. Signo. Barcelona 1999.

Hulley SB, Cummings SR. Diseño de la investigación clínica. Un enfoque epidemiológico. Barcelona: Doyma; 1993.

Kelsey IL, Thompson WD, Evans A. Methods in observational epidemiology. New York. Oxford University Press 1986.

Kleinbaum DG., Kupper LL., Morgenstern H. Epidemiologic Research. Principles and Quantitative Methods. Belmont, California: Lifetime Learning Publications. Wadsworth; 1982.

Martín Andrés A, Luna del Castillo JD. Bioestadística para las ciencias de la salud. 4ª ed. Madrid: NORMA; 1993.

Milton JS, Tsokos JO. Estadística para biología y ciencias de la salud. Madrid: Interamericana McGraw Hill; 2001.

Explorable.com (Oct 10, 2009). Muestreo probabilístico y aleatorización. Apr 30, 2018 Obtenido de Explorable.com: <https://explorable.com/es/muestreo-probabilistico>

Pita Fernández, S. Determinación del tamaño muestral. Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Universitario de A Coruña. CAD ATEN PRIMARIA 1996; 3: 138-14. <https://www.fisterra.com/mbe/investiga/9muestras/9muestras2.asp> Actualizada el 01/12/2010

## **CAPÍTULO 6**

### **Diseño y administración de instrumentos de recogida de datos**

**Autores:** Juan Hoyos Miller  
José Pulido Manzanero  
Enrique Regidor

## Resumen

En este tema se expondrán algunos elementos clave que se deben tener en cuenta para el diseño y posterior administración de un cuestionario. Se explicará cuál es el lenguaje a utilizar en un cuestionario y se ofrecerán algunas pautas para ordenar correctamente las preguntas a incluir. Seguidamente, se explicará que es una pregunta cerrada y abierta, cuáles son sus ventajas e inconvenientes y se hará lo mismo con los principales métodos de administración de una encuesta: auto administrada y cara a cara.

## Objetivos

- Establecer unas pautas básicas que permitan al alumno la elaboración de un instrumento de recogida de datos de calidad.
- Describir las principales ventajas e inconvenientes de los distintos métodos de administración de recogida de datos.

## Introducción

Existen infinidad de ejemplos de excelentes estudios epidemiológicos realizados con datos secundarios, esto es, que analizan datos que no han sido recogidos por el propio investigador, sino que utilizan en sus análisis, bases de datos ya existentes (Ver capítulo 7). Para aquellos que se encuentran en las fases iniciales de su carrera investigadora, el trabajo con datos ya existentes ofrece la posibilidad de elaborar análisis complejos sin necesidad de invertir las enormes cantidades de tiempo y esfuerzo que supone diseñar un instrumento de recogida de datos y administrarlo.

Sin embargo, tarde o temprano, para todo epidemiólogo e investigador en salud pública llega el momento en que se tienen que recoger datos primarios, ya sea porque con los datos se puede dar respuesta a su pregunta de investigación, o porque se quieren explorar nuevas hipótesis.

No trabajar con datos primarios también implica algunas debilidades en el proceso formativo. Haber estado involucrado en el diseño del instrumento y en el proceso de recogida de datos ofrece al investigador junior un marco en el que puede hacer una interpretación válida de los resultados. Su presencia en este proceso, permitirá conocer cómo se obtuvo la muestra, que dificultades se encontraron a la hora de recoger los datos, que implicaciones tienen las preguntas utilizadas etc. Todo ello facilitará conocer a fondo las limitaciones del estudio y sus potenciales implicaciones en el análisis e interpretación de los resultados.

A pesar de su importancia, las habilidades requeridas para el diseño y construcción de un buen cuestionario son poco tenidas en cuenta. Si nuestra materia prima es de mala calidad, nuestros resultados no pueden ser buenos. Es decir, lo que los anglosajones tan gráficamente describen como “garbage in-garbage out”. En este tema introduciremos algunos conceptos claves para diseñar un buen cuestionario y describiremos los distintos métodos de administración que existen y cuáles son sus ventajas y limitaciones.

## Diseño del cuestionario

Antes de comenzar a diseñar un cuestionario, deberemos tener clara cuál es nuestra pregunta de investigación y cuáles son los objetivos que nos planteamos. Son estos los que nos servirán de guía para la elaboración del mismo. De alguna forma, el cuestionario traslada la hipótesis y los objetivos de investigación a preguntas concretas. Las respuestas que se den a esas preguntas nos permitirán comprobar nuestra hipótesis de investigación.

Cuando un cuestionario es administrado a una persona lo que se pretende es que se produzca una transmisión precisa, válida y fiable de la información que nos interesa de dicha persona al investigador. Para ello un cuestionario debe:

- Ser apropiado
- Ser comprensible
- No dar lugar a ambigüedades
- No contener sesgos
- Las categorías de respuesta deben contener todas las posibles respuestas
- Haber sido validado

## Pautas a seguir para elaborar un cuestionario

El tono utilizado en el cuestionario deberá ser cordial y no excesivamente formal. Generalmente, el cuestionario se inicia con una breve presentación para explicar los objetivos del estudio y un breve agradecimiento por la participación al sujeto de estudio. En esta introducción también se le debe informar al participante de que el estudio (y el cuestionario) cumple los requisitos éticos, de la importancia de la encuesta y de la participación del encuestado. Deberemos informar de por qué le hemos elegido, qué se hará con los resultados, cual es la institución financiadora y cuál es la persona de contacto.

Los cuestionarios generalmente se inician con una serie de preguntas sobre características sociodemográficas que son de enorme importancia. Sexo, edad, lugar de nacimiento, etc. Estas preguntas se sitúan al principio por dos razones: sirven para “romper el hielo” y en segundo porque suelen ser variables muy relevantes para el análisis. Las preguntas sobre temas sensibles o muy personales no se deben situar nunca al principio.

Es normal que, en mayor o menor medida, se produzcan abandonos por agotamiento o desinterés. Esto es especialmente cierto en los cuestionarios auto-administrados. Conviene no dejar el bloque de preguntas más relevantes para el final, con el fin de asegurar que los valores perdidos en estas preguntas sean bajos. El cuestionario deberá finalizarse con un agradecimiento a la participación.

### Preguntas abiertas y cerradas

En las preguntas cerradas, las posibles respuestas son convertidas en un número variable de categorías estándar mientras que en las abiertas se le pide al entrevistado que use sus propias palabras para responder.

Tanto unas como otras tienen ventajas y limitaciones:

- Preguntas cerradas:

Ventajas:

- La información es fácilmente codificable
- Rápidas y fáciles de administrar

Inconvenientes:

- Pueden sugerir respuesta
- Pueden omitirse respuestas importantes
- Pueden introducir sesgos derivados de las categorías construidas
- No son útiles para informaciones complejas

- Preguntas abiertas:

Ventajas:

- Permiten observaciones o comentarios
- Pueden abarcar mejor la posible variabilidad o riqueza de las respuestas
- Menor riesgo de inducir respuestas

Inconvenientes:

- Llevan más tiempo
- Codificación más difícil
- La codificación a posteriori también puede introducir sesgos
- 

Normalmente, la gran mayoría de las preguntas que se incluyen en un cuestionario son preguntas cerradas.

### Redacción de las preguntas y categorías de respuestas

A la hora de elaborar las preguntas a incluir en el cuestionario, deberemos realizarlas de tal forma que sean inteligibles para las personas que van a contestarlo. El idioma y el lenguaje que se utilice deberán estar adaptados a la población a la que nos dirigimos. No tiene mucho sentido elaborar un cuestionario en inglés para administrarlo en zonas rurales de España y tampoco tiene sentido utilizar terminología médica en un cuestionario que no va a ser administrado a médicos. A ser posible, es mejor que las preguntas sean breves. Si vamos a introducir un concepto que sospechamos puede ser novedoso para el entrevistado, conviene hacer una breve explicación del mismo. Dicha explicación debe ir siempre antes de la pregunta y debe estar claramente diferenciada para no afectar la comprensión de la misma. Introducirla en un recuadro en gris u otro color ligeramente más oscuro que el blanco sería una buena manera de conseguirlo. Asimismo, la redacción de la pregunta debe estar libre de toda ambigüedad. La pregunta debe significar lo mismo tanto para el investigador como para todas las personas que contestan. Deben estar redactadas de forma sencilla y no contener dobles negaciones en su redacción. Hay que evitar las dobles preguntas y si la pregunta a realizar es muy complicada, es conveniente dividirla en varias preguntas sencillas.

A la hora de redactar las preguntas deberemos evitar generar sesgos o errores sistemáticos. Es importante tratar de evitar que la redacción de las preguntas y de las categorías de respuestas induzca la respuesta de las mismas. Otro de los sesgos a tener en cuenta en la redacción de las preguntas es el de memoria. Determinados acontecimientos importantes pueden ser mejor recordados (por ejemplo, el diagnóstico de una enfermedad grave). Para limitar este sesgo conviene ser cuidadoso en la introducción del marco temporal de referencia en la pregunta. Así, no tendrá mucho sentido preguntar por el número de veces que se ha consumido pan en los últimos 12 meses, pero sí podría tener sentido preguntar cuántas veces ha sido ingresado en un hospital en los últimos 12 meses.

Una pregunta y sus categorías de respuesta deben contener todo el universo de posibles respuestas. En realidad, esta es una aspiración imposible por lo que normalmente se incluye como última categoría de respuesta “otros/as: especificar”. Con ella se le da al encuestado la oportunidad de incluir una respuesta que considere que no está incluida en las categorías de respuesta cerrada. No es infrecuente que muchas de las respuestas incluidas en este “otros: especificar” puedan ser imputadas a posteriori por el investigador en alguna de las anteriores categorías. Si bien es un trabajo tedioso, es importante realizarlo antes de iniciar el análisis de los datos. Asimismo, hay que asegurarse de que las categorías de respuesta no se solapen y que sean mutuamente excluyentes.

## Estudio Piloto

Antes de poner en marcha el proceso de recogida de datos, deberemos llevar a cabo un estudio piloto con el fin de evaluar y validar el cuestionario. Podemos enviar el cuestionario a colegas expertos en el tema investigado y pedirles que revisen críticamente el cuestionario. El investigador deberá revisar a su vez los comentarios de los expertos y decidir si incorporarlos al cuestionario.

Una vez consideremos que tenemos un buen borrador, deberemos seleccionar a un grupo de personas de características similares a la población de estudio y administrarles el cuestionario. Podremos controlar el tiempo que tardan en cumplimentar y cuando hayan finalizado, podremos preguntarles acerca de aspectos relacionados con el formato o lenguaje que se han utilizado. El objetivo de este proceso es minimizar errores en el diseño que se le hayan podido escapar al investigador.

## ADMINISTRACIÓN DE CUESTIONARIOS

### Encuestas autoadministradas

Los cuestionarios auto administrados pueden cumplimentarse a través del tradicional “papel y lápiz” y mediante dispositivos electrónicos como teléfonos móviles, tabletas u ordenadores. Los programas existentes en el mercado para el diseño de cuestionarios online son cada vez más numerosos y se encuentran en mejora constante. Además, son relativamente baratos o incluso gratuitos. Cada vez permiten mayor flexibilidad en el diseño de cuestionario lo que era uno de sus puntos débiles. Estos avances, junto con el uso masivo de nuevas tecnologías en la población han hecho que cada vez más los cuestionarios auto administrados sean cumplimentados mediante este tipo de metodología. Esta vía de cumplimentación, ahorra mucho tiempo ya que las repuestas de los participantes van directamente a la base de datos evitando así el tedioso trabajo de introducir a mano en la base de datos las respuestas reflejadas en el cuestionario en papel. La introducción de los datos a mano es también una fuente de error ya que el encargado del mismo, puede confundirse durante el proceso dando lugar a un error de transcripción.

Cada vez más, las encuestas son administradas vía web. Se invita a la participación mediante banners o “mailing lists”. Este método de diseminación permite llegar a un amplio número de participantes sin encarecer el proceso. La principal desventaja se encuentra a la hora de llegar a poblaciones que puedan tener dificultades en la utilización de máquinas (pensemos en personas de más de 70 años por ejemplo), en cuyo caso tal vez sean más convenientes los cuestionarios en papel.

## Encuestas por entrevista cara a cara

Este tipo de encuestas se suelen utilizar en encuestas a población general. Suele ser encuestas extensas y que se realizan, a ser posible mediante visitas domiciliarias. También se pueden realizar por teléfono o Skype, estas modalidades están más sujetas a interrupciones.

En las entrevistas cara a cara es fundamental realizar una formación y entrenamiento previo a los entrevistadores, para eliminar al máximo el riesgo de que puedan influir en las respuestas. Deben existir unas instrucciones claras para el entrevistador, un manual de administración.

- Sus principales ventajas son:
  - Más seguridad de que el que responde es el seleccionado
  - Mayor tasa de respuesta al contestar
  - Posibilidad de ayuda si existen dificultades de comprensión, aunque esto es un arma de doble filo, pues puede introducir sesgos del entrevistador.
  - Buen método para plantear preguntas abiertas
  - Flexibilidad en la presentación de los ítems si es necesario presentar tarjetas con respuestas
- Sus desventajas:
  - Lentas y costosas. Requieren anuncios previos y visitas repetidas
  - Necesidad de Supervisión de entrevistadores lo que requiere mucho tiempo.
  - Sesgo del entrevistador (influencia de las características del entrevistador, deseabilidad social, deferencia interpersonal)
  - Requieren citar a la persona a encuestar lo que, dependiendo de su motivación y disponibilidad, puede resultar complicado.

¿Qué elementos tendré que considerar para elegir el tipo de encuesta?

Como hemos visto, cada método de administración tiene sus ventajas y sus inconvenientes y a la hora de elegir uno u otro tendremos que valorar algunos puntos como los siguientes.

- Contenido y tipo de datos recogidos
  - Si necesitamos realizar preguntas sensibles o no. En caso afirmativo, el investigador debería inclinarse por uno método auto administrado.
  - Necesidad o no de observar o realizar exámenes o medidas objetivas sobre los respondientes. Obviamente, en caso de necesitarlo es necesaria la presencia de un entrevistador entrenado.

- Se pueden hacer métodos mixtos. La entrevista puede ser cara a cara, pero al llegar a las preguntas sensibles, el entrevistador puede entregar al entrevistado una Tablet o un cuestionario en papel para que este cumplimente el bloque de preguntas sensibles a solas.
  
- Aspectos relacionados con la población. ¿Cuál es el tamaño de la muestra que tiene que cumplimentar el cuestionario? ¿Cuáles son sus características? ¿Saben leer? ¿Es esperable que necesiten aclaraciones acerca de algunos conceptos que aparecen en el cuestionario? ¿a través de qué medios podremos acceder a la población de estudio? Todas estas preguntas se deberán plantear a la hora de escoger un método de administración.
- Disponibilidad de marco muestral. La necesidad de que nuestra muestra sea representativa influirá en la selección de un método u otro.
- Recursos disponibles. La disponibilidad de recursos humanos, económicos y de tiempo tendrá que ser tenida en cuenta para una elección realista del método de administración.
- Tiempo y molestias de realización de la encuesta para el respondiente. Siempre conviene que el cuestionario sea lo más breve posible. No obstante, si la longitud es importante, cabe la posibilidad de usar incentivos siempre que nuestro presupuesto lo permita. Asimismo, si el cuestionario es muy largo es esperable que la tasa de abandonos sea menor si el cuestionario es cumplimentado por el entrevistador.
- Posibilidad de formar y supervisar el trabajo de los entrevistadores durante el trabajo de campo. Una correcta formación y supervisión de los entrevistadores minimizará la posible inclusión de sesgos por parte del mismo.

## **Bibliografía**

Buring JE. Primary data collection: what should well-trained epidemiology doctoral students be able to do? *Epidemiology (Cambridge, Mass)*. 2008;19(2):347-9.

Casas Anguita J, Repullo Labrador JR, Donado Campos J. [Surveys as a research technique. Composition of questionnaires and statistical processing of data (I)]. *Atencion primaria*. 2003;31(8):527-38.

Stone DH. Design a questionnaire. *BMJ (Clinical research ed)*. 1993;307(6914):1264-6.

## **CAPÍTULO 7**

### **Datos Secundarios**

**Autores:** Lucía Cea Soriano  
Paloma Astasio Arbiza  
Paloma Ortega Molina

## Objetivos

En el presente capítulo se abordan los siguientes apartados

- a) Transmitir y asimilar el concepto de datos secundarios
- b) Tipos de fuentes de información secundarias para la investigación epidemiológica
- c) Aplicabilidad, ventajas e inconvenientes de la utilización de estas fuentes en la investigación clínica-epidemiológica.

## Introducción

En primer lugar, es importante discernir entre datos primarios y datos secundarios a la hora de realizar una investigación en medicina clínica. El primero, se origina a propósito de un tema concreto de investigación, es por tanto que recae directamente en el equipo investigador, no únicamente su diseño, recopilación y manejo, sino también su coste y factibilidad. Un ejemplo de datos primarios serían las entrevistas, exámenes y exploraciones médicas que se realizan de forma directa. Por el contrario, los datos secundarios, incluye información ya sea de diagnóstico, estilo de vida etc., que ya han sido recopilados y registrados con anterioridad, normalmente con un objetivo diferente a la investigación clínica. Es por ello que la calidad de los mismos no está sujeta al equipo investigador. El presente capítulo, tiene como objeto de interés, este segundo tipo de información.

## Desarrollo

Los sistemas de información sanitaria (SIS) disponibles en España, se diferencian y clasifican en función de su representatividad, es decir, su denominador (1). Si éste incluye a toda la población o bien a una muestra representativa, siendo el marco de referencia toda la población, hablaríamos de SIS de base poblacional, un ejemplo sería el registro de defunciones. Mientras que si el denominador lo constituye aquellos que conforman la institución sanitaria como es el Conjunto Mínimo Básico de Datos del Hospital (CMDDB), entonces nos referiríamos a los sistemas de información sanitaria procedente de los servicios sanitarios.

### Sistemas de información sanitaria de base poblacional

Entre ellos se identifican tres tipos: los registros, las encuestas y los sistemas de notificación poblacional. En cuanto a los registros, el Instituto Nacional de Estadística (INE) tiene a disposición de todos los ciudadanos los datos referentes a mortalidad, nacimientos y muertes fetales tardías (2). Estos tres registros constituyen, no sólo la base del conocimiento del movimiento natural de la población de nuestro país, sino también, la evaluación del estado de salud de la población, la vigilancia epidemiológica y planificación sanitaria. La clasificación de defunciones por causa, donde también se integra la muerte fetal tardía comenzó en 1861-

1870, pero no fue hasta 1900 cuando se estableció una clasificación internacional. En cuanto a las muertes fetales tardías, y su notificación se ha adoptado el criterio de considerar como tal aquellas muertes en las que el feto tenga un peso igual o superior a 500 g. El registro de nacimientos también tiene utilidad en el cálculo de la mortalidad infantil (indicador sanitario), esta información está disponible a través del registro civil y el boletín estadístico de partos.

Dentro de los registros también destacan aquellos centrados en las enfermedades, como es el registro de cáncer o el de VIH (3,4). Gracias a la existencia de los mismos es posible conocer medidas de frecuencia como es la incidencia de la enfermedad. A la hora de tener en cuenta que tipo de enfermedad puede constituir un registro, estas deben de reunir una serie de requisitos como son: i) una alta importancia sanitaria y social, ii) alta mortalidad y iii) un gran impacto asistencial. El registro de cáncer es imprescindible para determinar la supervivencia de los pacientes a nivel poblacional y para conocer la prevalencia. También tienen otras muchas finalidades siendo una de ellas la evaluación de los programas de diagnóstico precoz y por tanto la evaluación de las medidas preventivas y tratamientos, En España, los dos registros de cáncer de población más antiguos son el Registro de Cáncer de Zaragoza, creado en el año 1960, y el de Navarra, creado en 1970. En la actualidad, existen 14 RCP de Población cuyos datos están incluidos en la publicación de referencia Cancer Incidence in Five Continents editada por la International Agency for Research on Cancer (IARC)

En relación con las encuestas de base poblacional, éstas pueden ser llevadas a cabo por medio de entrevista personal o por teléfono. Tradicionalmente las entrevistas personales obtienen una mayor tasa de respuesta, posibilidad de mostrar material de apoyo (i.e. material visual) que ayude a cumplimentar las preguntas de una forma más precisa. Como desventaja, estas son costosas. La validez de la información proporcionada va a estar sujeta al tipo de información a extraer (ej. enfermedad crónica vs. enfermedad transmisible) que dependerá del recuerdo del entrevistado. Hoy en día las encuestas son una herramienta útil para el estudio de enfermedades crónicas de baja letalidad. A su vez, éstas permiten recoger una gran cantidad de información no exclusivamente clínica, sino datos referentes a estatus socioeconómico, datos demográficos, de estilo de vida y conducta y de incapacidad, donde los registros quedan limitados. Esta información confiere la base del conocimiento sobre morbilidad, utilización de servicios sanitarios, hábitos de vida. En España destacan diferentes encuestas de base poblacional como es la encuesta de salud, la encuesta de población activa, la encuesta de presupuestos familiares, la encuesta de condiciones de vida (5,6)

Por último, dentro del bloque de SIS de base poblacional se encuentran otras fuentes como pueden ser las víctimas de accidente de tráfico (7), realizada por la dirección general de tráfico del ministerio de interior, donde se incluye diversa información sobre las características accidentes, víctimas y vehículos implicados. La notificación de accidentes de trabajo, elaborados por el ministerio de trabajo y seguridad social a partir de la

notificación que realizan las empresas sobre el tipo de accidente y trabajador accidentado, aunque hay que tener en cuenta que existe la posibilidad de infra-notificación dado que no todo accidente causa baja. Y, por último, las enfermedades profesionales, también realizada por la autoridad laboral. Para estas últimos España cuenta con un listado definido.

### **Sistemas de información sanitaria procedente de los servicios sanitarios**

Desde comienzos de los años 90, los hospitales como normativa ministerial, registran los datos de cada paciente atendido en el hospital. Recientemente se ha incorporado los episodios de cirugía sin ingreso y los hospitales de día. Toda esta información está integrada en el Conjunto Mínimo Básico de Datos (CMBD) (8). Como fuente de información está constituido por la historia clínica del paciente y el informe al alta. Esto proporciona información valiosa para conocer la realidad sanitaria de la población incluyendo datos demográficos (edad, sexo, localidad de residencia), diagnóstico que ha motivado el ingreso, factores de riesgo, comorbilidades y complicaciones durante el ingreso. Además de facilitar técnicas diagnósticas e intervenciones terapéuticas, con especial relevancia: las quirúrgicas. Para su codificación, siguen la clasificación internacional de enfermedades (CIE v.9). Para poder establecer la temporalidad de cada uno de los episodios consta la fecha de ingreso y alta. Además, incluye tipo de ingreso y circunstancia al alta, datos que pueden facilitar información acerca de las necesidades y recursos hospitalarios.

Otros sistemas de información a destacar son los sistemas de notificación obligatoria de una serie determinada de enfermedades de los que se encarga la Red Nacional de Vigilancia Epidemiológica (9), donde tienen a disposición protocolos y reglamentos en como declarar cada caso. Generalmente incluyen enfermedades de origen infeccioso. Al tratarse de un sistema de vigilancia su análisis suele incluir una estimación sobre las tendencias de aparición y mortalidad.

### **Base de datos de atención primaria: BIFAP (10,11)**

Por último, cabe destacar una fuente de información que se encuentra a caballo entre los dos SIS que se han mencionado, es la base de datos de atención primaria BIFAP (Base de Investigación Farmacoepidemiológica en Atención Primaria) perteneciente a la Agencia Española del Medicamento y Productos Sanitarios. España tiene una asistencia sanitaria universal, lo cual permite al médico de atención primaria recabar la información de sus pacientes de forma integral, al recoger comorbilidad y diagnóstico clínico, utilización de medicamentos incluyendo detalles de dosis, duración y posología en sus prescripciones, datos de edad, sexo, indicadores de estilo de vida y resultados de pruebas analíticas y complementarias. Es esta característica de universalidad lo que hace que dicha base de datos de atención primaria, pese a que contenga información procedente de un sistema sanitario, pueda constituir una muestra representativa de la población dado que la mayoría de población española está adscrita a un médico de atención primaria del sistema nacional de salud. BIFAP

contiene la información registrada en la práctica clínica habitual aportada por médicos de familia y pediatras de atención primaria del Sistema Nacional de Salud de 9 Comunidades Autónomas, desde el año 2000, con una cobertura del 16% de la población española, actualizándose de forma anual. Específicamente, BIFAP contiene datos de aspectos sociodemográficos (edad, sexo, indicadores de estilo de vida, problemas sociales), clínicos y diagnósticos, interconsultas, pruebas complementarias de laboratorio, prescripciones, y antecedentes clínicos personales y familiares de pacientes anónimos. Es importante destacar que la única

variable identificativa de los pacientes, siendo ésta el número de historia clínica es sometida a un procedimiento de anonimización/criptación por capas, efectuada por los respectivos responsables de los ficheros de las diferentes organizaciones. Cada paciente constituye un registro, con toda la información longitudinal contenida desde la primera consulta con su médico hasta la fecha. Por último, para que los datos se utilicen en los estudios de investigación, deben cumplir unos criterios mínimos de calidad de registro exigidos a la información del cupo, y del paciente (10, 11)

## Resumen

Los datos primarios, surgen a propósito de un tema concreto de investigación, es por tanto que recae directamente en el equipo investigador, no únicamente su diseño, recopilación y manejo, sino también su coste y factibilidad. Por el contrario, los datos secundarios, incluye información ya sea de diagnóstico, estilo de vida, que ya han sido recopilados y registrados con anterioridad, normalmente con un objetivo diferente a la investigación clínica. Los sistemas de información sanitaria (SIS) disponibles en España, se diferencian y clasifican en función de su representatividad, es decir, su denominador. Pudiéndose distinguir por tanto los SIS de base poblacional cuando el denominador constituye a toda la población (ej. registro de defunciones) y los SIS procedentes de servicios sanitarios cuyo denominador lo conforma la institución sanitaria. Los SIS de base poblacional incluyen los registros, las encuestas y los sistemas de notificación poblacional entre los que se encuentran las víctimas de accidente de tráfico o la notificación de accidentes de trabajo. En cuanto a los SIS procedentes de los servicios sanitarios, destacan el Conjunto Mínimo básico de datos (CMBD) desde los años 90 y los sistemas de notificación obligatoria de una serie determinada de enfermedades de los que se encarga la Red Nacional de Vigilancia Epidemiológica. Por último, la base de datos de atención primaria BIFAP (Base de Investigación Farmacoepidemiológica en Atención Primaria) perteneciente a la Agencia Española del Medicamento y Productos Sanitarios se encuentra a caballo entre los dos SIS mencionados constituyendo una potente herramienta de investigación no sólo por la amplia información contenida si no por la alta representatividad de sus datos.

## Bibliografía

1. Regidor Poyatos E, Miqueleiz Autor E, Calle Purón ME, de Juanes Pardo JR, Domínguez Rojas V, Rodríguez Artalejo F. Sistemas de Información Sanitaria: Registros, encuestas de salud, y sistemas de notificación. En: Medicina Preventiva y Salud Pública Piedrola Gil.Fernandez-Crehuet Navajas J, Gestal Otero JJ, Delgado Rodríguez M, Bolúmar Montrull F, Herruzco Cabrera R, Serra Majem L, Rodríguez Artalejo F. Elsevier Masson. Barcelona 2015. p 217-227.
2. Movimiento natural de las poblaciones. Instituto Nacional de Estadística (INE). [http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica\\_P&cid=1254735573002](http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735573002)
3. REDECAN. <http://redecn.org/es/index.cfm>

4. Vigilancia epidemiológica del VIH. <https://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/home.htm>
5. Encuesta Nacional de Sanidad Instituto Nacional de Estadística (INE) <https://www.msssi.gob.es/estadEstudios/estadisticas/encuestaNacional/home.htm>
6. Encuestas mercado laboral, Instituto Nacional de Estadística (INE) [http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica\\_P&cid=1254735976595](http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735976595)
7. Víctimas de tráfico <http://www.dgt.es/es/seguridad-vial/unidad-de-victimas-de-accidentes-de-trafico/>
8. CMBD. <https://www.msssi.gob.es/estadEstudios/estadisticas/cmbdhome.htm>
9. Red Nacional de Vigilancia Epidemiológica <http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/vigilancias-alertas.shtml>
10. Salvador Rosa A, Moreno Pérez JC, Sonogo D, García Rodríguez LA, de Abajo Iglesias FJ. El Proyecto BIFAP: Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria. Aten Primaria. 2002 Dec;30(10):655-61
11. BIFAP. Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria. <http://www.bifap.org/ques.php>

## **CAPÍTULO 8**

### **Revisiones sistemáticas y Metaanálisis. Conceptos básicos**

**Autores: Luis Sordo del Castillo  
José Pulido Manzanero  
David Martínez Hernández**

## Introducción

El número de publicaciones sobre cualquier tema relacionado con la salud es tan elevado que a veces es complejo llegar a conclusiones ante la heterogeneidad de sus resultados. Esto hace que, en muchas ocasiones, más que plantear estudios que den respuesta a una hipótesis, cada vez esté cobrando más importancia la revisión y evaluación crítica de los ya existentes. Es de ahí de donde viene la importancia y eclosión de las revisiones sistemáticas y los metaanálisis.

### 1. Concepto y utilidad

Una revisión sistemática es una investigación sobre una pregunta concreta, cuya unidad de análisis son los estudios originales que dan respuesta a esa pregunta, en la que se sigue un proceso sistemático y explícito de selección e interpretación de sus resultados para la obtención de conclusiones globales. Dicho de otra manera, una revisión sistemática es una investigación no original a partir de otras investigaciones previas que compila la información de las mismas, la evalúa y la valora para llegar a resultados generales.

Los términos “revisión sistemática” (RS) y “metaanálisis” a veces se emplean como sinónimos, pero no lo son. En una revisión se eligen y seleccionan aquellos estudios que dan respuesta a la pregunta de investigación. Si esta compilación decide analizarse, estaremos ante un metaanálisis, pero en ocasiones los resultados obtenidos de diferentes estudios son tan disímiles que no permiten ningún análisis. En tal caso estaremos ante una simple RS sin metaanálisis. Aunque no se trata de una definición que sea aceptada por el 100% de la comunidad científica sí es la más usada y operativa para entender de qué hablamos. Así pues, todo metaanálisis parte de una RS, pero no todas las RS permiten la realización de un metaanálisis.

### 2. Revisiones sistemáticas

- A. Antes de empezar:** Al igual que cuando hacemos una investigación primaria, lo primero que tenemos que hacer antes de iniciar una RS es ver si tal revisión ya ha sido realizada. La base de datos de revisiones más aceptada internacionalmente es la Cochrane, y allí debemos buscar si lo que nosotros queremos saber ya ha sido estudiado de esa misma manera. Además de buscar en la Cochrane (no todas las revisiones siguen su metodología), es bueno comprobar que no existen revisiones en otros lugares (si es en el campo biomédico se recomienda pubmed). Una vez que nos hemos asegurado de que la revisión no existe, es conveniente registrar que la estamos haciendo para que otras personas lo sepan y no se duplique el trabajo. Existen diferentes registros siendo el más conocido “Prospero”. <https://www.crd.york.ac.uk/prospero/>
- B. La pregunta:** Una RS se inicia desde la pregunta de investigación. Toda RS pretende “poner orden” en la respuesta a tal pregunta y su misión es la de seleccionar aquellos artículos que se plantearan en su investigación esa misma pregunta o la contuvieran entre sus resultados. La pregunta de investigación de una RS tiene que ser muy precisa para poder ser respondida de la misma manera por aquellos artículos que se la plantearan. Por ello, tras establecerla (por ejemplo: ¿Se relaciona el consumo de cocaína con el

C. aumento de convulsiones epilépticas? O ¿Protege la aspirina del infarto de miocardio?), hay que determinar qué tipo artículos vamos a elegir.

D. **Las fuentes primarias: Criterios de selección de los artículos:** Ya con una pregunta que responder, una de las dificultades con las que nos encontramos es que existen DEMASIADOS artículos que dan respuesta a tal pregunta y que lo hacen de una manera heterogénea. Así que habremos de definir qué tipo de artículos seleccionaremos. Lo normal es que las RS se centren en los estudios primarios de máxima calidad en función de su diseño. Y en este sentido los ensayos clínicos y los estudios de cohortes son los elegidos. Esto no quiere decir que no podamos hacer RS de preguntas que hayan sido respondidas por estudios de peor calidad que los señalados. Podemos y se hace. Pero solo cuando no tenemos estudios “mejores”. Así pues, deberemos especificar qué tipo de estudios van a formar parte de la RS por su diseño.

Además, es frecuente que existan otras especificaciones que hacer a la hora de seleccionar los estudios en función de la pregunta. Las más frecuentes se refieren a criterios geográficos o temporales. Por todo ello, muchas revisiones tienen “criterios de inclusión y exclusión” que van a definir qué se selecciona y qué no. Podemos excluir aquellos estudios realizados antes de una determinada fecha o un determinado entorno que esté fuera de nuestra pregunta de interés (por ejemplo, prisiones); podemos seleccionar solo aquellos estudios que tuvieron cuidado en definir las variables como nosotros queremos. En definitiva, al igual que en un estudio primario se establecen los criterios para definir qué personas participan en un estudio, en una RS se determinan las que han de tener los artículos a seleccionar.

E. **Extracción de los artículos:** Una vez decididos los criterios que buscamos, idealmente hemos de ser capaces de encontrar TODO lo que se haya publicado de este tema con tales características. Para ello, existen diferentes repositorios virtuales o bases de datos de artículos científicos que lo permiten. No es objeto de este documento establecer las estrategias de búsqueda, pero sí citaremos las bases más frecuentemente empleadas: Pubmed, PsylInfo, Lilacs, Embase y Google Scholar. En todas ellas podremos buscar por palabras clave y términos libres aquellos artículos que tengan las características que hemos predefinido.

F. **Selección de los artículos:** Por mucho que se refine la búsqueda, lo normal es que seleccionemos menos de un 5% de los artículos que inicialmente extraigamos. Eso hace que la labor más importante (y tediosa) de una revisión consista en “tamizar” los artículos iniciales en busca de aquellos que verdaderamente cumplen los criterios predefinidos. En ocasiones nos valdrá con leer el título, otras veces habrá que recurrir al resumen y en ocasiones tendrá que ser el texto completo el que nos dé respuesta a si el artículo cumple o no para dar respuesta a la revisión.

G. **Valoración de la calidad de los artículos seleccionados:** Una vez que tenemos todos los artículos seleccionados, habremos de establecer la calidad de los mismos. Todos ellos responden a la pregunta planteada, pero es evidente que no todas las investigaciones son igual de rigurosas. Para la evaluación de

la calidad de los diferentes estudios, existen diferentes herramientas. Todas ellas “califican” a los artículos seleccionados en función de diferentes ítems. Desde el nivel de desarrollo de sus análisis a las variables

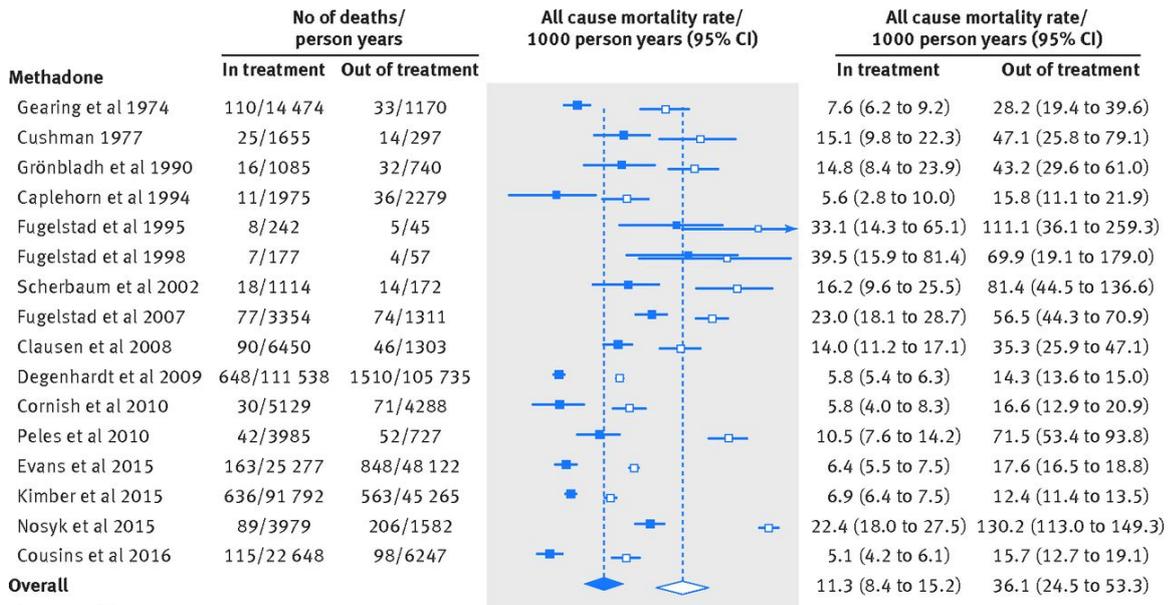
- H. incluidas o la posibilidad de sesgos. De esta forma a cada artículo seleccionado se le suele añadir una “calificación” basada en ello. La valoración de la calidad se hace siguiendo ítems preestablecidos. Existen muchas herramientas que nos facilitan esta evaluación siendo una de las más usadas la NHS Quality Improvement Scotland
- I. **Resultados:** Toda revisión sistemática muestra las características de los estudios seleccionados, y sus resultados SIN AGREGARLOS. Se hace en forma de tabla (más allá de que luego se incluya cierta información en el texto) y en ella deben verse, como mínimo las siguientes características de cada uno de los estudios seleccionados: Año de publicación, lugar, número de pacientes, características principales de los mismos, variables de interés seleccionadas y resultados referidos a la pregunta de interés de la RS ajustados y sin ajustar.

Si la RS enseña artículos similares que pueden ser analizados de manera conjunta, se realizará un metaanálisis. Lo que ocurre es que, en muchas ocasiones, la respuesta a la pregunta es tan diversa, tan heterogénea, que no permite hacer esto. En este segundo caso, los resultados de la RS terminan en este punto, teniendo que ser mediante la discusión de los mismos la manera en que saquemos conclusiones globales.

## 2. Metaanálisis

Como ya se ha dicho, el metaanálisis es la culminación “cuantitativa” de una revisión sistemática. Se produce al atribuir a cada artículo seleccionado un valor (como si cada artículo fuera un paciente) y al analizar de manera conjunta estos valores. Su objetivo es integrar estadísticamente los resultados. Las técnicas estadísticas requieren del uso de programas de software específicamente diseñados para este propósito. Los análisis de la información de los artículos son complejos y aprender a hacerlos está fuera de los objetivos de este documento. No obstante, señalaremos que el más empleo y con diversos tutoriales para aprender a utilizarlo es el programa Review Manager (RevMan 5.3), elaborado por el grupo Cochrane. Este programa permite obtener las estimaciones combinadas y obtener su representación gráfica.

**Figura 1: Ejemplo de la representación de resultados en un metaanálisis (de ejemplo práctico 2)**



En el ejemplo se muestra que cada línea azul representa el resultado de cada artículo seleccionado. Los rombos finales son el resultado agrupado, el resultado del metaanálisis.

A los metaanálisis no solo se les pide que nos den una estimación agrupada del resultado (ver figura 1). También a través de ellos podremos ver la variabilidad observada en los estudios. En ocasiones los artículos incluidos en el metaanálisis tienen resultados muy similares, pero otras veces estos son muy heterogéneos. En ambos casos obtendremos una determinación final, pero es importante saber si esta cifra final procede de muchos resultados similares o muy heterogéneos. Es decir, conocer su grado de variabilidad.

### 3. Ejemplos prácticos

Aquí pueden verse diferentes revisiones sistemáticas con o sin metaanálisis en las que pude comprenderse de manera práctica lo explicado:

- Revisión sistemática SIN metaanálisis de estudios transversales y casos y control en referencia a la correlación entre las convulsiones y el consumo de cocaína

[A systematic review of evidence on the association between cocaine use and seizures](#)

En este caso, la variabilidad enorme en los estudios que estudiaban la relación buscada y sus diferentes metodologías hizo que no fuera recomendable agrupar los datos (metaanalizarlos), por lo que los resultados se discutieron de manera “cualitativa” (a las Revisiones sistemáticas si metaanálisis a veces se las llamaba “metaanálisis cualitativos”)

- Revisión sistemática CON metaanálisis de estudios de cohortes sobre la mortalidad de los pacientes en tratamiento con metadona y tras salir de él

[Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies](#)

### Bibliografía

1. Ferreira González I, Urrútia G, Alonso-Coello P. Revisiones sistemáticas y metaanálisis: bases conceptuales e interpretación. *Rev Esp Cardiol.* 2011;64:688-96
2. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283:2008-12.

3. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *J Clin Epidemiol.* 1995;48:167-71
4. Meseguer Guaita F. Lectura crítica de un Metaanálisis y de una revisión sistemática. Guía Salud Murcia. Disponible en: [https://www.murciasalud.es/recursos/ficheros/136630-capitulo\\_11.pdf](https://www.murciasalud.es/recursos/ficheros/136630-capitulo_11.pdf)

## **CAPÍTULO 9**

### **Investigación de un brote epidémico**

**Autores:** Luis Sordo del Castillo  
Juan Hoyos Miller  
Rosa Villanueva Orbaiz

**Objetivos:** Establecer las pautas de estudio de un brote epidémico para: (1) conocer sus causas, (2) evitar su extensión y agravamiento en la población mediante medidas de control, (3) conocer mejor las características clínico-epidemiológicas de la enfermedad (4) Ayudar a detectar enfermedades nuevas (emergentes o reemergentes) e (5) Identificar factores de riesgo que podrían haber pasado desapercibidos en su origen.

### Introducción y concepto

Entendemos como **brote epidémico** la *existencia de un número de personas afectadas por una determinada enfermedad o que presentan factores de riesgo específicos u otra característica relacionada con la salud, con una frecuencia claramente superior a la esperada en condiciones normales, en un ámbito geográfico y período de tiempo determinados*. Es decir, determinamos la sospecha de un brote en función de lo esperado para un determinado lugar y tiempo, y no solo lo centramos en enfermedades. Un aumento brusco en el consumo de sustancias psicoactivas o de hábitos relacionados con la salud pueden ser definidos como brote.

### Fases de un brote epidémico

**1. Sospecha de brote epidémico:** La sospecha de la existencia de un brote epidémico conlleva una serie de actuaciones que tienen que tomarse de manera rápida:

- **Notificación urgente** al nivel sanitario inmediatamente superior.
- **Iniciar la investigación:** La información disponible en los primeros momentos es probable que sea mínima e incluso confusa. De ahí la importancia de recoger la información que pueda servir de orientación acerca de la magnitud del problema. Para ello, se recogerán datos sobre:
  1. Características clínicas y gravedad de la enfermedad.
  2. Fecha de aparición de los primeros síntomas.
  3. Número de personas afectadas y expuestas y características personales
  4. Localización espacial del brote.
  5. Hipótesis causales iniciales, perfilándolas si esto fuera factible.
- **Establecer medidas de control inmediato:** Sin un mínimo conocimiento de las características de un brote es muy complejo establecer medidas de control. No obstante, desde el momento en el que se establece la sospecha del mismo hemos de priorizar toda actuación encaminada a finalizar el brote o evitar que sigan apareciendo afectados. Las situaciones en que más claramente está indicado abordar medidas de control inmediato son:
  - Aquéllas en que la investigación preliminar sugiere una fuente común de infección
  - Aquéllas en que la enfermedad se transmite persona a persona, y clínicamente son lo suficientemente graves como para tomar medidas de control inmediato

### Recogida de información

Se debe obtener información sobre las personas afectadas, expuestas y no expuestas, mediante entrevistas. Si es factible, recoger información de sus pruebas analíticas. Esto nos ayudará a establecer las hipótesis de trabajo y a esclarecer las causas y factores que han condicionado la ocurrencia del brote. Cada brote es diferente, pero esta es la información que en general se necesita recoger:

- A. **DEMOGRÁFICA:** la distribución de la población, por grupos de edad y áreas geográficas, será fundamental para establecer las tasas de incidencia de la enfermedad.
- B. **SOCIOECONÓMICA Y CULTURAL:** El objetivo es conocer la posible existencia de contextos de marginación, hábitos sociales, etc., que impliquen un riesgo para la salud.
- C. **SANEAMIENTO BÁSICO:** Deberemos indagar acerca del origen, tratamiento y distribución de agua potable, y red de aguas residuales.
- D. **DISTRIBUCIÓN, MANIPULACIÓN Y CONSUMO DE ALIMENTOS:** La ocurrencia de tox infecciones alimentarias exige una detallada investigación acerca de:
  - a. Cuándo y dónde se han obtenido las materias primas utilizadas en la elaboración de los alimentos, y cómo se han transportado y almacenado.
  - b. Técnicas y utensilios utilizados en la preparación de los mismos.
  - c. Cantidad de alimentos elaborados, así como fecha de elaboración, conservación, distribución y consumo.
  - d. Condiciones higiénico-sanitarias de los establecimientos en los que se han elaborado y consumido los alimentos implicados.
- E. **CONTAMINACIÓN AMBIENTAL E INDUSTRIAL**
- F. **ANTECEDENTES DE CASOS Y BROTES DE LA MISMA ENFERMEDAD EN LA ZONA**

## 2. Definición de caso

Una vez establecida la sospecha de brote a partir de la información recogida, hay que verificar el diagnóstico de la enfermedad. Por ello han de definirse aquellos criterios por los cuales un individuo puede ser considerado CASO bajo las condiciones de estudio. Estos no son rígidos y varían mucho de un brote a otro. Lo importante es determinar en lo posible aquellos criterios que se consideren necesarios para definir la enfermedad según los siguientes ejes:

1. **Características clínicas:** Son los signos y síntomas del enfermo; su magnitud, extensión, duración de cada uno y secuencia de aparición.
2. **Información procedente de pruebas de laboratorio:** Puede incluirse desde el aislamiento de un agente específico a la identificación de su serotipo, la evidencia de anticuerpos u otros indicios biológicos o químicos.

3. Características epidemiológicas: Estas deben incluir: Fecha de comienzo de la enfermedad, exposición del caso a una comida, agua u otra fuente sospechosa, presencia en un lugar determinado durante un tiempo dado, etc.

A partir de ahí se realizará la definición de caso dependiendo de lo avanzado del diagnóstico:

**CASO CONFIRMADO:** Individuo de quien se ha obtenido evidencia definitiva de laboratorio en cuanto al agente etiológico o enfermedad. Ej.: Serología positiva.

**CASO PROBABLE:** Individuo del que se tiene sospecha de laboratorio de infección o de enfermedad. Ej.: Un simple título de anticuerpos indicativo de enfermedad o infección.

**CASO SOSPECHOSO:** Individuo que tiene signos y síntomas compatibles (que podrían estar asociados) con la enfermedad.

En ocasiones es necesario distinguir entre una definición inicial y otra definitiva, si estamos aún pendientes de tener toda la información. La inicial suele ser más sensible para asegurarnos de no tener casos sin detectar (evitar los falsos negativos).

### 3. Confirmación de la existencia de un brote

La confirmación del brote epidémico se basará en la comparación entre el número de casos ocurridos («observados») y el número de casos «esperados», por unidad de tiempo y lugar determinado; es decir, confirmar que la presencia de un determinado número de casos de la enfermedad diagnosticada en la comunidad aparece como excesiva con relación a la frecuencia habitual de dicha enfermedad en la misma población y tiempo.

Con las cifras de casos esperados y observados se puede calcular el **índice epidémico**, cociente entre los casos observados (incidencia actual) y casos esperados (incidencia habitual).

$$\text{Índice Epidémico} = \text{Nº casos observados} / \text{Nº casos esperados}$$

Se confirma la existencia de un brote epidémico cuando el número de casos observados es mayor que el número de casos esperados. Es decir, cuando el **índice epidémico** es mayor de 1.

### 4. Estudio descriptivo del brote

La caracterización del brote es imprescindible para tomar las medidas de control si desconocemos la causa. Este estudio descriptivo se realizará sobre tres ejes:

A. Patrón temporal de la enfermedad: Nos sirve para determinar cuánto dura o ha durado el brote, lo cual nos ayuda a confirmar o rechazar si los casos que estamos observando pueden responder a la conducta epidémica de esa enfermedad determinada (por ejemplo: una alta incidencia sostenida de procesos respiratorios agudos en los meses de primavera y verano es improbable que sea una gripe, que cursa en ondas estacionales y generalmente en invierno).

El tiempo como variable epidemiológica sirve para la construcción y uso de la curva epidémica, gráfico en el cual los casos de una enfermedad se representan según el momento de su inicio. La información fundamental que nos suministra una «curva epidémica» es el discernimiento acerca de si el brote se ha producido por fuente común como podría ser el agua (en este caso la curva epidémica sigue una distribución normal) o por propagación, cuando se transmite persona a persona (con una curva con distribución no normal). En algunos brotes ambos mecanismos se combinan. Si se conoce el período de incubación de la enfermedad y las fechas de aparición de los casos, se puede identificar el momento más probable de exposición de los casos a la fuente.

B. Lugar: La zona o zonas en las que localizamos el brote. El objetivo es elaborar el «mapa de casos» que tiene como finalidad permitirnos identificar la pauta de distribución del mismo.

C. Características personales de los afectados: Estas pueden describirse en función de sus características inherentes o adquiridas (edad, sexo, raza, estado inmunitario...), actividades (trabajo, costumbres...), condiciones en las que viven (sociales, económicas...) y otras.

## 5. Estudio analítico del brote

El análisis epidemiológico nos va a permitir demostrar que las personas que estuvieron expuestas a la fuente sospechosa tuvieron un riesgo de enfermar significativamente mayor que aquellas personas que no estuvieron expuestas. El método más simple, para demostrar un mayor riesgo entre las personas expuestas, es la comparación de las tasas de ataque. Estas tasas se han de calcular tanto para las personas expuestas a la fuente sospechosa como para las no expuestas:

**Tasa de ataque expuestos (TAE):** probabilidad de que una persona expuesta al factor de riesgo (FR) desarrolle la enfermedad.

$TAE = \frac{\text{Nº de personas expuestas al FR que enferman}}{\text{Nº de personas expuestas al FR}}$

**Tasa de ataque no expuestos (TANE):** probabilidad de que una persona no expuesta al factor de riesgo (FR) desarrolle la enfermedad

$TANE = \frac{\text{Nº de personas no expuestas al FR que enferman}}{\text{Nº de personas no expuestas al FR}}$

La diferencia de estas tasas (o tasa de ataque diferencial) nos indicará para cada factor de riesgo (por ejemplo, el consumo de un alimento, en una toxiinfección alimentaria) cuanto más frecuente es que los expuestos a él enfermen que los que no lo están.

No obstante, puede ocurrir que sean varias las fuentes sospechosas y las tasas de ataque no indiquen claramente cuál de ellas es la implicada. Por ello, en la mayoría de las ocasiones se prefiere utilizar un método más completo y preciso como es el **estudio de casos y controles**.

Su diseño consiste en la selección de un grupo de personas que presentan la enfermedad investigada y que se denominan «CASOS», y de otro grupo de personas de iguales características a las de los casos pero que difieren de ellos en que no presentan la enfermedad y que se denominan «CONTROLES»; tanto en los casos como en los controles se investiga su historia de exposición a las fuentes sospechosas y los datos que se obtengan se ordenan en tablas de 2 x 2 EXPOSICION A UNA FUENTE ESPECIFICA

	Casos/Enfermos	Controles/Sanos
Expuestos	a	b
No expuestos	c	d

Una vez que se ha tabulado toda la información, se procederá al cálculo del Odds ratio (OR) de la siguiente manera:  $\text{Odds Ratio} = (a \times d) / (b \times c)$

El resultado se interpreta de la siguiente forma:

- Si la OR es igual a la unidad no hay asociación entre la exposición y la enfermedad.
- Si la OR es superior a 1, existe asociación.
- Si la OR es inferior a 1, no existe asociación, e incluso se podría interpretar que el factor de exposición actúa como factor protector ante la enfermedad.

Por último, siempre hay que calcular la significación estadística, para comprobar si tal asociación puede deberse al azar. Suele emplearse la prueba Chi-cuadrado de Mantel-Haenzel y calcular los IC al 95% para poder afirmar que la exposición se comporte como un factor de riesgo.

En caso de que haya más de una fuente asociada (p. e. dos alimentos) hay que hacer un análisis estratificado de los datos, u otros análisis más complejos como el multivariante.

Una vez finalizado este análisis se puede llegar a:

- Aceptar la hipótesis formulada y, por tanto, llevar a cabo las medidas de control.
- Rechazar la hipótesis actual por lo que se habrán de generar nuevas hipótesis.

## 6. Medidas de control

Como se dijo en el apartado 2, en muchas ocasiones las medidas de control se toman urgentemente con la información inicial existente sobre el brote.

Hay unas medidas standard de control que generalmente son relativamente fáciles de identificar. Es necesario valorar lo apropiado de cada medida para una situación particular e identificar aquéllas otras medidas que pudieran ser necesarias.

Medidas standard:

1. Disminuir la susceptibilidad o aumentar la resistencia de los susceptibles (p.e vacunas, etc.).
2. Control de la fuente (ej. eliminar la comida contaminada, potabilizar aguas, etc.)
3. Control del mecanismo de transmisión y vehículos contaminados (ej. alimentos, aire, etc)

Hablar de medidas específicas por tipo de brote es muy complejo por la variabilidad de las mismas. Cada medida adoptada debe ser valorada en función del conocimiento de la epidemiología de la enfermedad que se sospecha y fundamentado en torno a su efectividad, costo, tiempo de implantación y aceptación por parte de las personas implicadas.

**Resumen:** Se define como brote epidémico la aparición de un número de personas afectadas por una característica relacionada con la salud (sobre todo enfermedades), con una frecuencia claramente superior a

la esperada en condiciones normales, en un ámbito geográfico y período de tiempo determinados. Al estudiar un brote pretendemos conocer sus causas para evitar su extensión y agravamiento en la población mediante medidas de control. Un brote se estudia desde la descripción de los casos para determinar rápidamente qué se considera caso. Una vez determinados los mismo, mediante técnicas analíticas podremos determinar qué lo causa, siendo las dos más empleadas la tasa de ataque diferencial o el estudio de casos y controles.

### **Bibliografía:**

El presente documento es una adaptación y resumen de los siguientes documentos:

1. Protocolo de investigación de un brote epidémico. Sistema de Vigilancia Epidemiológica Andaluz. Grupo de mejora de comunicación de brotes nosocomiales. 2008. Disponible en: [http://www.juntadeandalucia.es/salud/export/sites/cs salud/galerias/documentos/p\\_4\\_p\\_1\\_vigilancia\\_de\\_la\\_salud/Protocolos\\_actuacion/protocolo\\_brote\\_epidemico.pdf](http://www.juntadeandalucia.es/salud/export/sites/cs salud/galerias/documentos/p_4_p_1_vigilancia_de_la_salud/Protocolos_actuacion/protocolo_brote_epidemico.pdf)
2. Organización Panamericana de la Salud. Investigación epidemiológica de campo: aplicación al estudio de brotes. [monografía en Internet]. Washington: Organización Panamericana de la Salud; 2002 [acceso 14 de Febrero de 2014]. Disponible en: [http://new.paho.org/col/index.php?option=com\\_docman&task=doc\\_download&gid=857&Itemid=](http://new.paho.org/col/index.php?option=com_docman&task=doc_download&gid=857&Itemid=).

Otros documentos de apoyo y consulta:

3. Centro Nacional de Epidemiología. Instituto de Salud Carlos III. Red Nacional de Vigilancia Epidemiológica. Protocolos de enfermedades de declaración obligatoria. Madrid, 2013. Disponible en: <http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-procedimientos/protocolos.shtml>
4. Cortés García M. Actividades de Inteligencia Epidemiológica. [monografía en internet]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad de España; 2012. Disp. en: <http://www.msc.es//profesionales/saludPublica/ccayes/inteligenciaepidemiologica/doc/ieCcaes2012.pdf>

## **CAPÍTULO 10**

### **Farmacoepidemiología. Farmacovigilancia**

**Autores:** Lucía Cea Soriano  
Paloma Ortega Molina  
Paloma Astasio Arbiza

## Objetivos

El presente capítulo aborda los siguientes apartados

- a) Introducción a la disciplina de farmacoepidemiología
- b) Fuentes de información y tipos de diseños epidemiológicos aplicados en la investigación farmacoepidemiológica
- c) Introducción a la disciplina de farmacovigilancia

## Introducción y Desarrollo

La farmacoepidemiología estudia el uso y efectos de los medicamentos, en grandes grupos poblacionales, por medio de la aplicación del método epidemiológico. Esta disciplina engloba dos ramas bien diferenciadas: la farmacología y la epidemiología. Por un lado, la farmacología se encarga de estudiar los efectos de los fármacos en los humanos (farmacocinética y farmacodinamia) por otro lado la epidemiología sirve de base para el estudio de la distribución de la enfermedad, cuantificación de la misma y factores asociados. La farmacoepidemiología ha utilizado la parte clínica de la farmacología en el estudio y el entendimiento de los efectos adversos producidos por los fármacos, valiéndose del método epidemiológico entendiéndose como el conocimiento de la distribución y los determinantes de las enfermedades en la población (1). Sin embargo, la farmacoepidemiología engloba otras áreas de conocimiento como son el estudio epidemiológico de la evolución natural de enfermedades en diferentes áreas terapéuticas y estudios de utilización de medicamentos. Es por tanto una herramienta de apoyo a las actividades farmacéuticas en diversas áreas, abarcando desde la fase de desarrollo de un medicamento (estudios experimentales: ensayo clínico) como del tiempo de vida comercial de un medicamento. Durante el desarrollo clínico, se evalúan por un lado los efectos beneficiosos y adversos de los fármacos en grupos de pacientes de número reducido y selectividad en cuanto a su perfil mórbido, durante un tiempo determinado. Tras su comercialización, el número de usuarios con diferentes características aumenta de forma considerable ampliando la indicación del medicamento objeto de estudio a grupos poblacionales no evaluados. Para detectar de forma precoz los efectos adversos de un medicamento, existe la notificación espontánea de sospecha de reacciones adversas a los medicamentos, sin embargo, ésta está sujeta al profesional sanitario y no permite cuantificar medidas de frecuencia, riesgo o impacto al no disponer de un denominador. En contraposición, los estudios epidemiológicos proporcionan estimadores sobre la frecuencia de la enfermedad, cuantificando también la frecuencia de efectos adversos en personas que utilizan un medicamento de interés y su comparativa con fármacos de la misma clase terapéutica o con aquellas personas no expuestas a ningún tratamiento, además de permitir estudiar subgrupos poblacionales de riesgo por medio del estudio de características del estado de salud y variaciones de riesgo. Por tanto, facilitan la planificación y desarrollo de otros fármacos, permiten un seguimiento de los

efectos de los medicamentos e impulsa la elaboración de directrices para el uso eficaz y seguro de los medicamentos (2).

Entre los estudios epidemiológicos que emplea la farmacoepidemiología, los estudios tradicionales se basaban en la entrevista personal por medio de encuestas, o de archivos de historias clínicas. Sin embargo, este tipo de recogida de datos tiene como consecuencia la insuficiente calidad de información en parte por la subjetividad y en ocasiones, la falta de información completa y detallada, alto coste económico y el prolongado tiempo empleado, sin olvidar destacar el limitado tamaño muestral, lo que llevaba a afectar a la potencia del estudio. Todas estas características, han restringido su utilización a la investigación farmacéutica de forma ocasional pero no así en estudios de asociaciones causales (2). La investigación farmacoepidemiológica precisa de poblaciones numerosas para poder evaluar de forma eficaz la aparición de efectos adversos, características del paciente y patrón de uso de medicamentos y en especial de información a largo plazo que permita estudiar la duración de tratamiento y apariciones de reacciones adversas retardadas. Es por todo ello, que las bases de datos sanitarias informatizadas se han convertido en la herramienta clave para la realización de estudios de farmacoepidemiología y han supuesto y suponen un avance en el conocimiento de la relación causal entre la exposición a un fármaco y un episodio no deseado determinado. Estas fuentes, recogen de forma rutinaria información individualizada sobre el uso de los servicios sanitarios tales como la prescripción de determinados fármacos, las visitas ambulatorias, los ingresos hospitalarios y las vacunaciones, entre otras informaciones, respetando en todo momento la confidencialidad de los datos de carácter personal por medio de la encriptación de la información de cada individuo. Esta recogida de datos suele ser llevada a cabo por mutuas de seguros médicos, sistemas sanitarios gubernamentales o los propios profesionales de la salud en su rutina de trabajo diaria, integrando en su sistema informático la información clínica; resolviendo, por tanto, el problema de la baja eficiencia de los métodos tradicionales descritos.

Existen dos modelos básicos de bases de datos sanitarias informatizadas: las *enlazadas* por un identificador (*record linkage databases*) y las *integradas*, en estas últimas la información es registrada por parte del personal sanitario, en especial, por los médicos de atención primaria con acceso al historial completo de sus pacientes. Las bases de datos enlazadas recogen la información referente a recetas de fármacos dispensados en la farmacia (incluyendo información detallada), información acerca del diagnóstico de ingreso y alta hospitalaria, diagnósticos realizados por los médicos especialistas, e información demográfica (edad, sexo, residencia, etc.). Cada individuo en la base de datos tiene adjudicado un código único y personal de identificación que permite fusionar la información procedente de diferentes fuentes sanitarias, permitiendo así generar una historia clínica de forma cronológica preservando la privacidad de cada sujeto incluido. Como ejemplo, destacan las creadas por los sistemas federales y estatales de los Estados Unidos, de asistencia sanitaria al anciano (Medicare) y al colectivo con bajo recurso económico (Medicaid) (2). En cuanto a las bases de datos integradas, en ellas se registra toda la información sanitaria utilizada por el médico de atención

primaria, por lo que este tipo de bases de datos sólo pueden desarrollarse en aquellos países en los que la cobertura sanitaria sea prácticamente universal, como es el caso de España, siendo el médico de atención primaria el principal responsable de los pacientes, y requieren de una adecuada informatización de los sistemas sanitarios. Entre alguno de los ejemplos destacan las bases de datos pioneras de Reino Unido, the Clinical Practice Research Datalink (CPRD) y *The Health Network Improvement (THIN)*. La base de datos PHARMO (Holanda) desde 1990, integrando un sistema de bases de datos relacionadas creado en las Universidades de Rotterdam y de Utrecht, entre las que se incluye la base de datos de medicamentos U-Expo, en la que se registra información detallada de más de 200 farmacias. (1,2). En los últimos años en España ha acontecido lo que puede considerarse el inicio de la investigación farmacoepidemiológica con bases de datos integradas: el proyecto BIFAP (*véase el capítulo de “Datos secundarios. Sistemas de información sanitaria “para más información”*)

A continuación, se repasan brevemente los principales diseños de estudio y estrategias que se pueden llevar a cabo dentro de la disciplina de la farmacoepidemiología. En cuanto a estudios de índole descriptiva, destacan los **estudios de utilización de medicamentos (EUM)**. Éstos proporcionan una información muy amplia referente a *patrones de uso*, incluyendo tendencias en la utilización y costes de tratamiento; a la *calidad de uso* por medio de la comparación con las guías clínicas y la realización de protocolos de actuación y otras recomendaciones; a *determinantes de uso* como parámetros sociodemográficos, y de las características del fármaco y prescriptor entre otros. Hay que tener en cuenta que con estos estudios, la relación temporal causa efecto no queda determinada (3,4). Con respecto a estudios de índole analítica, el **ensayo clínico aleatorizado (ECA)** constituye la mejor herramienta metodológica en la evaluación de la eficacia y la seguridad de nuevos fármacos. Según la fase de investigación en la que se encuentre el fármaco a evaluar, así como su finalidad (estudios de superioridad, estudios de equivalencia), el diseño a realizar requerirá de determinados criterios y consideraciones. Dentro de los estudios de post-autorización tanto los estudios de cohortes como los estudios caso control son los estudios más ampliamente utilizados (1). Los **estudios de caso control** se caracterizan por comparar un grupo de individuos con una determinada característica de interés (patología, intervención quirúrgica, tratamiento) con otro grupo sin esa misma característica, con el objetivo de buscar diferencias en exposiciones anteriores al evento de interés (5). Los **estudios de cohorte** tienen como objetivo identificar individuos de una población bajo una determinada exposición y seguirlos hasta la ocurrencia del evento de interés. La gran diferencia a nivel conceptual de este diseño frente al caso-control es la selección de los individuos reclutados en el estudio, éstos seleccionan a los individuos por una determinada exposición presente mientras que los estudios de caso control, por el evento a estudiar. Es por este motivo que los estudios de cohortes se encuentran, por lo general, libres del sesgo de selección. No obstante, este diseño también tiene limitaciones, pues requieren tamaños poblacionales grandes y tiempos de seguimiento muy prolongados, lo que los convierte, por lo general, en estudios menos eficientes. Por último, el diseño de **estudio**

**caso control anidado** constituye, a día de hoy, una de las mejores alternativas a la hora de elegir un diseño para un estudio de farmacoepidemiología. Tiene como peculiaridad ser un diseño que toma las ventajas de los estudios de cohortes y los estudios caso control, así en este diseño tanto los casos como los controles proceden de la misma población fuente con lo que se estudian a todos los casos identificados en esa población y una muestra aleatoria de controles realmente pertenecientes a esa población fuente (6). El hecho de que los controles constituyan una muestra aleatoria de la cohorte seleccionada está basado en el esquema de muestreo conocido como muestreo de densidad de incidencia, lo que permite la comparación de los casos con un subconjunto de los miembros de la cohorte que están en riesgo de convertirse en casos en el momento en que cada caso ocurre. Mediante esta técnica los sujetos que se convierten en casos posteriormente en el seguimiento son elegibles como controles a los efectos de su comparación con individuos que pasaron a ser casos más tempranamente, es así que bajo este modelo el OR es un estimador estadísticamente no sesgado del RR o razón de tasas (7)

La farmacoepidemiología se engloba a su vez en la farmacovigilancia, constituyendo la columna vertebral de ésta última. La farmacovigilancia es una actividad de salud pública cuyo objetivo es identificar, evaluar y prevenir los riesgos asociados a los medicamentos una vez comercializado. Los procesos que integra la farmacovigilancia pueden englobarse en 1) la identificación del riesgo, 2) el análisis de riesgos y 3) la gestión de riesgos. La identificación y el análisis de riesgos tiene que ver con la obtención y manejo de los datos, su cuantificación y la evaluación de su aceptabilidad social. La gestión de riesgos, lo cual implica toma de decisiones, abarca la adopción de medidas reguladoras para minimizar el riesgo, comunicación de riesgos a los profesionales y población y evaluación de la efectividad de dichas medidas. Por tanto, son acciones para acomodar el uso del medicamento a las condiciones en las que la evidencia científica indica que el balance beneficio-riesgo se mantiene favorables (8).

El Sistema Español de Farmacovigilancia de uso humano (SEFV-H) integra las actividades de las administraciones sanitarias para recoger y elaborar información sobre reacciones adversas a medicamentos. El SEFV-H, se organiza en Centros Autonómicos de Farmacovigilancia (CAFV) y, para el desarrollo permanente de la actividad de Farmacovigilancia, mediante un método técnico y científico. Todos los CAFV integran el Comité Técnico del Sistema de Farmacovigilancia El núcleo fundamental de trabajo del SEFV es la notificación de sospechas de reacciones adversas a través del Programa de notificación espontánea. Estas notificaciones adversas son enviadas por los profesionales sanitarios a través de unos formularios estandarizados de recogida de datos y en los Centros de Farmacovigilancia son evaluadas, codificadas y registradas en una base de datos común, denominada FEDRA (Farmacovigilancia Española Datos de Reacciones Adversas). La información contenida en FEDRA es evaluada periódicamente por los técnicos del SEFV con el fin de identificar de forma precoz posibles problemas de seguridad derivados del uso de los medicamentos (generación de señales de alerta). Dichas señales son discutidas en el Comité Técnico y trasladadas cuando se considere oportuno al Comité de Seguridad de Medicamentos de Uso Humano (CSMH) incluido en la división de Farmacovigilancia y

Farmacoepidemiología de la Agencia del Medicamento. Por último, ante la identificación de un problema de seguridad, la AEMPS solicitará un informe al Titular de la Autorización de Comercialización (TAC) sobre dicho problema y sobre el balance beneficio/riesgo del medicamento (8,9).**Resumen**

La farmacoepidemiología estudia el uso y efectos de los medicamentos, en grandes grupos poblacionales, por medio de la aplicación del método epidemiológico. La investigación farmacoepidemiológica precisa de poblaciones numerosas para poder evaluar de forma eficaz la aparición de efectos adversos, características del paciente y patrón de uso de medicamentos y en especial de información a largo plazo que permita estudiar la duración de tratamiento y apariciones de reacciones adversas retardadas. Estas cuestiones se resuelven por medio de la aplicación de los siguientes diseños: estudios de utilización de medicamentos (EUM), el ensayo clínico aleatorizado (ECA), Los estudios de caso control, los estudios de cohorte y los estudios de caso control anidado en una cohorte. Cabe destacar, cómo las bases de datos sanitarias informatizadas se han convertido en la herramienta clave para la realización de estudios de farmacoepidemiología y han supuesto y suponen un avance en el conocimiento de la relación causal entre la exposición a un fármaco y un episodio determinado. La farmacoepidemiología se engloba a su vez en la farmacovigilancia. Ésta constituye una actividad de salud pública cuyo objetivo es identificar, evaluar y prevenir los riesgos asociados a los medicamentos una vez comercializado. Los procesos que integra la farmacovigilancia pueden englobarse en 1) la identificación del riesgo, 2) el análisis de riesgos y 3) la gestión de riesgos, cuyo encargado es el Sistema Español de Farmacovigilancia de uso humano (SEFV-H) que integra las actividades de las administraciones sanitarias para recoger y elaborar información sobre reacciones adversas a medicamento

### **Bibliografía**

1. Strom B. What is Pharmacoepidemiology. En: Pharmacoepidemiology 3rd Edition. 2000. Publisher: John Wiley & Sons, Incorporated. Chichester. England p.3-16
2. Pérez-Gutthann S, García Rodríguez LA, de Abajo Iglesias FJ, Varas-Lorenzo C, Gutiérrez Cornejo L, Castellsague J. Farmacoepidemiología. La epidemiología aplicada a la seguridad y gestión de los medicamentos y al desarrollo farmacéutico. En: Tratado de Medicina Farmacéutica. Hernández Moreno, Zaragoza Porrás Editorial Médica Panamericana. Madrid 2010. p 649-660.
3. Laporte JR, Tognoni G. Principios de epidemiología del medicamento. 2ª Edición: Masson Salvat; Barcelona 1983
4. Figueras A, Vallano A, Narváez E. Fundamentos metodológicos de los estudios de utilización de medicamentos. Manual Práctico. 2003. Disponible [http://www.icf.uab.es/eums/manual\\_e.html](http://www.icf.uab.es/eums/manual_e.html). Fecha de acceso: Abril 2011.
5. Strom BL. Study Designs Available for Pharmacoepidemiology Studies. En: Brian L. Pharmacoepidemiology. Strom 3rd Edition. John Wiley & Sons, Incorporated. Chichester. England 2000, p 17-30.

6. González Pérez A, García Rodríguez LA. Estudios de cohortes y de casos y controles qué podemos esperar de ellos. GH Continuada. 2003; 2(1) 44:48. Disponible en <http://213.4.18.135:48080/98.pdf>.
7. Szklo M, Nieto FJ. Epidemiología Intermedia. Conceptos y aplicaciones. Ediciones Díaz de Santos, Madrid 2003. p 109-151
8. Los primeros 25 años del sistema español de farmacovigilancia de Medicamentos de Uso Humano. Gloria Manso, Agustín Hidalgo, Alfonso Carvajal, Francisco José de Abajo.  
<https://www.unioviedo.es/gaife/documentos/libro25aniversario/libro.pdf>
9. Abajo Iglesias FJ, Madurga Sanz M, Rodríguez Pascual, Martín Serrano-García G, Maciá Martínez MA. La Farmacovigilancia en España. En: Tratado de Medicina Farmacéutica. Hernández Moreno, Zaragoza Porras Editorial Médica Panamericana. Madrid 2010. p 605-619.