

2.9. Farmacoepidemiología y estudios de utilización de medicamentos

J. ALTIMIRAS

J. BAUTISTA

F. PUIGVENTÓS

1 INTRODUCCIÓN

La toma de decisiones para conseguir un uso eficiente de los recursos terapéuticos, requiere disponer de información y aplicar unos conocimientos técnicos. Los conceptos y las herramientas metodológicas provenientes del campo de la epidemiología han mostrado su gran utilidad, de forma que en los últimos años se han incorporado de una forma masiva a disciplinas como la farmacoeconomía, la evaluación de tecnologías médicas o la medicina basada en la evidencia.

En el campo del medicamento, la aplicación de métodos epidemiológicos/poblacionales ofrece una alternativa para aumentar el grado de información disponible. Por una parte se trata de integrar al máximo la información proveniente de la experiencia de uso cuando los fármacos se utilizan en condiciones habituales de la práctica clínica, y por otra parte conocer los determinantes de su utilización.

La Epidemiología Clínica, que Spitzer define como “el estudio de los determinantes y de los efectos de las decisiones clínicas” consiste en una forma de aplicar el método epidemiológico a la po-

blación “pacientes” y centrar el punto de observación en la atención clínica⁽¹⁾.

También para dar respuesta a cuestiones muy similares, la Farmacoepidemiología centra la observación en los fármacos, aplicando el método epidemiológico a la población en general o a la población “pacientes”⁽²⁾. Porta y Hartzema han definido esta disciplina como “La aplicación de los conocimientos, métodos y razonamiento de la epidemiología al estudio de los efectos –positivos y negativos– y usos de los fármacos en grupos de poblaciones”⁽³⁾.

Para el farmacéutico de hospital, los conocimientos de farmacoepidemiología son necesarios para ayudar en la interpretación crítica de la literatura científica sobre fármacos y a su conversión en parámetros aplicables a la toma de decisiones. Los criterios y métodos de la farmacoepidemiología son útiles para desarrollar sus actividades en áreas como la evaluación de medicamentos, la farmacovigilancia y en el diseño de estudios y programas.

El objetivo de este capítulo es proporcionar unos conceptos básicos de epidemiología, como un instrumento necesario para luego pasar revisión al abanico a sus aplicaciones dirigidas al estudio de los efectos y usos

de los fármacos. En este sentido se hará especial énfasis en los estudios de utilización de medicamentos, por la relevancia práctica que tienen para los Servicios de Farmacia de hospital.

2 CONCEPTOS DE EPIDEMIOLOGÍA

Tradicionalmente se ha definido la Epidemiología como el estudio de la distribución y determinantes de la salud y de las enfermedades en las poblaciones humanas. En la actualidad, este concepto se ha ampliado también al estudio de la intervención de los propios servicios sanitarios. Así mismo, el cambio en las patologías más prevalentes y en las características demográficas de muchas poblaciones, así como los avances que se han producido en el campo biomédico han contribuido a cambiar los centros de interés y a evolucionar los métodos propios de esta disciplina.

2.1. Objetivos de la epidemiología

En resumen, los objetivos básicos de la epidemiología actual son:

1. El análisis de la información sobre el estado de salud de la población (entendiendo el concepto estado de salud en una forma amplia), su distribución y tendencias, para que sea útil en la planificación y gestión de los programas de salud.
2. El análisis de la causalidad, con el fin de hacer posibles las intervenciones de prevención, control y/o erradicación.
3. La evaluación de las intervenciones, cuyo fin es el de mejorar el estado de salud de la comunidad.

En base a estos tres objetivos, se acostumbra a dividir la epidemiología en:

1. Epidemiología descriptiva.
2. Epidemiología analítica.
3. Epidemiología evaluativa o de intervención.

2.2. Método

Etimológicamente el vocablo epidemiología proviene del griego “epi” (acerca de) y “demos” (pueblo). La epidemiología basa su método en razonar en términos de grupos, no de individuos y, en función de ello y de la aceptada causalidad multifactorial, acostumbra a trabajar en términos de proba-

bilidad. El método epidemiológico está basado en la observación sistemática y rigurosa de los hechos que ocurren de forma natural.

De hecho, el desarrollo del método epidemiológico ha surgido de la necesidad de buscar métodos de investigación que soslayan las limitaciones de los estudios experimentales. No obstante, actualmente la frontera entre los clásicos estudios observacionales utilizados en epidemiología y los estudios experimentales es menos definida.

2.2.1. Medición

En general, tanto si se trata de una variable continua como de una variable discreta, en la mayoría de casos interesará disponer de información de tipo cuantitativo. Un suceso se puede expresar mediante un número, una proporción, una tasa o una razón.

Una proporción es una fracción en la que todos los elementos del numerador están incluidos en el denominador; es adimensional y su rango va de 0 a 1, aunque también se pueden expresar en relación a 100, 1.000, etc.

Una tasa es una fracción que expresa el cambio de una magnitud en relación al cambio de otra magnitud, generalmente el tiempo. Un ejemplo clásico de una tasa es la velocidad instantánea. Es siempre dimensional y lleva incorporado estructuralmente el tiempo como $1/t$; su rango va de menos infinito a más infinito.

Una razón es una fracción en la que al menos no todo el numerador está incluido en el denominador. Una razón puede ser dimensional o adimensional.

La Odds es una razón adimensional de probabilidades y expresa la probabilidad de que ocurra un suceso en relación a la probabilidad de que no ocurra ($p/1-p$). Esta forma de expresión de una medida proviene del argot de las apuestas en el mundo anglosajón.

Por ejemplo, podemos decir que en nuestro hospital se atendieron, en el Servicio de Urgencias, 20 casos de intoxicación por el fármaco A (número). O podemos decir que el 10% de las intoxicados atendidos en el año 2001 en nuestro hospital, lo fueron a causa de A (proporción). También, si dispusiéramos de suficiente información, podríamos calcular el número de personas intoxicadas por cada 100 personas expuestas a A en un año (tasa).

a) Medidas de frecuencia

Las medidas básicas de frecuencia⁽⁴⁾ en epidemiología son las que hacen referencia a morbilidad y mortalidad. En lo referente a las medidas de morbilidad, éstas siempre están asociadas a los conceptos de incidencia y prevalencia.

La incidencia hace referencia a los nuevos casos, es decir, a los individuos que pasan de ser sanos a enfermos. La prevalencia se refiere a los casos presentes, es decir a los individuos enfermos en un momento dado. La incidencia, por lo tanto, nos da una idea de la velocidad o fuerza con la que se propaga la enfermedad en un determinado colectivo, mientras que la prevalencia nos da idea del peso de la enfermedad en el colectivo. Estos dos conceptos y su relación dinámica han sido expresados muy gráficamente por Barker y Rose⁽⁵⁾:

Matemáticamente la prevalencia (P) se expresa como la proporción de individuos que tienen la enfermedad en un momento determinado.

$$P = n^{\circ} \text{ de casos existentes} / \text{población total}$$

El hecho de que se haga referencia a un “momento dado”, no significa que éste tenga que ser necesariamente un “instante”; puede ser un periodo de tiempo relativamente largo, pero medida en un sentido estático, es decir se estudia el fenómeno en un periodo de tiempo, no sus cambios en relación al transcurso del tiempo. Aunque habitual-

mente nos referiremos a enfermedades, obviamente cualquiera de los conceptos que se irán tratando pueden ser referidos a otros acontecimientos relacionados con la salud. Por ejemplo, si consideramos como hecho observable la prescripción de medicamentos y consideramos la “población” de prescripciones de los médicos de un área determinada, podemos hablar de la “prevalencia de prescripción” de antibióticos en esta área.

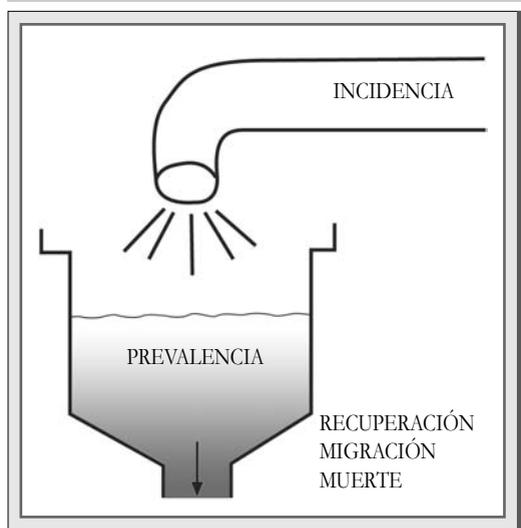
La expresión numérica de la incidencia toma dos formas distintas. La incidencia acumulada (IA) mide la proporción de individuos que pasan de sanos a enfermos durante un periodo determinado; es por lo tanto una proporción referida a un periodo de tiempo y nos da una estimación del valor promedio de la velocidad con la que se produce una enfermedad o suceso. Se expresa:

$$IA = n^{\circ} \text{ de nuevos casos} / n^{\circ} \text{ de individuos a riesgo al inicio del periodo}$$

La incidencia acumulada expresa la probabilidad o riesgo individual de aparición de la enfermedad, condicionada a que no se sufra la enfermedad al inicio o de que no se muera por otras causas.

La densidad de incidencia (DI) o “tasa de incidencia” mide la fuerza instantánea de ocurrencia de una enfermedad, considerando el número de casos nuevos en un periodo determinado en relación al tiempo que cada individuo estuvo a riesgo de contraer la enfermedad en el citado periodo. En este caso, el denominador se expresa por la magnitud “personas-tiempo” a riesgo. Por ejemplo, 100 personas-año a riesgo de intoxicación por digoxina, pueden representar a 100 personas que estuvieron expuestas a la digoxina durante un año, a 50 personas que estuvieron expuestas durante dos años o a 1 persona que estuvo expuesta 100 años al fármaco.

Figura 1. Relación entre incidencia y prevalencia: (Modificado de: Barker, D.J.P. y Rose G.A.).



Relación entre incidencia y prevalencia

Asumiendo que la incidencia y la duración media de la enfermedad son constantes durante el periodo de estudio, la prevalencia de una enfermedad es aproximadamente igual a la incidencia por la duración media de la misma.

$$P = I \times D$$

Veamos ahora con un ejemplo los conceptos de IA, DI y P:

Supongamos que realizamos un estudio de 5 años en el que seguimos a 10 individuos expuestos a un fármaco

co y en los que estudiamos la aparición de toxicidad renal (Esquema 1). La “I” -significa la aparición de dicha toxicidad, la “MI” la muerte por esta causa y la “M” la muerte por otras causas. La línea discontinua representa el periodo de seguimiento en el que cada individuo está a riesgo y la línea de puntos el periodo en el que los individuos no están a riesgo, sea porque ha sufrido toxicidad, porque se ha muerto o porque no ha entrado aún en el estudio.

La IA sería en este caso $IA = 5/10 = 50$ intoxicados por 100 individuos expuestos en 5 años, es decir 10 intoxicados por 100 individuos expuestos, por año.

Las personas-año expuestas serían $5 + 2 + 2,5 + 4 + 4 + 3 + 1,5 + 5 + 1 + 2,5 = 30,5$ personas-año, la DI sería

$$DI = 5 / 30,5 = 0,164 = 16,4 \text{ casos por 100 personas-año.}$$

En sentido estricto, como en el cálculo de la IA se considera como denominador el número de personas a riesgo al inicio del estudio, el denominador en nuestro ejemplo debería ser 8. En un estudio con poblaciones reales lo haríamos así. Pero en nuestro caso se trata de un ejemplo irreal y consideramos la población total de estudio. En el caso de estudios con grandes números de individuos y

una cierta movilidad poblacional, se puede elegir como denominador la población a mitad de periodo de estudio o la media de la población al inicio y al final del estudio.

En este mismo caso la prevalencia de intoxicaciones a mitad del periodo de estudio (año 2,5) sería

$$P = 2 / 10 = 20\%$$

a) Medidas de frecuencia de mortalidad

Aunque en realidad no lo son, se acostumbra a denominar a estas medidas como tasas de mortalidad. En efecto, se calculan como la incidencia acumulada:

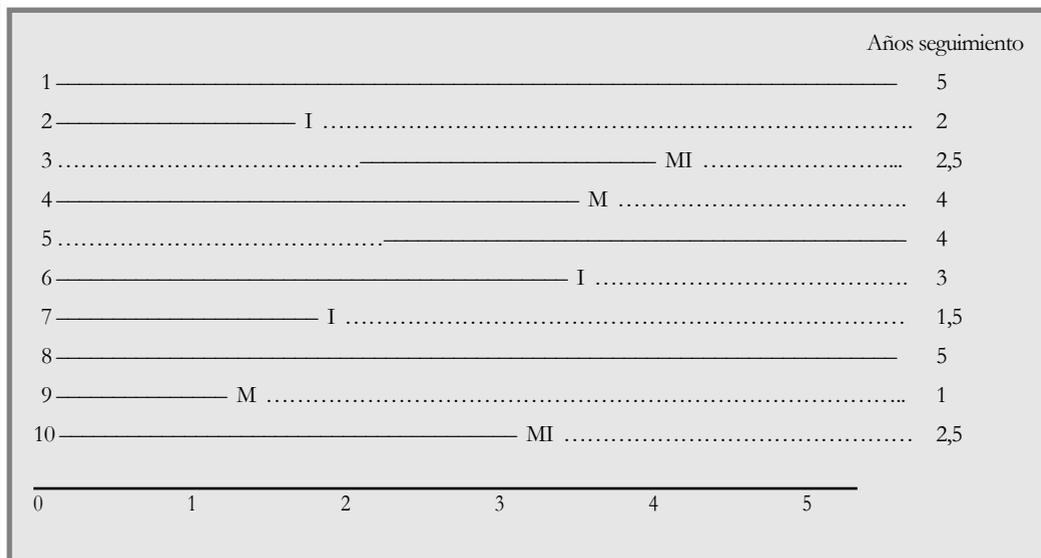
$$\text{Tasa de mortalidad} = \frac{\text{n}^\circ \text{ de muertos en el periodo}}{\text{n}^\circ \text{ de personas vivas al inicio}}$$

Evidentemente, también se emplean tasas de mortalidad específicas por grupos de edad, sexo, y otras como la tasa de mortalidad infantil, la tasa de mortalidad perinatal, etc.

b) Medidas de asociación o comparación

En general, uno de los intereses de las medidas de frecuencia es el poder compararlas. Las medidas que a continuación se citan son las que habitualmente se emplean en epidemiología y en las que se

Esquema 1. Funciones farmacéuticas clínicas.



encuentran expresados los resultados de la mayoría de estudios epidemiológicos.

Riesgo relativo:

Es la razón entre dos tasas de incidencia. Estrictamente, cuando se comparan dos incidencias acumuladas se la denomina razón de riesgos, aunque en la práctica en cualquiera, de los dos casos se acostumbra a expresar como riesgo relativo. El riesgo relativo (RR) es adimensional y su rango es de 0 a más infinito.

Generalmente se utiliza esta medida para comparar las incidencias de una determinada enfermedad en los individuos expuestos; en los no expuestos a un determinado factor de riesgo.

$$RR = DI \text{ en expuestos} / DI \text{ en no expuestos}$$

Expresa la fuerza de la asociación entre la exposición al factor y la aparición de la enfermedad. Tiene su utilidad en la investigación etiológica.

Si $RR > 1$ El factor está asociado positivamente con la enfermedad.

Si $RR = 1$ No existe asociación.

Si $RR < 1$ El factor está asociado inversamente a la enfermedad, es decir dicho factor “protege” contra la enfermedad.

Por ejemplo, si al comparar las incidencias de agranulocitosis entre los enfermos expuestos al fármaco A y los no expuestos es de 4, podemos decir que el estar expuesto al fármaco A supone un riesgo cuatro veces mayor de padecer agranulocitosis. El RR puede transformarse en una medida de exceso de riesgo (en %).

$$\text{Exceso de riesgo} = (RR - 1) \times 100$$

En nuestro caso, en el que $RR = 4$, $(RR - 1) \times 100 = (4 - 1) \times 100 = 300\%$ de exceso de riesgo. Al tratar los estudios de cohortes se profundizará sobre el sentido e interpretación de esta medida.

Riesgo atribuible (RA)

Es la diferencia entre dos tasas de incidencia o IA's. Es dimensional y su rango es de menos a más infinito. Mide el efecto absoluto del factor de riesgo en la incidencia de la enfermedad. Es útil como medida de la magnitud (impacto) que el efecto de un factor de riesgo tiene sobre la salud de una población. Tiene un interés eminente en salud pública y, especialmente es útil en la

toma de decisiones. En algunas publicaciones el Riesgo Atribuible recibe también el nombre de Reducción. Absoluta del Riesgo (RAR) o el de Diferencias de Riesgos.

Veamos ahora un ejemplo de interpretación del RR y del RA. Supongamos que la incidencia anual de una enfermedad en la población expuesta a un determinado factor de riesgo es de 10 por 1.000 personas expuestas, mientras que en los no expuestos es de 5 por 1.000 personas.

$$\text{El RR sería } 10/5 = 2$$

$$\text{El RA sería } 10 - 5 = 5 \text{ por 1.000 personas expuestas.}$$

Supongamos ahora que la incidencia anual de otra enfermedad en los expuestos a otro factor de riesgo es de 100 por 1.000, personas expuestas, mientras que en las no expuestas es de 50 por 1.000 personas. En este caso, el RR y el RA serían respectivamente:

$$RR = 100/50 = 2$$

$$RA = 100 - 50 = 50 \text{ por 1.000 personas expuestas}$$

(recuérdese que el RA al ser una diferencia conserva las unidades de los valores de las incidencias).

En ambos casos la fuerza de la relación causal entre los factores de riesgo y las dos enfermedades es la misma, pero en el segundo caso el impacto en la población es mucho más relevante. Si, conseguimos evitar el segundo factor de sesgo x riesgo, conseguiremos un impacto mucho mayor en la salud de la población.

En ocasiones, el considerar los resultados de un estudio desde la perspectiva del riesgo relativo o del riesgo atribuible puede influir en la interpretación práctica de los resultados. Así por ejemplo, en el “Helsinki Heart Study” el porcentaje de muertes por enfermedad cardíaca isquémica fue, a los 5 años, de 4,1% en el grupo placebo y de 2,7% en el grupo tratado con gemfibrozil. El decir que el gemfibrozil produjo una reducción del 34% (obtenido al aplicar la fórmula del exceso de riesgo) en la mortalidad por enfermedad isquémica es distinto que decir que al fármaco se le puede atribuir una reducción de 1,4 muertes por cada 100 pacientes tratados. O, dicho de otra forma, que de cada 1.000 personas tratadas; 14 se beneficiarían, 27 morirían por enfermedad isquémica y 959 tampoco se beneficiarían porque tampoco morirían aunque no recibieran el tratamiento^(6,7). Para ello, y aplicando el mismo concepto a la expresión

de los resultados en los estudios de eficacia de los fármacos se aplica el concepto de “número necesario de pacientes a tratar (NNT)”, que se considera una herramienta sencilla y fiable que permite al clínico tener una interpretación intuitiva y comparable del resultado de un estudio⁽¹²⁻¹³⁾.

El NNT es el inverso de la reducción del riesgo absoluto (1/RRA) y se puede calcular a partir de los resultados de eficacia. Volveremos sobre este concepto más adelante.

$$\text{NNT} = 1 / \text{Incidencia en G. control} - \text{Incidencia en G. tratado}$$

Razón de Odds (RO)

Es una medida de comparación entre las exposiciones a un determinado factor de riesgo en los individuos afectados de una determinada enfermedad (casos) y en los sanos (controles), calculado mediante la razón de las Odds de exposición en ambos grupos. La RO es una estimación del RR cuando no se disponen de datos de incidencia y es, por ello, la medida de asociación utilizada en los estudios de casos y controles. Es adimensional y su rango es de 0 a más infinito. La forma de calcularla así como su interpretación se tratarán extensamente cuando se haga referencia a los estudios de casos y controles.

c) Medidas de impacto

Fracción etiológica (FE)

Expresa la proporción de casos atribuibles a un determinado factor de riesgo en la población expuesta. Viene a representar la probabilidad de que un caso sea debido a la exposición al factor de riesgo. Se puede calcular a partir del valor del RR:

$$\text{FE} = [(\text{RR} - 1) / \text{RR}] \times 100$$

d) Estandarización o ajuste

Cuando comparamos dos medidas de frecuencia de morbilidad, mortalidad u otro suceso en dos poblaciones, pueden existir algunos atributos que se distribuyan de forma distinta en las dos poblaciones y que induzcan a una interpretación incorrecta de la comparación. Por ejemplo, si estamos comparando la mortalidad en dos poblaciones, puede ser que las mismas tengan distinta composición en cuanto a la edad, al sexo o a la clase social y que estos atributos, por sí solos, puedan influir directamente en los valores de la mortalidad en las dos poblaciones, con lo que su compa-

ración directa estará sesgada.

En este caso, para evitar la comparación incorrecta de las dos “tasas crudas” de mortalidad en las dos poblaciones, se analizan las tasas en ambas subdivididas en grupos o estratos. Por ejemplo, agrupando a los individuos en grupos de edad definidos. A partir de estos datos se pueden obtener las tasas de mortalidad global en las dos poblaciones ajustadas por la edad y efectuar de una forma más correcta la comparación. Existen dos métodos de estandarizar o ajustar tasas, el método directo y el método indirecto, cada uno de ellos con su utilidad y sus limitaciones. El lector interesado en conocer estos dos métodos puede encontrar su descripción en los libros de Rothman o de Kramer, citados en la bibliografía.

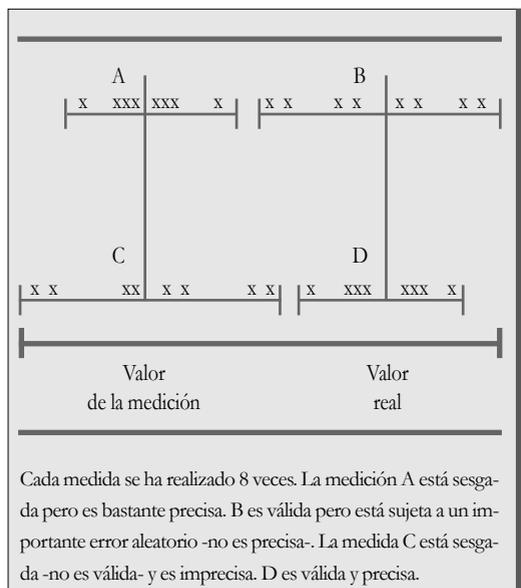
e) Calidad de las medidas

Cuando se realizan mediciones en biomedicina[®], las cuales van a ser la principal fuente de información en los estudios epidemiológicos, existen fuentes de variabilidad fruto del acto de medir. Es esencial conocerlas y evitarlas cuando ello sea posible, para asegurar la calidad de la información utilizada y evitar una interpretación distorsionada de la misma. Las dos fuentes principales de variabilidad son la variabilidad biológica y el error de medida. La variabilidad biológica refleja la naturaleza dinámica de las variables biológicas y se manifiesta en la propia variabilidad interindividual, así como en la intraindividual. El error de medida siempre presente en cualquier medición, se subdivide en el error aleatorio y el error sistemático o sesgo. El error aleatorio es aquél que no tiene dirección, es fruto del azar y tiende a reducirse por compensación en sucesivas mediciones. El error sistemático o sesgo siempre se produce en la misma dirección, es fruto del instrumento de medida o del individuo que la realiza, y no se reduce por repetición. La Figura 2 expresa gráficamente la relación entre dichos errores.

Las cualidades de una medición están en relación con los errores de la misma. Se dice que una medida es precisa o repetible cuando los resultados de sucesivas mediciones están muy próximos entre sí. La precisión es, por tanto, la ausencia de error aleatorio. Una medida es válida, en la medida en que sus resultados se acercan al valor real. La validez de una medición está asociada a la ausencia de sesgos.

La validez y la precisión no tienen por que estar directamente relacionadas: una medición puede ser muy precisa y no ser válida o viceversa. Supongamos que

Figura 2. Relación entre error aleatorio y sistemático



estamos midiendo la presión arterial de un sujeto con un esfigmomanómetro y que realizamos mediciones sucesivas. La validez de este método de medición vendría determinado por la proximidad de los resultados con la medición realizada mediante un catéter intraarterial. Mientras que la precisión vendría determinada por la proximidad de los resultados de las distintas mediciones.

No se debe confundir el significado que toma aquí la palabra precisión, con otro de los posibles significados de la misma y que alude al “detalle” con el que se realiza la medición, por ejemplo al número de decimales con el que se expresa la medición, o sea la cantidad de información proporcionada y que es otra cualidad de las mediciones.

f) Sensibilidad y especificidad

En pruebas diagnósticas en las que habitualmente se manejan variables categóricas dicótomas, enfermo o sano, etc, la validez se conceptualiza en los términos sensibilidad y especificidad. La sensibilidad de una prueba es la capacidad de identificar todos los casos positivos, mientras que la especificidad es la capacidad de identificar como positivos sólo a los casos positivos. Los índices de sensibilidad y especificidad son fácilmente calculables si distribuimos los resultados de nuestro test y los datos reales o de un método estándar o patrón oro en una tabla de contingencia 2 x 2.

		Metodo estándar o patrón oro	
		Positivo	Negativo
Test que se evalúa	Positivo	a (verdadero +)	b (falso +)
	Negativo	c (falso -)	d (verdadero +)

Cálculos:

Sensibilidad = $a/(a+c)$, es decir verdaderos positivos del test / total de positivos del patrón oro.

Especificidad = $d/(b+d)$, es decir verdaderos negativos del test /total de negativos del patrón oro.

Valor predictivo positivo (VPP): $a/(a+b)$, es decir verdaderos positivos/total de positivos identificados por el test.

Valor predictivo negativo (VPN): $d/(c+d)$, es decir verdaderos negativos/total de negativos identificados por el test.

Veamos esto con un ejemplo. Supongamos que estamos midiendo el cumplimiento de unos enfermos hipertensos mediante el recuento de pastillas, sirviendo como método standard (patrón oro) la obtención de niveles plasmáticos del fármaco.

Recuento de pastillas (test que se evalúa)		Niveles plasmáticos del fármaco (patrón oro)		
		Cumplidores	No cumplidores	
Cumplidores	20	4		
No cumplidores	10	66		
	30	70	100	

Los 20 efectivos de la primera casilla representan los verdaderos positivos –los cumplidores verdaderos, identificados por el test–, los 4 efectivos de la segunda casilla son los falsos positivos, los 10 de la tercera los falsos negativos y los 66 de la cuarta son los verdaderos negativos.

La sensibilidad y especificidad serían respectivamente:

$$s = 20/(10+20) = 66,6\%; e = 66/(4 + 66) = 94,3\%$$

Por lo que podemos concluir que el sistema de detección de cumplidores por recuento de pastillas es bastante específico (identifica muchos no cumplidores como no cumplidores, o lo que es lo mismo pocos cumplidores como no cumplidores) y moderadamente sensible (hay muchos cumplidores que no son identificados por este método).

Estos índices no corresponden a la clásica forma de razonamiento diagnóstico y por ello en clínica puede ser más útil presentar los resultados en términos del valor predictivo del test. Si estamos identificando los cumplidores, el valor predictivo positivo (VPP) es definido como la proporción de sujetos cumplidores identificados por el test que realmente son cumplidores. El valor predictivo negativo (VPN) se define como la proporción de sujetos no cumplidores identificados por el test que son realmente no cumplidores.

En el ejemplo anterior VPP y VPN serían:

$$\text{VPP} = 20 / (20 + 4) = 83,3\%$$

y

$$\text{VPN} = 66 / (10 + 66) = 86,8\%$$

Mientras la sensibilidad y especificidad dependen sólo del test, los valores de VPP y de VPN dependen también de la prevalencia de lo que queremos determinar (el cumplimiento) en la muestra de pacientes que estudiamos. En el ejemplo anterior si multiplicamos por 10 el valor de la primera columna ($a=200$ y $c=100$), y repetimos los cálculos, nos encontramos que la sensibilidad y especificidad del test serán los mismos, en cambio el VPP = 98,0% y el VPN = 39,7%.

2.2.2. Tipos de diseño de estudios

Los elementos básicos que configuran los distintos tipos de diseños de estudios en biomedicina⁹, son la asignación controlada del factor de estudio por el investigador y la asignación de los individuos a uno de los grupos de comparación, la existencia o no de secuencia temporal y la direccionalidad del estudio, en caso de que ésta exista. En la Tabla 1 se clasifican los diseños básicos conforme a estos elementos.

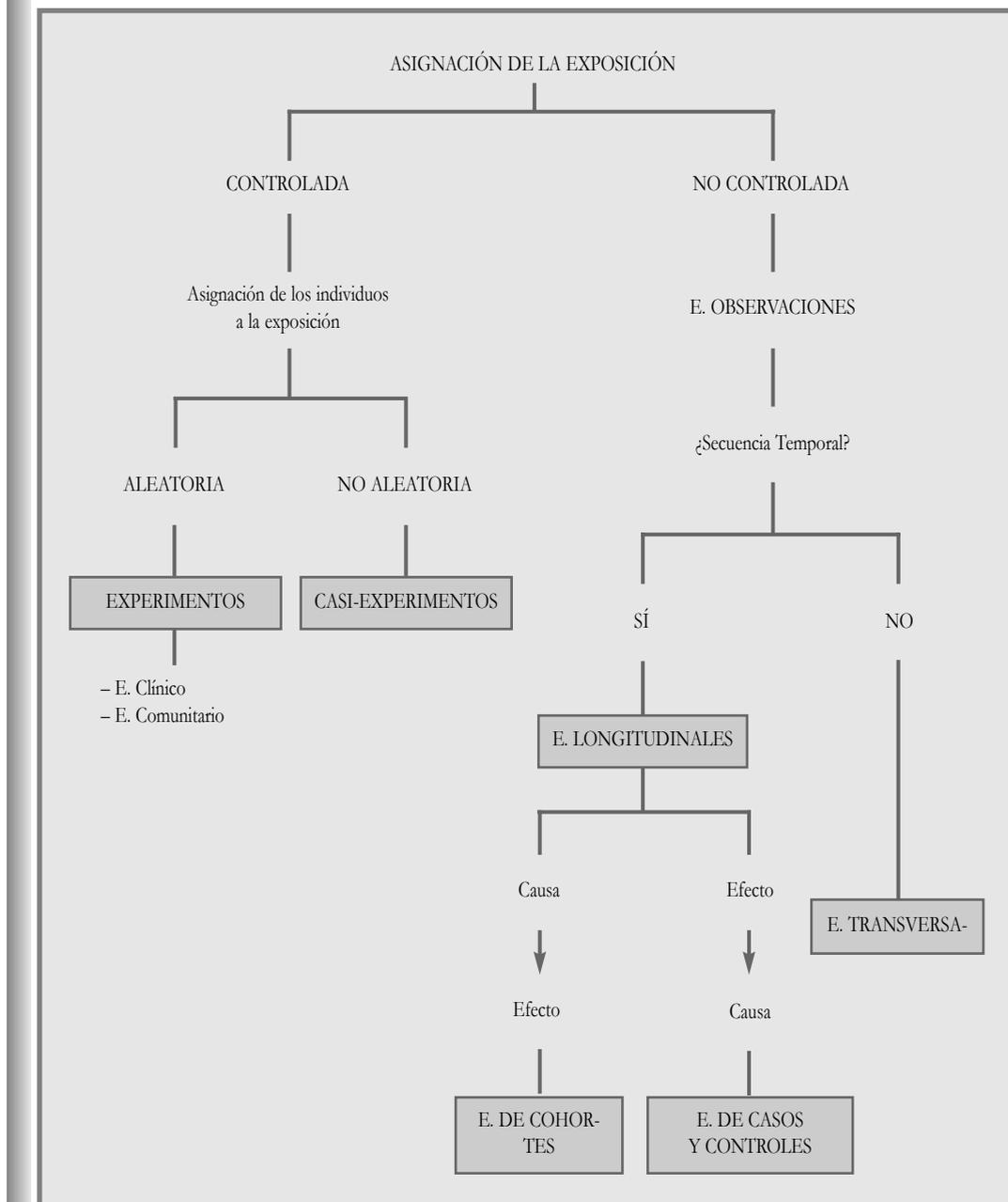
Si la asignación es controlada por el investigador nos encontramos ante los clásicos estudios experimentales, que representan la aplicación a la investigación humana del clásico experimento de laboratorio. Pero ello no siempre es posible, no podemos por ejemplo asignar un hábito tóxico a un grupo de individuos y evitarlo en el grupo control. En estos casos lo que es factible es la observación sistemática del “laboratorio natural” en el que la asignación del factor de estudio se produce de forma natural, sin control por el investigador y, en este caso, estamos ante los estudios observacionales.

Los estudios experimentales, en los que la asignación de los individuos es controlada por el investigador, son el ensayo clínico –cuando los individuos son pacientes– y el ensayo comunitario –cuando los individuos son sanos–. Cuando en un experimento la asignación controlada de los sujetos no es posible o pertinente nos encontramos ante un estudio casi-experimental.

Un ejemplo de estudio casi-experimental podría consistir en una intervención educativa de pacientes hospitalizados en los que no es posible realizar una asignación al azar por la posibilidad de “contaminación”, en el sentido de que los individuos educados pueden transmitir total o parcialmente la educación a otros individuos ingresados próximos a ellos. Otro ejemplo de estudio casi-experimental es el experimento en el que se estudió la salud dental de dos poblaciones, una que recibía agua fluorada y otra que no. En los casos como el citado, en los que el estudio se realiza en comunidades amplias, los estudios casi experimentales se acostumbra a denominar ensayos comunitarios de intervención.

Un estudio observacional en el que no existe secuencia temporal, es decir, en el que la identificación del suceso y de los factores de estudio se realiza en el mismo momento es un estudio transversal. Cuando existe secuencia temporal, los estudios se denominan estudios longitudinales. Si en un estudio longitudinal se parte de la identificación de existencia o no de la posible causa (por ejemplo, el uso de un determinado fármaco) y se investiga la posterior aparición de un efecto (por ejemplo, una reacción adversa) estamos ante un estudio de cohortes. Cuando por el contrario partimos de la identificación del efecto (en nuestro anterior ejemplo, los individuos que sufren una determinada entidad patológica compatible con una reacción adversa) y se investiga retrospectivamente la exposición a determinados factores posibles de riesgo (en nuestro caso, fármacos) nos encontramos ante un estudio de casos y controles.

Como decíamos anteriormente, la frontera entre la investigación clásicamente epidemiológica y las investigaciones experimentales en ciencias de la salud, en lo referente al tipo de diseño de estudio, es cada vez menos definida. De hecho, cuando es posible y pertinente, la epidemiología utiliza el estudio experimental. Sin embargo, en este capítulo no trataremos este tipo de estudios ya que existe un capítulo de ensayos clínicos.

Tabla 1. Modificado de Liliefeld AM y Lilienfeld DE⁽¹⁰⁾.

a) Estudios transversales

Un estudio transversal es aquél en el que se observa, en un momento dado, la presencia de la enfermedad o cualquier otro suceso, así como la presencia de determinados factores que "a priori" se consideran posiblemente asociados al suceso estudiado. La medida de fre-

cuencia que se utiliza es, evidentemente, la prevalencia; por ello, a los estudios transversales también se les conoce como estudios de prevalencia.

Una encuesta epidemiológica para conocer en un momento dado cuantos enfermos de SIDA existen en una población y la distribución de esta enfermedad en

distintos grupos de riesgo es un ejemplo de estudio transversal.

Métodos básicos

La selección de los sujetos en este tipo de estudios se realiza generalmente por muestreo aleatorio aunque, en ocasiones, el objetivo del estudio puede aconsejar la selección por el suceso o por algún factor de estudio, cuando éstos son poco frecuentes.

Se estudia la presencia de la enfermedad y se analizan los resultados subdividiendo los sujetos seleccionados en función de los factores estudiados y, de esta forma, pueden compararse las prevalencias en cada subgrupo. La Figura 3, representa en esquema este tipo de diseño.

Las prevalencias en este caso para los expuestos y no expuestos serían respectivamente:

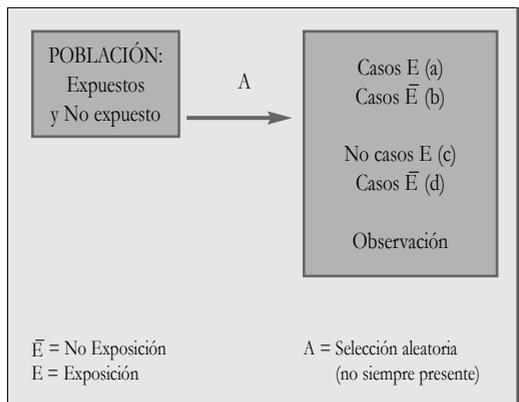
$$PE = a/(a+b) \text{ y } PNE = c/(c+d)$$

A partir de ellas se podría calcular la Razón de Prevalencias (RP) = PE / PNE.

La ambigüedad temporal es una de las limitaciones más importantes de los estudios transversales, en el sentido de que al no conocer la secuencia temporal entre posible causa y efecto, difícilmente pueden dar lugar, a explicaciones etiológicas. En algunos casos como por ejemplo el sexo y la herencia genética, es evidente que si se puede atribuir una secuencia temporal. En enfermedades de corta duración, especialmente cuando existe una distribución temporal irregular, los estudios transversales pueden dar lugar a estimaciones erróneas.

En resumen, los estudios transversales tienen una

Figura 3. Representación esquemática de los estudios transversales.



función generalmente descriptiva, no son válidos para contrastar hipótesis etiológicas, pero sí para sugerirlas, función en la que su relativa rapidez y su bajo costo los hacen muy útiles y efectivos.

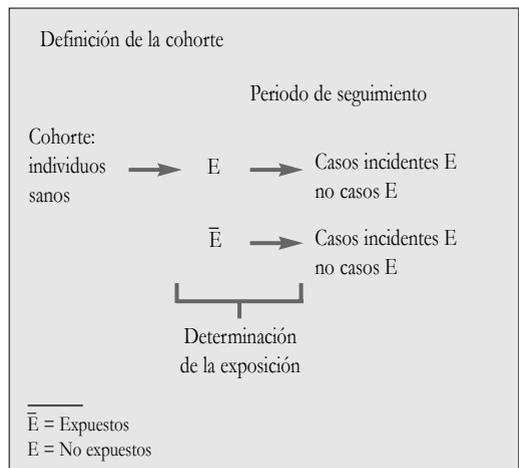
b) Estudios de Cohortes

El estudio de cohortes es un diseño de investigación epidemiológica en el que se parte de la identificación, en la población de estudio, de los individuos sanos, es decir, que no han desarrollado la enfermedad, y a los que se divide en el grupo de individuos expuestos al factor que se desea estudiar y los no expuestos. A ambos grupos se les sigue durante el tiempo que se establezca en el diseño, con el fin de observar los "acontecimientos" —léase enfermedad, muerte, etc.— y que son los resultados que el investigador sospecha puede causar la exposición al factor de estudio. La Figura 4 muestra de forma esquemática este tipo de diseño.

Métodos básicos

La selección del grupo expuesto se realiza en base a criterios diversos, como el haber estado sometidos a un determinado nivel de exposición, o por ofrecer algún recurso especial que facilite la determinación de la exposición o el seguimiento y obtención de información sobre los acontecimientos de interés. En el diseño del estudio debe quedar definido muy estrictamente lo que se considera "exposición". El grupo de no expuestos debe provenir de la misma población que los expuestos y sólo diferenciarse de él en la ausencia de la exposición. Este es uno de los aspectos importantes en el

Figura 4. Representación esquemática de los estudios de cohortes.



diseño de un estudio de cohortes: el asegurar la comparabilidad de ambos grupos. El ensayo clínico aleatorizado, que de hecho es un estudio de cohortes experimental, asegura esta comparabilidad mediante la aleatorización.

La información se obtiene a partir de la información clínica ya registrada, de la información proporcionada por los participantes –cuestionarios y de los procedimientos de examen que se determinen en función de los acontecimientos que deseen estudiarse. Se ha de evitar el hecho de que se obtenga la información en ambos grupos de forma distorsionada. Ejemplo de ello es el que el grupo expuesto, en razón de su exposición (por ejemplo el uso de anticonceptivos orales) puede tener más contactos con el sistema de atención sanitaria y ello puede causar un aumento de la probabilidad de detección de determinados acontecimientos. En este caso, el “efecto” podría estar sobrerrepresentado en el grupo de expuestos.

Los resultados se analizan de forma distinta según que la información de la exposición y de los acontecimientos sea cuantitativa o cualitativa. A continuación vamos a ver la forma de análisis cuando tanto la exposición como el resultado son variables cualitativas dicótomas.

Los datos se organizan en una tabla 2 x 2 para su análisis:

		Resultado	
		R	No R
Exposición	E	a	b
	No E	c	d
		a + c	b + c

Incidencia en expuestos y en no expuestos: $IE = a/a+b$; $I\bar{E} = c/c+d$

$$RR = IE / I\bar{E} = a/a+b/c/c+d$$

$$RA = (a/a+b)-(c/c+d)$$

En el caso de un estudio de densidad de incidencia, se puede realizar el análisis sustituyendo “b” y “d” por las “personas-tiempo” de seguimiento.

– Si $RR = 1,0$ no hay asociación. Si $RR > 1,0$ la exposición está asociada con un mayor riesgo. Si $RR < 1,0$ La exposición está asociada con un riesgo menor.

– Si $RA = 0$ no hay asociación. Si $RA > 0$ la exposición está asociada con un mayor riesgo. Si $RA < 0$ la exposición está asociada con un riesgo menor.

En la interpretación de los resultados se deben considerar el cumplimiento de unos criterios de causalidad tales como:

1. Fuerza de la asociación (tamaño del efecto) y su significación estadística.
2. La consistencia con otros resultados.
3. La relación “dosis-respuesta”.
4. La especificidad de la asociación.
5. La falta de ambigüedad temporal.
6. La plausibilidad biológica.

La magnitud de la asociación, expresada por la magnitud del RR, necesita ser corroborada por una prueba de significación estadística, que nos asegure una probabilidad muy baja de que la asociación encontrada sea debida al azar. La evaluación de la precisión estadística se puede realizar mediante una prueba de significación estadística, se expresa como un valor de la p ($p < 0,05$) o bien mediante la estimación del intervalo de confianza. La ventaja de esta segunda opción es que en lugar de reducir el problema a una expresión del resultado en forma dicótoma, es decir “significativo o no”, se da más información al dar una idea de la magnitud del efecto. Con el intervalo de confianza se da una idea de entre qué valores límite se pueden encontrar los valores reales de RR y RA, con un determinado nivel de confianza estadística. Para un cálculo aproximado del intervalo de confianza (IC), se puede utilizar el “método basado en el test” de Miettinen, en base al test de Ji cuadrado⁽¹¹⁾:

$$\chi^2 = \frac{n (ad - bc)^2}{(a+c)(b+d)(a+b)(c+d)} \quad (n = a+b+c+d)$$

Los límites del IC para el RR y el RA son respectivamente:

$$\text{lim. inf.} = RR^{1-(z/\chi^2)} \quad \text{lim. sup.} = RR^{1+(z/\chi^2)}$$

$$\text{lim. inf.} = RA - z / -\chi^2 \quad \text{lim. sup.} = RA + z / -\chi^2$$

Existen otros métodos de estimación aproximada, así como un método exacto de cálculo del intervalo de confianza, que el lector puede encontrar en el libro de

Rothman⁽¹²⁾. Si, para el RR el IC no incluye al 1,0, se puede rechazar la hipótesis nula con un riesgo a prefijado. Para el RA, se puede rechazar la hipótesis nula cuando el IC no contiene el 0.

Los estudios de cohortes, por la existencia, de secuencia temporal permiten contrastar hipótesis etiológicas. Frente a otros estudios observacionales, ofrecen un mayor grado de evidencia causal y proporcionan datos de incidencia y una estimación de la proporción de los casos prevenibles si se puede controlar la exposición. Por otra parte, los problemas principales que comportan este tipo estudios son el largo tiempo de seguimiento, el gran número de individuos enrolados y, por lo tanto, su costo. El largo periodo de seguimiento que se necesita en muchos casos, facilita la posibilidad de pérdidas de seguimiento y el que se produzcan otros cambios (ambientales, culturales, en el tipo de atención, etc.), que influyan en los resultados. Lógicamente, en estos estudios, como la exposición se identifica al inicio del estudio, no sirven para generar hipótesis etiológicas.

c) Estudios de Casos y Controles

Estos estudios se caracterizan porque los grupos a estudiar se identifican en base a si presentan el resultado (casos) o no lo presentan (controles). Ejemplo de este tipo de diseño es el estudio de la imputación de la hemorragia digestiva alta a la exposición a determinados fármacos. Para ello se identifican los individuos que desarrollen la enfermedad y el grupo de controles deberá ser elegido en función de su comparabilidad a los casos, excepto para el desarrollo de la misma. En este caso, por ejemplo, se debe vigilar que los controles y los casos no tengan distintas probabilidades de exposición a los fármacos imputables. Así, si se eligen como controles enfermos de reumatología ingresados en el hospital, pueden tener mayor probabilidad de estar expuestos a antiinflamatorios no esteroideos.

Métodos básicos

La selección de los controles es uno de los problemas metodológicos más importantes en este tipo de estudios. Algunas veces para evitar los problemas de comparabilidad se seleccionan los controles a partir de dos fuentes distintas, por ejemplo, ingresos hospitalarios de los propios centros de donde se han elegido los casos, y cuyo motivo de ingreso sea diverso y no relacionado con los temas de estudio y controles de origen comunitario. Si la identificación de la exposición se realiza mediante cuestionario, puede

ser necesario que el nivel cultural sea semejante al de los casos para evitar que la información sobre la exposición sea distinta para los casos y los controles. Debemos asegurarnos de que si los controles hubieran sido casos serían detectados (regla de oro). A continuación, en la Figura 5, se representa esquemáticamente este tipo de diseño.

El análisis de los resultados se puede realizar mediante la distribución de los resultados en una Tabla 2 X 2 en la que se compara la frecuencia de exposición en los casos y en los controles:

		Resultado	
		Casos	Controles
Exposición	E	a	b
	No E	c	d
		a + c	b + c

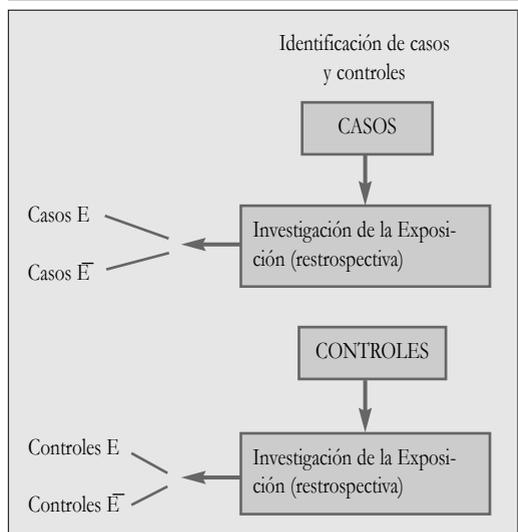
La razón de las Odds (RO) de exposición entre los casos y los controles es:

$$RO = a \times d / b \times c$$

(Nota: El lector interesado puede encontrar en los textos de referencia la forma en que se llega a esta fórmula)

A la RO también se la denomina “razón de los productos cruzados” o “razón de ventaja”. Se debe observar que lo que en este caso se compara son las proba-

Figura 5. Representaciones esquemáticas de los estudios de casos y controles.



bilidades de exposición al factor de estudio, siendo un caso o un control, sin que exista un periodo de seguimiento de una población definida, por lo que este tipo de estudios no proporciona datos de incidencia. Sin embargo, en muchas ocasiones, la RO puede ser considerada como una estimación del RR, y muchas veces se expresan los resultados de este tipo de estudios hablando de RR. En eventos que se produzcan con baja incidencia (<10%), el valor numérico del OR y el del RR serán semejantes. Cuando el suceso es más frecuente del 10% pero la diferencia entre grupo casos y grupo controles es pequeña también resultan similares las estimaciones del RR y de la OR.

La interpretación de los resultados se realiza de igual forma que la descrita en el caso de los estudios de cohortes, incluyendo el cálculo de los límites de confianza. En muchas ocasiones la selección de controles se realiza mediante apareamiento (matching) individual con los casos y entonces el método de análisis y los cálculos estadísticos son distintos, aunque la extensión de este capítulo no nos permite entrar en ellos.

Así como los estudios de cohortes son poco factibles cuando se trata, de enfermedades de baja frecuencia, los estudios de casos y controles son muy eficientes para el estudio de este tipo de enfermedades. En cambio, este tipo de diseño no es útil en el caso de frecuencias bajas de exposición. Al mismo tiempo, al no requerir largos periodos de seguimiento, son muy efectivos en el estudio de la asociación entre exposición y enfermedad cuando existen largos periodos de latencia—por ejemplo, inducción de algún tipo de cáncer—. Por todo ello, son estudios mucho menos costosos que los de cohortes, más eficientes estadísticamente, pero tienen su debilidad en la gran probabilidad de sesgos de selección o factores de confusión—ver más adelante— en la selección de los controles y en la dificultad que supone obtener información válida sobre la exposición, al ser ésta investigada retrospectivamente. Los estudios de casos y controles, en general, proporcionan un grado de evidencia menor que los estudios de cohortes y no dan información tan precisa sobre los efectos potenciales en la población, pero por su factibilidad está siendo el tipo de diseño más utilizado y más perfeccionado, especialmente en el estudio de los efectos de los medicamentos.

Diseño híbrido

En la actualidad, para combinar las ventajas de eficiencia estadística de los estudios de casos y controles

y reducción de sesgos potenciales en la selección del grupo control y en la determinación a el (los) factor(es) de riesgo, se han desarrollado diseños de estudio, llamados, en general, “híbridos”. Este tipo de diseños se caracterizan por emplear en el análisis estadístico tan sólo una muestra de los componentes de una cohorte previamente definida y seguida en el tiempo. Los distintos tipos de muestreo, a partir de los componentes de la cohorte inicial, dan lugar a dos tipos básicos de diseño híbrido.

En los estudios de casos y controles híbridos o anidados (nested en inglés) el análisis se realiza como en un estudio de casos y controles, tomando todos los individuos que han desarrollado la enfermedad o desenlace de interés como casos y una muestra de los controles.

En los estudios de cohortes y casos o estudios cohorte-caso, el método de muestreo se basa en que, a partir de los componentes iniciales de una cohorte, se selecciona aleatoriamente una subcohorte que servirá como grupo control (de comparación) para todos los casos que aparezcan durante el periodo de seguimiento. Las formas concretas de muestreo para ambos diseños, así como el análisis estadístico, en el que deben realizar ciertas asunciones relacionadas con el tipo de selección de la muestra, exceden la extensión de este capítulo; el lector interesado puede consultar el excelente artículo de E. Guallar⁽¹⁵⁾.

Este tipo de diseños pueden tener una especial consideración en los estudios farmacoepidemiológicos, cuando éstos se van a realizar aprovechando la información obtenida a partir de la conexión de archivos clínicos (ver más adelante). El último ejemplo de este tipo de estudios es el realizado por los componentes del Departamento de Epidemiología y Bioestadística de la Universidad de McGill en Canadá, utilizando la Base de Datos de Saskatchewan, acerca de la posible asociación entre la mortalidad en individuos asmáticos y el uso de agonistas betaadrenérgicos (N Engl J Med 1992. 326:501-6).

2.3. Validez de los estudios

La validez de un estudio valora la medida en que no existe error sistemático o sesgo en la estimación del efecto, objeto de nuestro estudio. En otras palabras, si la asociación encontrada es real o ha sido influida por la distorsión introducida en la selección de los sujetos, en la obtención de la información o en el análisis de los resultados.

Ni que decir tiene, que el estudio debe también tener precisión, característica asociada a la ausencia de error aleatorio, el cual depende del tamaño de la muestra, la varianza del parámetro estimado y la eficiencia estadística –cantidad de información estadística por individuo, referida al costo de su obtención–. La extensión de este capítulo no nos permite tratar de forma sistemática los problemas estadísticos de los estudios epidemiológicos. La validez de un estudio tiene dos componentes, la validez interna y la validez externa o aplicabilidad, que es la validez del estudio para los sujetos incluidos en él y la validez externa, que se refiere a la validez del estudio para los individuos de otras poblaciones no incluidas en el estudio.

- La validez interna se plantea si los resultados obtenidos proporcionan una respuesta correcta a la pregunta planteada en el objetivo del estudio, con ausencia de sesgos en los procesos de diseño, ejecución y análisis. La validez interna es el grado en el que las conclusiones de un estudio reflejan correctamente lo que sucedió.
- La validez externa de un estudio nos informa sobre la aplicabilidad de las conclusiones del mismo a otros pacientes distintos de los que participaron.

2.3.1. Validez interna

En la literatura epidemiológica y clínica de las dos últimas décadas han aparecido no pocas referencias a las posibles fuentes de sesgo en los estudios e, incluso, se han generado amplias controversias en torno a la validez de algunos estudios publicados. En general, se ha tendido a agrupar la multiplicidad de sesgos a los que el investigador se debe enfrentar en tres categorías: sesgos de selección, sesgos de información y factores de confusión^(13,14).

a) Sesgos de selección

Los sesgos de selección se producen cuando los procedimientos empleados en la selección de los individuos incluidos en el estudio, causan una distorsión del efecto medido. En otras palabras cuando el sistema de selección de los individuos no garantiza la comparabilidad en los dos grupos, diremos que existe sesgo de selección.

En los estudios de cohortes, la principal fuente de sesgos de selección reside en que existan pérdidas de seguimiento y éstas se produzcan por algún factor rela-

cionado con la exposición, con lo que los resultados de los dos grupos analizados al final estarían distorsionados por la falta de comparabilidad entre los mismos. En los estudios de casos y controles, la principal fuente de sesgo de selección reside en la selección de los controles, como ya se comentó anteriormente.

b) Sesgos de información

Los sesgos de información se producen cuando la forma de obtención de la información distorsiona el efecto medido. En otras palabras, cuando la información obtenida para ambos grupos no es comparable.

En los estudios de cohortes, el riesgo de presencia de este tipo de sesgo se encuentra principalmente en la obtención de la información sobre el resultado. Por ejemplo si estudiamos a una cohorte de mujeres expuestas a anticonceptivos orales y una no expuesta, el hecho de tomar el fármaco aumenta la probabilidad de que exista un mayor control médico. Un mayor número de contactos con el médico –visitas, análisis, etc.– puede aumentar la probabilidad de diagnóstico de acontecimientos clínicos en el grupo de expuestas, lo que podría producir una sobreestimación del riesgo asociado a la exposición a los anticonceptivos. En los estudios de casos y controles, la probabilidad de sesgo de información se encuentra básicamente en la investigación retrospectiva de la exposición. Por ejemplo, en un estudio de la asociación entre malformaciones y toma de determinados fármacos durante el embarazo, si la investigación de la exposición se realiza sólo por cuestionarios, es muy probable que las madres de niños malformados recuerden mejor que tomaron durante el embarazo que las madres de niños sanos, con lo que habría una sobreestimación de la exposición en el grupo de casos. A este tipo de sesgo de información se le llama sesgo de memoria y se puede controlar mediante la obtención de la información por dos fuentes distintas –por cuestionarios y por revisión de las historias clínicas– o con alguna estrategia para unificar el recuerdo en ambos grupos. Otro tipo de sesgo sería el relacionado con el investigador que realiza el cuestionario, el cual puede tener una actitud más inquisitiva al realizar los cuestionarios en los casos.

c) Factores de confusión

El factor de confusión es un factor de riesgo que está asociado al mismo tiempo a la enfermedad y a la

exposición. Un factor de confusión ha de ser un determinante de la enfermedad o bien de su diagnóstico; hallarse asociado a la exposición y por lo tanto distribuirse desigualmente entre los expuestos y los no expuestos y no ser simplemente un elemento intermedio en la cadena etiológica.

Spongamos un estudio en el que se ha encontrado una asociación entre la ingesta de alcohol y el cáncer de pulmón, mediante la comparación de las incidencias de este tipo de cáncer de bebedores y abstemios. En realidad, se conoce que existe un factor de riesgo para el cáncer de pulmón, como es el consumo de tabaco que se distribuye desigualmente entre bebedores y abstemios, es decir, existe una proporción mucho más alta de fumadores en el grupo de bebedores que en el de los abstemios. En este caso, el hábito de fumar está relacionado a la vez con un mayor riesgo de cáncer de pulmón y con una mayor ingesta de alcohol y, por ello, es un factor de confusión, que tiende a distorsionar la relación entre cáncer de pulmón e ingesta de alcohol.

Un caso particular en farmacoepidemiología lo representa el factor de confusión por la indicación, en el que una determinada patología está asociada al fármaco en estudio, en razón de su indicación en aquella y, a la vez, está asociada al resultado. Por ejemplo, la enfermedad coronaria, así como otras indicaciones de los betabloqueantes, aumentan el riesgo de muerte súbita, por lo que estas patologías son un factor de confusión en la asociación entre el uso de los betabloqueantes y la muerte súbita. La presencia de este tipo de factores de confusión es la causa de las dificultades de la valoración epidemiológica de los efectos de los fármacos mediante estudios observacionales, como comentaremos más adelante.

Control de los factores de confusión

Con el control de los factores de confusión se trata de neutralizar su influencia en la estimación del efecto estudiado.

El control se puede realizar en la fase del diseño del estudio mediante tres técnicas: la restricción, por la cual se incluyen en el estudio a aquellos individuos que presentan, para el factor de confusión, unos determinados valores, la randomización o asignación aleatoria, por la cual se garantiza la distribución equitativa en los dos grupos para el factor de confusión - técnica más propia del ensayo clínico- y el apareamiento, mediante la cual a cada individuo de uno de los grupos se le asigna otro individuo del grupo de comparación, que presente iguales o muy parecidos valores para el factor de con-

fusión.

El control de los factores de confusión al contrario que para los sesgos de selección y de información, también se puede realizar en la fase de análisis de los resultados. Las dos técnicas empleadas son:

– **Estratificación:** se agrupa a los sujetos según categorías, definidas para unos determinados valores del factor de confusión, o para la presencia o no del mismo. Si la estratificación es suficientemente adecuada, se puede neutralizar el efecto del factor de confusión en cada estrato. El procedimiento más común de análisis estratificado es el de Mantel-Haenszel, por el cual los resultados de cada estrato son ponderados de acuerdo con el tamaño del estrato, obteniéndose una estimación del efecto global y ponderada por el peso de los distintos estratos.

Otro método posible, cuando existe una población de referencia, es el método anteriormente comentado de estandarización o ajuste.

El análisis por estratos es la técnica de elección para el control de los factores de confusión, cuando las variables confundentes son pocas y categóricas.

– **Modelización:** la modelización o regresión múltiple consiste en la aplicación de un modelo matemático que permite obtener estimaciones de efectos que pueden, considerarse ajustados para varios factores de confusión al mismo tiempo. Sin embargo, aunque la modelización es muy atractiva, sus resultados pueden ser difíciles de interpretar. Conviene aplicarla con criterios muy claros y previa definición de la hipótesis de partida.

En la Tabla 2 se resumen las ventajas e inconvenientes de las distintas estrategias de control de los factores de confusión.

2.3.2. Validez externa

Definíamos la validez externa como la capacidad de generalizar los resultados del estudio a otra población distinta de la estudiada, por el lugar o el momento en que se ha realizado aquél. Condición necesaria para la presencia de validez externa es que el estudio tenga validez interna y, obviamente, que tengan precisión estadística. Generalmente, el hecho de que se asegure la validez del estudio comporta el reducir la validez externa, ya que, en muchos casos, para asegurar una alta validez interna que nos garantice la respuesta correcta a preguntas muy específicas, es necesaria una selección muy

Tabla 2. Ventajas e inconvenientes de los distintos tipos de control de los factores de confusión.

	Ventajas	Inconvenientes
Restricción	<ul style="list-style-type: none"> – Simple de aplicar y analizar – Aumenta la eficacia del estudio de la homogeneidad de las unidades de observación. 	<ul style="list-style-type: none"> – Alarga el tiempo del estudio. – Reduce la validez externa. – Si no está bien hecha puede quedar residual algo del efecto del factor de confusión.
Apareamiento	<ul style="list-style-type: none"> – Aumenta la eficiencia del estudio con estimaciones más precisas. 	<ul style="list-style-type: none"> – Alarga el tiempo del estudio. – Problemas de factibilidad.
Asignación al azar	<ul style="list-style-type: none"> – Garantiza la comparabilidad de los grupos. – Aumenta la eficacia del estudio. No es necesario el control en el análisis. 	<ul style="list-style-type: none"> – No es posible realizarla en muchos casos por razones éticas.
Estratificación	<ul style="list-style-type: none"> – Simple; no impone estructuras a los datos y favorece la interpretación de los resultados. 	<ul style="list-style-type: none"> – Categorizar puede perder alguna información. El número de estratos es limitado. – Con muchas variables de confusión la interpretación se hace más compleja.
Modelización	<ul style="list-style-type: none"> – No se pierde información y no se ha de categorizar – Pueden proporcionar estimaciones individuales de riesgo. 	<ul style="list-style-type: none"> – Las asunciones de base no siempre se cumplen o son plausibles. – No facilitan el contacto directo con los datos y pueden dar resultados no interpretables fácilmente.

restringida de sujetos. La forma habitual de aumentar la validez externa es la de replicar el estudio en otras poblaciones o subpoblaciones, siempre que la replicación sea necesaria y el costo asumible. De hecho, cuando factores como el tiempo, los procedimientos, las características de las poblaciones o las variables culturales difieren de las del estudio original, mayor interés tiene la replicación de aquél, con el fin de asegurar la aplicación de los resultados a otra población. También, cuanto más inesperado o contradictorio con otros resultados sea un hallazgo, mayor será la necesidad de asegurar la validez externa mediante la replicación del estudio.

Recomendamos la utilización de las “listas guía” para la evaluación crítica de validez de los estudios publicados, que el lector podrá encontrar en la bibliografía.

3 FARMACOEPIDEMIOLOGÍA

En la introducción ya se definió el concepto de farmacoepidemiología. Veamos ahora como se concreta en el ámbito de sus aplicaciones. En realidad, se trata de considerar las tres categorías de la investigación epidemiológica, descriptiva, analítica y evaluativa, tomando como punto de referencia al medicamento, sea para poner el punto de observación en su utilización, sea para ponerlo en sus efectos. El medicamento, dentro de un contexto socio-cultural y sanitario es un recurso terapéutico, un agente capaz de producir yatrogenia y un indicador socio-sanitario⁽¹⁶⁾.

3.1. Estudios de utilización de medicamentos

La investigación sobre la eficacia de los medicamentos, mediante los ensayos clínicos, ha crecido mucho en los últimos años. Sin embargo, las

condiciones de los ensayos clínicos no corresponden exactamente a las de la práctica habitual, porque las poblaciones incluidas en los primeros no son necesariamente representativas de la población que recibe el fármaco una vez comercializado, y porque las condiciones en las que se realizan estos estudios no son las mismas que las habituales.

Los estudios de utilización de medicamentos (EUM) tienen por objeto examinar en qué grado se pueden transferir a la práctica habitual los conocimientos adquiridos en los ensayos clínicos. También son una forma de auditoria terapéutica, cuyo fin es identificar áreas de intervención informativa, educativa o de otro tipo, para mejorar la calidad de la terapéutica en la práctica clínica. Podemos definir los EUM como aquellos estudios epidemiológicos descriptivos con los que se pretende determinar cuáles son los patrones, perfiles y otras características de la oferta, la prescripción, la dispensación, el consumo, el cumplimiento terapéutico y cualquier otra faceta relacionada con los medicamentos en una población determinada, con el objetivo de conseguir, mediante el análisis de los mismos, su uso racional. La OMS los definió como el estudio de la comercialización, distribución, prescripción y uso de medicamentos en una sociedad, con acento especial en las consecuencias médicas, sociales y económicas resultantes⁽¹⁷⁾. Típicamente, los EUM pueden abarcar todas o algunas de las siguientes etapas:

- Descripción de la utilización de medicamentos: proporciona una fotografía de cuál es la situación respecto a la prescripción, consumo, etc.
- Análisis cualitativo de los datos obtenidos, para identificar posibles problemas de inadecuación en términos de utilización insuficiente o excesiva en una determinada indicación o en una población de características dadas, en comparación con sus alternativas en función de la relación beneficio/riesgo y coste/efectividad, en relación al esquema terapéutico considerado ideal o de referencia, etc.
- Identificación de áreas de intervención en función del análisis anterior.
- Evaluación del impacto de las intervenciones sobre los problemas identificados.

a) Estudios de la oferta de medicamentos

Proporcionan una descripción de cuál es el arsenal farmacoterapéutico disponible en un determinado ám-

bito (un país, un hospital, una organización sanitaria, etc.). Su interés radica en que:

- Ofrecen un marco de referencia para los estudios de consumo y pueden ayudar a entender los resultados de aquéllos, ya que la oferta puede ser un condicionante importante de los mismos.
- Permiten identificar problemas cualitativos (por ejemplo, de accesibilidad) derivados de la propia oferta.

Un caso especial de estudios de oferta es, en los hospitales, el estudio de la prescripción o solicitud de medicamentos no incluidos en la guía farmacológica; su registro y análisis de rutina permite disponer de un indicador del grado de conocimiento y cumplimiento de la guía, así como de un indicador de posibles, necesidades terapéuticas no cubiertas en la misma. Ello permite una adecuación del proceso de selección de medicamentos y de las intervenciones educativas según las necesidades del hospital (véanse referencias generales sobre EUM).

b) Estudios de consumo

Suponen un método relativamente sencillo de seguimiento para detectar problemas potencialmente importantes, cuantitativamente, de inadecuación en la utilización de medicamentos en un primer nivel de análisis, si bien su grado de fiabilidad no siempre es grande, porque no se tiene en cuenta la patología o la indicación en la que los medicamentos son utilizados. Por ello, para que los estudios de consumo cumplan su objetivo, es imprescindible que se realicen de forma sistemática y continuada a lo largo del tiempo, lo que permitirá determinar tendencias de consumo en una población o ámbito dados.

Los estudios de consumo también pueden ser útiles para estimar la prevalencia de ciertas enfermedades (cuando se trata de medicamentos utilizados específicamente en indicaciones muy concretas, como por ejemplo los fármacos antituberculosos). Por ejemplo, ya en 1988 se publicó en el BMJ⁽¹⁸⁾ la validación de los datos de consumo de los nitritos como estimador de la prevalencia de ángor en una región del Reino Unido.

También pueden servir para tener una estimación aproximada de la incidencia de las reacciones adversas a un determinado fármaco, al proporcionar el denominador del cociente que cuantifica dicha magnitud.

b.1) Clasificaciones de medicamentos

Para que los estudios de consumo sean válidos, es esencial que los resultados se expresen utilizando una clasificación universal que tenga sentido terapéutico y que permita la comparación entre distintos principios activos o grupos terapéuticos.

Ante la necesidad de la existencia de una metodología internacionalmente aceptada que permitiera la realización de estudios de consumo de medicamentos, se creó en 1969 el Drug Utilisation Research Group (DURG) con la misión de desarrollarla. Al mismo tiempo, mediante la modificación de la clasificación de medicamentos adoptada por la European Pharmaceutical Market Research Association (EPhMRA), el Norwegian Medicinal Depot (NMD) desarrolló un nuevo sistema conocido como Anatomical Therapeutic Chemical (ATC) Classification. En esta clasificación, los medicamentos son divididos en 14 grupos de acuerdo con el órgano o sistema anatómico sobre el que actúan (primer nivel), los cuales se subdividen a su vez en dos niveles más que corresponden a subgrupos terapéutico/farmacológicos. El Nordic Council on Medicines (NCM), establecido en 1975, modificó esta clasificación añadiendo dos niveles más, el cuarto hace referencia al subgrupo químico-terapéutico y el quinto a la entidad química.

El siguiente ejemplo ilustra la estructura de la clasificación:

A	Aparato Digestivo y Metabolismo (1 ^{er} nivel, grupo anatómico principal).
A10	Medicamentos utilizados en la Diabetes (2 ^o nivel, grupo farmacológico /terapéutico principal).
A10B	Antidiabéticos orales (3 ^{er} nivel, subgrupo químico/farmacológico /terapéutico).
A10B A	Biguanidas (4 ^o nivel, subgrupo químico/farmacológico/terapéutico).
A10B A02	Metformina (5 ^o nivel, subgrupo para entidad química).

Por lo tanto, en la clasificación ATC, los preparados de Metformina tienen el código A10A02.

Los medicamentos son clasificados de acuerdo con el uso terapéutico más importante de su principal principio activo, bajo la regla de un solo código ATC para cada formulación farmacéutica (es decir, principios activos, dosis y forma farmacéutica similares). Un principio activo puede tener más de un código ATC si se utiliza en

dosis o formulaciones diferentes para usos terapéuticos claramente distintos.

Un principio activo utilizado en dos o más indicaciones igualmente importantes, o en los que la indicación principal difiere de unos países a otros, generalmente recibe un sólo código, aunque se incluyen referencias cruzadas en la guía de utilización de la clasificación ATC para indicar los distintos usos de estos principios activos.

Los principios activos clasificados en el mismo subgrupo de 4^o nivel no pueden considerarse terapéuticamente equivalentes, ya que su mecanismo de acción principal, su efecto terapéutico y su perfil de interacciones y efectos secundarios pueden diferir.

Sólo se asignan códigos ATC a asociaciones de principios activos en casos excepcionales, y siempre que dicha asociación esté ampliamente extendida a nivel internacional (por ejemplo, asociaciones de betabloqueantes y diuréticos).

En España, la clasificación oficial sigue básicamente la ATC, aunque en algunos casos presenta variaciones relevantes respecto a ésta. Aunque recientemente el Insalud ha publicado una lista actualizada y nuevos criterios.

b.2) Unidades de medida

Para poder medir la utilización de los medicamentos, tan importante como tener un sistema de clasificación adecuado es disponer de una unidad de medida apropiada. Los resultados cuantitativos de un estudio de consumo pueden expresarse de muy diversas formas: en valor económico, en número de prescripciones, en unidades físicas (gramos, envases, etc.), pero todas ellas carecen de un valor epidemiológico adecuado, o dificultan mucho la posibilidad de comparaciones temporales, entre distintos ámbitos geográficos o entre distintos principios activos o grupos terapéuticos.

Para evitar los problemas e inconvenientes de las unidades de medidas tradicionales, el NMD desarrolló una unidad técnica de medida denominada Dosis Diaria Definida (DDD), que se define como “la dosis promedio de mantenimiento en adultos para la indicación principal del principio activo considerado”.

La asignación del valor de la DDD a los distintos medicamentos se rige por los siguientes principios:

- Salvo excepciones, las DDDs están basadas en las dosis utilizadas en monoterapia.
- Nunca se asigna una DDD antes de que el medica-

- mento esté registrado y comercializado en al menos un país.
- La indicación principal es la derivada del código ATC del medicamento.
 - Cuando la dosis recomendada de un principio activo esté en función del peso corporal, se considera la dosis correspondiente a una persona de 70 kg.
 - Cuando el medicamento se utiliza fundamentalmente en niños (por ejemplo, hormona de crecimiento) se utiliza la dosis pediátrica.
 - Generalmente se utiliza la dosis de tratamiento, salvo que la indicación principal sea de profilaxis (por ejemplo, algunos antipalúdicos).
 - Aunque un fármaco se emplee a dosis diferente cuando se inicia el tratamiento, se elige siempre la dosis de mantenimiento.
 - La DDD se establece de acuerdo con el contenido declarado del componente activo, de manera que las diferentes sales de un mismo principio activo tienen una sola DDD.
 - A veces, la DDD es la media de dos o más dosis utilizadas comúnmente en la misma indicación.
 - La DDD que se asigna a combinaciones de varios principios activos es generalmente la misma que tiene el principio activo que se encuentra en su mismo subgrupo terapéutico ATC.
 - Las DDDs asignadas por primera vez se revisan al cabo de 3 años, y posteriormente cada 5 años.

En 1981, la Oficina Regional para Europa de la OMS recomendó el uso del sistema ATC/DDD como estándar internacional para la realización de estudios de utilización de medicamentos. Para conseguir el uso generalizado de esta metodología se creó en 1982 el WHO Collaborating Centre for Drugs Statistics Methodology, con sede en Oslo, que se encarga de coordinar la revisión y actualización de la clasificación ATC y de las DDD.

Debe quedar claro que la DDD es una unidad de medida que no refleja necesariamente la dosis diaria recomendada o prescrita, ya que éstas deben basarse en características individuales (edad, peso) y en consideraciones farmacocinéticas. Los datos de consumo expresados en DDD sólo proporcionan una estimación aproximada, y no una fotografía exacta de su utilización real. Las DDDs constituyen una unidad de medida única, independiente del precio y de la forma farmacéutica, que permite evaluar las tendencias en el consumo de medicamentos y llevar a cabo compara-

ciones entre diferentes poblaciones.

No se establecen DDDs para preparados de uso tópico, sueros, vacunas, agentes antineoplásicos, extractos alérgicos, anestésicos locales o generales, ni medios de contraste.

El Centro Colaborador de la OMS de Oslo es quien establece la inclusión de nuevos medicamentos en la clasificación ATC a petición de los usuarios del sistema (laboratorios farmacéuticos, agencias reguladoras e investigadores), generalmente cuando ha sido autorizado su registro en algún país, y quien asigna un valor de DDD a los mismos, además de estimular y fomentar el uso práctico del sistema ATC/DDD cooperando con los investigadores en el campo de la utilización de los medicamentos. Anualmente publica una nueva edición del índice ATC con las correspondientes DDDs, consistente en dos listados, clasificados por orden alfabético de principio activo y de código ATC, respectivamente. Para algunos grupos de medicamentos para los que no se han establecido DDDs, se recomienda la utilización de unidades de medida alternativas. Por ejemplo, el consumo de preparados dermatológicos puede expresarse en gramos pomada, crema, etc.

Cuando existen discrepancias conocidas entre la DDD y la dosis diaria prescrita (PDD), es importante tener este hecho en cuenta a la hora de interpretar los datos de consumo (por ejemplo, la Ampicilina en el ámbito hospitalario)⁽¹⁹⁾. También debe tenerse precaución en situaciones donde la dosis recomendada difiere de una indicación a otra (por ejemplo, las heparinas de bajo peso molecular en la profilaxis y el tratamiento de la enfermedad tromboembólica).

Ejemplos: para el caso de estudios de consumo ambulatorio, supongamos que las ventas anuales de ácido acetil salicílico han sido de 200 millones de comprimidos de 300 mg en una población de 4.035.000 habitantes y la DDD para el fármaco es de 3 g; la expresión de los resultados sería:

Tomando el mismo fármaco del ejemplo anterior, si

$$\text{DDD}/1.000 \text{ hab./día} = \frac{\text{N}^\circ \text{ de comp.} \times \text{peso de cada comp.} \times 1.000 \text{ hab}}{\text{DDD} \times \text{núm. hab.} \times 365}$$

$$= \frac{200.000.000 \times 300 \times 1.000}{3.000 \times 4.035.000 \times 365} = 13,66 \text{ DDD}/1000 \text{ hab./día}$$

el consumo anual en un hospital fue de 100.000 comp., el hospital tiene 1.000 camas y el índice de ocupación fue del 80%:

$$\text{DDD} / 100 \text{ est./día} = \frac{\text{N}^\circ \text{ de comp.} \times \text{peso de cada comp.} \times 100 \text{ est.}}{\text{DDD} \times \text{Núm. de camas} \times \text{índice ocup.} \times 365 \text{ días}}$$

$$= \frac{100.000 \times 300 \times 100}{3.000 \times 1000 \times 0,8 \times 365} = 10,27 \text{ DDD} / 100 \text{ est./día}$$

b.3) Aplicaciones del sistema ATC/DDD

El uso del sistema ATC/DDD permite la estandarización de la agrupación de los medicamentos y de los métodos de medida de su utilización, para hacer posible la comparación del uso de los medicamentos entre países, regiones y ámbitos de atención sanitaria, así como examinar las tendencias en el uso de los medicamentos a lo largo del tiempo.

Los datos de consumo de medicamentos deberían presentarse preferiblemente como número de DDD por 1.000 habitantes y día, o cuando se trata del consumo hospitalario, como número de DDD por 100 estancias. Los datos sobre ventas o sobre prescripción presentados en DDD/1.000 habitantes/día proporcionan una estimación grosera de la proporción de la población tratada diariamente con los medicamentos en cuestión. Por ejemplo, el dato de 10 DDD/1.000 habitantes/día indica que el 1% de la población estudiada, en promedio, recibe ese tratamiento diariamente.

Para antiinfecciosos y otros fármacos usados normalmente durante cortos periodos de tiempo, se considera más apropiado presentar los datos como número de DDDs por habitante y año, lo que proporciona una estimación del número de días durante los cuales cada habitante, por término medio, es tratado anualmente. Por ejemplo, 5 DDD/hab./año indica que el consumo es equivalente al tratamiento de cada habitante durante de 5 días a lo largo de un año. Por otro lado, si se conoce la duración normal de un tratamiento, el número total de DDDs puede transformarse en el número de tratamientos.

Consumo en hospitales: no existe hasta hoy un banco de datos de consumo de medicamentos en los hospitales en nuestro país, aunque existe alguna información procedente de estudios publicados, algunos de ellos multicéntricos y en su mayoría referentes al con-

sumo de antiinfecciosos⁽²⁰⁻²⁵⁾. Por otra parte, la informatización de los Servicios de Farmacia irá permitiendo disponer fácilmente de este tipo de información. El grupo de estudios de utilización de medicamentos de la FEFH esta desarrollando iniciativas en este tema (ver www.sefh.es).

Con los estudios de consumo puede avanzarse un poco más y pasar de la mera descripción de la situación al análisis cualitativo de la misma, utilizando indicadores, generalmente de dos tipos:

- Calidad farmacológica intrínseca de los medicamentos. Dentro de éstos podemos considerar la cuantificación del consumo de medicamentos o de subgrupos terapéuticos de utilidad terapéutica baja o de valor intrínseco poco elevado, es decir, aquellos fármacos de los que existen pocas evidencias de eficacia o con una mala relación beneficio/riesgo (por ejemplo, subgrupos terapéuticos como el C04A vasodilatadores cerebrales, el C05 antivaricosos o determinadas asociaciones a dosis fijas como las del subgrupo R05 de antiinfecciosos y mucolíticos). La cuantificación de estos subgrupos puede permitir el cálculo del porcentaje de prescripciones potencialmente inefectivas o innecesarias sobre el total de prescripciones, y sobre las cuales intervenir prioritariamente. La realización sistemática de estudios de utilización en los que se cuantifica el consumo de medicamentos de valor intrínseco no elevado (VINE), da lugar al establecimiento de indicadores de primer nivel de la calidad de la prescripción, de carácter oficial en determinados ámbitos, y que se utilizan para el establecimiento de objetivos asistenciales y de incentivos de los prescriptores.
- Grado potencial de uso de los medicamentos. Podemos comparar los datos de consumo de principios activos o subgrupos terapéuticos de primera elección frente a los alternativos o a los de uso excepcional en el ámbito que estemos considerando. Un ejemplo claro lo constituye la cuantificación del consumo de fármacos no incluidos en la guía o formulario, o la comparación del consumo de penicilinas de amplio espectro frente al de cefalosporinas en atención primaria.

Todos estos indicadores, y especialmente los que valoran la calidad intrínseca de los medicamentos, presentan limitaciones evidentes, derivadas

fundamentalmente de que no se conoce en qué indicación se han utilizado o cuáles son las características de los pacientes que puedan modificar la adecuación de las prescripciones. Hay que ser conscientes de que lo único que proporcionan son datos de consumo que pueden reflejar un bajo beneficio intrínseco o una escasa adecuación al ámbito asistencial donde se origina la prescripción. Por ejemplo, un alto consumo de medicamentos de valor intrínseco elevado no quiere decir necesariamente que la calidad de la prescripción sea elevada, porque puede tratarse de la utilización de medicamentos de alto valor intrínseco pero en indicaciones no adecuadas (por ejemplo, amoxicilina para tratar el resfriado común). La utilidad básica de estos estudios es la identificación de potenciales problemas terapéuticos, y también la identificación de potenciales ineficiencias en la asignación de recursos.

A la hora de interpretar los datos de consumo de medicamentos, debe tenerse en cuenta aquellos factores que puedan distorsionar dicha interpretación. Por ejemplo, en los estudios de utilización hospitalaria de medicamentos, se ha recomendado separar los datos de las unidades que no generan estancias hospitalarias (consultas externas, urgencias, etc.), o tenerse en cuenta la magnitud cuantitativa de servicios como el de pediatría que presentan unas características específicas en las dosis utilizadas.

c. Estudios cualitativos y Auditorías terapéuticas

Este tipo de estudios tiene como objetivo, no sólo describir el uso de los medicamentos en un ámbito determinado, sino identificar los factores relacionados con dicho uso, de manera que es un paso más respecto a los estudios de consumo, a efecto de contrastar las hipótesis que de ellos se han derivado.

c.1) Estudios de Prescripción-Indicación y de Indicación-Prescripción

Estos estudios permiten profundizar en un potencial problema detectado y conocer algunas de las razones de la prescripción de fármacos. En ellos se parte de una patología o indicación (indicación-prescripción) y se identifican las prescripciones asociadas a la misma, o bien se parte de un fármaco o grupo de fármacos (prescripción-indicación) y se analizan las indicaciones para las que son utilizados. Los diseños pueden ser:

– Retrospectivo, mediante la revisión de historias clíni-

cas o de las historias farmacoterapéuticas obtenidas de los registros informatizados en un sistema de dosis unitaria.

- Transversal o de prevalencia, en muestras de pacientes que, en un momento dado, reciban un fármaco o presenten una patología.
- Prospectivo, mediante el seguimiento de una cohorte de pacientes con una patología determinada o la detección en tiempo real de todas las prescripciones de un determinado medicamento. Este diseño permite, a la vez, el análisis de otros factores relacionados con el uso de los medicamentos (por ejemplo, la evaluación del cumplimiento por parte de los pacientes), pero permite también, sobre todo, la realización de intervenciones que aumenten la calidad de la utilización de los medicamentos antes de que ésta se produzca (por ejemplo, sugerencias sobre la selección o la posología del fármaco, y evaluación del impacto de las mismas).

Los objetivos de estos estudios pueden estar dirigidos a aspectos concretos, como la evaluación del cumplimiento de la prescripción por parte de los pacientes, a la detección de errores de medicación, en todas sus vertientes, o a la detección de problemas relacionados con la medicación en un contexto de atención farmacéutica.

c.2) Auditorías terapéuticas

Se trata de un instrumento de gran utilidad para la mejora del uso de los fármacos en cualquier ámbito asistencial, cuyos elementos esenciales son:

- Definición precisa del problema potencial, detectado a partir de estudios descriptivos o analíticos previos.
- Establecimiento previo de los criterios de uso correcto de los fármacos a analizar (indicación, posología, duración del tratamiento) o de la patología a estudiar (guías de práctica clínica y protocolos de tratamiento).
- Determinación del grado de coincidencia de uso real de los fármacos objeto del estudio y las definidas previamente como correctas, o del grado de adecuación del perfil farmacoterapéutico de los pacientes que presentan determinada patología a las recomendaciones de tratamiento de la misma.

Es muy útil en este tipo de estudios implicar a los prescriptores en su diseño y realización, ya que, de esta

forma, estaremos planteando directamente una estrategia de intervención educativa. Es importante la aplicación de criterios de priorización a la hora de seleccionar los fármacos y las patologías objeto de auditoría, para que de su realización puedan derivarse consecuencias relevantes respecto a la mejora de la calidad del uso de los medicamentos:

- Fármacos que implican un mayor volumen de consumo en términos económicos.
- Patologías de gran prevalencia, y de las que se sospecha un cierto grado de inadecuación en su tratamiento.
- Fármacos potencialmente ineficientes en función de la indicación de uso o forma de utilización.
- Fármacos con problemas importantes de seguridad, debido a efectos secundarios y reacciones adversas potenciales graves, o a interacciones relevantes con otros fármacos.
- Fármacos de reciente introducción, sobre los que existen pocos datos de efectividad o seguridad en la práctica, especialmente en determinadas indicaciones.

En general, para que los estudios de utilización de medicamentos consigan su objetivo último, que no es más que mejorar la calidad de la farmacoterapia, además de que tengan un diseño adecuado, es fundamental que posean determinadas características:

- Que formen parte de programas que tengan una continuidad en el tiempo, lo que va a permitir evaluar la evolución de las tendencias de uso y el impacto de las actuaciones que se pongan en marcha como consecuencia de la actuaciones realizadas. De poco sirven los estudios puntuales llevados a cabo en un momento determinado.
- Que en el diseño y realización de los estudios participen los propios implicados, especialmente los médicos prescriptores o sus organizaciones científicas representativas, ya que esto les dará credibilidad a los resultados y legitimidad a las actuaciones y programas de mejora que se establezcan. De hecho, la propia realización del estudio orientado a un problema específico puede ser una eficaz estrategia de intervención⁽⁶⁰⁾.
- Que se establezca un sistema de retroalimentación de la información generada, de manera que los resultados obtenidos lleguen a los propios médicos, alcanzando la mayor difusión posible, lo que, en sí mismo, constituirá una estrategia de mejora de la calidad en el uso

de los medicamentos.

3.2. Estudio de los efectos –positivos y negativos– de los fármacos tras su comercialización

La investigación clínica humana en la fase de precomercialización se basa en la realización del ensayo clínico controlado. Los resultados de los ensayos clínicos bien diseñados permiten la evaluación básica de la eficacia y seguridad de un fármaco, garantizando la ausencia de sesgos. Es el punto de partida necesario pero no suficiente de la evaluación clínica.

En primer lugar, el ensayo clínico debe, por su costo y complejidad, estar dirigido a resolver cuestiones relevantes en clínica y no ser un método de promoción. Esta es una cuestión básica a tener en cuenta, tanto en el diseño como en la lectura de un ensayo publicado, tanto por cuestiones éticas como de racionalidad en la producción de conocimientos. Sin embargo, el ensayo clínico en fase III tiene una serie de limitaciones, que emergen de sus propias características y de su metodología. Por una parte, las limitaciones de tipo ético que imposibilitan cierto tipo de experimentaciones en humanos. Por otra parte, por cuestiones tanto de método como de factibilidad, existen una serie de limitaciones que han sido resumidas en los “cinco demasiados”⁽⁶¹⁾:

- Demasiado pocos: el número de individuos expuestos a un fármaco en esta fase de investigación clínica precomercialización suele ser alrededor de 2.000. Para detectar determinados efectos relevantes, pero de baja frecuencia o determinadas diferencias clínicamente importantes pero pequeñas, se necesitan muchos más individuos, con el fin de trabajar con niveles de confianza estadística suficientes.
- Demasiado simples: los pacientes con situaciones médicas complejas y/o que reciben tratamientos complicados son habitualmente excluidos de la investigación en fase III. Pero es muy probable que estos enfermos reciban el fármaco tras su comercialización, sin que se conozcan con precisión sus efectos en ellos.
- Demasiado “adultos jóvenes”: en este tipo de ensayos clínicos se excluyen generalmente los muy jóvenes o los muy ancianos, pero posteriormente estos grupos de edad también van a ser candidatos potenciales a recibir el fármaco ya comercializado.
- Demasiado estrechos: las indicaciones en las que la

eficacia y seguridad del fármaco es investigada, son estrictamente definidos. Una vez el fármaco está disponible, puede ser usado para otras indicaciones o en otras situaciones, sin que se conozca su perfil beneficio/riesgo en ellas.

– Demasiado breves: Los resultados, tanto de eficacia como de seguridad, son valorados en periodos de tiempo relativamente cortos. La eficacia en tratamientos crónicos o los efectos adversos que se producen tras una larga exposición no pueden ser conocidos.

En resumen, aquellos resultados positivos o negativos que se producen con frecuencias bajas, pero que son clínicamente muy relevantes, o que se manifiestan tras un largo periodo de tratamiento, o que resultan influidos por una serie de características personales o patológicas del paciente o por otros tratamientos que éste recibe, deben ser monitorizados una vez el fármaco haya sido comercializado y su uso extendido en la práctica diaria⁽³²⁾. Todo ello se traduce en que, una vez el fármaco ha sido comercializado, no estamos en disposición, en la mayoría de casos, de responder a cuestiones tales como:

- ¿Cuál es el beneficio atribuible de una terapia determinada, aplicada en condiciones reales?
- ¿Cuál es el riesgo “cotidiano” de un tratamiento, del cual disponemos de adecuada información procedente de ensayos clínicos controlados?
- ¿Cuál es la relación beneficio/riesgo de un nuevo tratamiento respecto a otros disponibles y empleados en la práctica habitual?
- ¿Cuál es el “destino” de los pacientes que no respon-

den a un tratamiento de eficacia documentada?

- ¿Cuál es el beneficio y/o riesgo acumulado de los tratamientos empleados en una determinada patología?
- ¿Cuál es la influencia del contexto cultural en el beneficio y el riesgo de una determinada terapia?

3.2.1. La utilización de los conceptos y métodos epidemiológicos en la toma de decisiones

Los métodos y conceptos epidemiológicos se emplean cada vez más en los estudios publicados y son una de las bases importantes de la medicina basada en la evidencia.

Frecuentemente la decisión de utilizar un tratamiento en un paciente, o aún más su inclusión en una guía farmacológica, en un protocolo, etc, supone tomar una decisión en condiciones de incertidumbre, aún manejando la mejor evidencia disponible. El transformar los resultados de los estudios en una medida del impacto que tendrá en nuestros pacientes se ha propuesto como un elemento de gran ayuda. Los RR y RO se han incorporado como elemento de expresión de resultados no sólo en los estudios observacionales, sino también a los ensayos clínicos y a los meta-análisis.

Anteriormente se ha definido como medida de asociación que proporciona una idea de la magnitud del efecto de un fármaco el Número de Necesario de Pacientes a Tratar (NNT). A continuación en las siguientes Tablas se expone un ejemplo del valor interpretativo de los resultados presentados de esta forma:

Este tipo de análisis nos permite valorar real-

Tabla 3. Forma de analizar los resultados de eficacia

A	Incidencia grupo control o riesgo base	Ejemplo, $20/100 = 20\% = 0,2$
B	Incidencia grupo intervención	Ejemplo, $15/100 = 15\% = 0,15$
A - B	RAR (Reducción Absoluta del Riesgo o Diferencias de Riesgo o Riesgo atribuible)*	$0,2 - 0,15 = 0,05$
$1 / (A-B)$	NNT (Numero Necesario a Tratar)	$1 / (0,2 - 0,15) = 20$
B / A	RR (Riesgo Relativo o Razón de Riesgos)	$0,15 / 0,20 = 0,75$
$[(A-B) / A] \times 100$	RRR (Reducción Relativa del Riesgo)**	$[(0,2 - 0,15) / 0,2] \times 100 = 25\% = 0,25$

Tabla 4. Medidas del efecto de un tratamiento en función del riesgo basal. Mismo ejemplo de la Tabla anterior, suponiendo riesgos basales de 40%, 20%, 10% y 5%.

	Riego basal del grupo control			
	40%	20%	10%	5%
Incidencia en el grupo tratado	30	25	7,5	3,75
RRR (Reduccion Relativa del Riesgo)	0,25	0,25	0,25	0,25
RAR (Reducción Absoluta del Riesgo)	10	5	2,5	1,25
NNT (Numero Necesario a Tratar)	10	20	40	80

mente el impacto de del tratamiento considerando el riesgo basal, evitando la ambigüedad de una medida de beneficio relativo. La Tabla 4 completa el ejemplo anterior:

3.2.2. Monitorización de los efectos positivos de los fármacos

¿Cómo pueden valorarse los efectos positivos de un fármaco que los ensayos clínicos precomercialización no han podido poner de manifiesto? Las estrategias empleadas pueden agruparse en tres grandes categorías:

1. Los métodos basados en la aplicación de los principios y métodos del ensayo clínico aplicados a la fase de postcomercialización.
2. El meta-análisis.
3. Los métodos basados en los estudios observacionales.

Métodos basados en el ensayo clínico

Para algunos fármacos el problema esencial en la valoración de su eficacia es la necesidad de tiempos prolongados de observación.

Peto et al.⁽⁶³⁾ han descrito de forma precisa y completa los métodos adecuados para realizar ensayos clínicos en estos casos, como por ejemplo en la valoración de los fármacos citostáticos, en los que los puntos de valoración importantes son los tiempos de remisión y supervivencia.

Por otra parte, la realización de ensayos clínicos con un número elevado de pacientes es en muchos casos factible y algunas veces muy necesario. Existen

enfermedades comunes, en los que los tratamientos a ensayar son ampliamente practicables, los endpoints, importantes en la evaluación del tratamiento, son simples de medir y en los que los resultados pueden ser aplicables a un número importante de pacientes, especialmente si, como decíamos, la intervención terapéutica es generalizada en las condiciones de la práctica habitual.

Actualmente son ya absolutamente habituales los estudios que están demostrando la utilidad y factibilidad de estos ensayos clínicos “poblacionales”. Vamos a tomar como ejemplo el clásico estudio del GISSI. Los ensayos clínicos de valoración de la estreptoquinasa como tratamiento trombolítico en el infarto fueron, en su mayoría, incapaces de demostrar efectos en la mortalidad de forma significativa. Pero también es cierto que ninguno de ellos estudió a un número de pacientes que llegara al millar. Yusuf et al.⁽⁶⁴⁾ empleando la técnica del meta-análisis (véase más adelante), concluyeron que la estreptokinasa reducía de forma significativa la mortalidad post-infarto ($22 \pm 5\%$, $p < 0,001$).

En base a todos estos resultados el Grupo Italiano per la Sperimentazione della Streptochinasa nell'Infarto (GISSI)⁽⁶⁵⁾. Mediante un diseño muy simple, se pudo incluir a 11.812 pacientes en el estudio, randomizados a: estreptoquinasa o a ningún tratamiento, permitiendo que la única cuestión que variaba era la randomización de los pacientes al tratamiento con el fibrinolítico o no. El resto de práctica clínica era igual que la que se practicaba de rutina en las distintas unidades coronarias, con lo que se realizaba el ensayo en condiciones reales de la práctica habitual. El estudio permitió de-

mostrar una reducción de la mortalidad post-infarto del 20% ($p < 0,001$). La contribución que el GISSI representaba era la capacidad de trasladar el método del ensayo clínico a un contexto cultural y con el suficiente poder estadístico como para obtener un parámetro de eficacia clínica altamente relevante.

Meta-análisis

El término meta-análisis⁽³⁶⁻⁴¹⁾ fue introducido por GU Glass en 1976 y se refiere al conjunto de técnicas que combinan de forma sistemática los resultados de diversos estudios sobre un mismo tema. Esta técnica había sido utilizada antes que en el campo de la biomedicina en otras disciplinas científicas como la psicología, sociología y la física, con el fin de aprovechar al máximo la información disponible. En biomedicina, no sólo se ha aplicado al ensayo clínico farmacológico, sino que también se ha utilizado en ensayos de otras intervenciones sanitarias, tales como las intervenciones educativas dirigidas a pacientes y también a estudios observacionales.

La diferencia esencial entre el meta-análisis y las revisiones tradicionales radica en que en el meta-análisis se valora con más rigor y sistemática la contribución ponderada de cada estudio, así como su consistencia interna. Estas clásicas revisiones han consistido en un análisis crítico con resumen de los resultados y conclusiones o en la contabilización de los resultados positivos y negativos estadísticamente significativos y, a partir del cálculo de su diferencia, extracción de conclusiones. De hecho el meta-análisis es como un “ensayo clínico de ensayos clínicos”, en el sentido de que se aplica el mismo método que en un ensayo clínico:

1. Definición del objetivo del meta-análisis y de los puntos de valoración.
2. Selección de los estudios con establecimiento de criterios de inclusión y exclusión, que implica la definición precisa del tipo de pacientes y tratamientos, los criterios de calidad interna de los estudios –validez, precisión, etc.– y de los criterios de consistencia clínica de la combinación de los resultados. En segundo lugar, revisión bibliográfica de todos los estudios publicados y selección según los criterios establecidos.
3. Establecimiento del método de análisis de los datos, los datos necesarios, la necesidad o no de realizar un análisis estratificado y del tipo de análisis estadístico a realizar.

4. Definición de la forma de obtención de los datos, incluyendo la valoración de la calidad interna de, la obtención de los datos de cada estudio.
5. Aplicación de las técnicas de análisis estadístico de los resultados combinados, incluyendo –en la mayoría de casos– test de homogeneidad de los resultados procedentes de los distintos estudios y contabilización del peso de cada uno de los estudios.

La extensión de este capítulo no permite la descripción de las principales técnicas empleadas en este tipo de análisis que, por otra parte, se pueden encontrar en la bibliografía sugerida. Las limitaciones de esta técnica radican fundamentalmente en la dificultad de obtener la información necesaria, que muchas veces no es publicada totalmente, y en la propia debilidad y dificultad de la combinación de resultados procedentes de estudios de diversa índole, con poblaciones distintas, tratamientos no exactamente iguales, distintos niveles de error muestral, etc. Actualmente existe una cierta corriente crítica acerca del abuso de esta técnica, y uno de los principales argumentos de crítica se basa en la dificultad de asegurar la homogeneidad estadística de la información que se maneja⁽⁴²⁾. Sin embargo, la contribución del meta-análisis en la respuesta a cuestiones controvertidas de la terapéutica puede considerarse positiva, especialmente si permite seleccionar aquellas cuestiones que merecen ser estudiadas con otro tipo de método como puede ser un ensayo clínico poblacional.

Métodos basados en los estudios observacionales

Básicamente consiste en aplicar los métodos de los estudios de cohortes y de los estudios de casos y controles a la determinación de la eficacia a largo plazo de los fármacos utilizados en el contexto real. Sin embargo, en estos casos y dependiendo del tipo de efectos en los que se debe valorar la eficacia del fármaco, los estudios epidemiológicos están sujetos a importantes problemas de validez, en los que la confusión por la indicación es uno de los elementos más importantes. Los estudios de cohortes ofrecen asimismo dificultades adicionales en razón de su costo y de los sesgos potenciales que un largo periodo de seguimiento implica. Por ejemplo, a lo largo del periodo de seguimiento pueden haberse introducido nuevas técnicas de diagnóstico o de atención que modifiquen el curso de los resultados sin que ello dependa de la eficacia del tratamiento sujeto a valoración. Se ha propuesto que la metodología

caso-control podría aportar ciertas ventajas, como ya ha aportado a la valoración de los efectos adversos, por su mayor eficiencia estadística y menor costo. Así, por ejemplo, se podría valorar la eficacia relativa de distintos tratamientos, mediante la asignación al grupo “casos” de los fallos de tratamiento, mientras que los controles serían aquellos pacientes que si han respondido a los tratamientos, para determinar en cada grupo las frecuencias de “exposición” a distintos tratamientos y prácticas⁽⁴³⁾. El problema principal en estos estudios es el disponer de adecuada información clínica previa para determinar retrospectivamente la exposición a los distintos tratamientos.

En algunos –seguramente muy pocos– casos, la utilización de los registros e información clínica y los datos de consumo de determinados fármacos o las técnicas de conexión de la información clínica de determinados grupos poblacionales y del registro de medicación tomada –record-linkage– pueden aportar información sobre la eficacia de determinados tratamientos, tal como también se intenta en la monitorización de los efectos adversos, pero los problemas de control de factores de confusión y otros tipos de sesgos son muy importantes. Sin embargo, Strom et al.⁽⁴⁴⁾ han propuesto unos criterios muy claros de cuando la eficacia de los medicamentos puede ser utilizada sin emplear el método experimental, sin que ello comporte problemas de validez a causa de la existencia de factores de confusión o de la necesidad de disponer de un grupo control. Por ejemplo, no sería necesario el método experimental cuando el efecto terapéutico es obvio, cuando la confusión por la indicación es controlable porque la indicación es dicótoma (la inmuno-globulina Anti-Rh (D) en madres Rh negativas, etc.

3.2.3. Vigilancia de las reacciones adversas a los medicamentos (RAM)

Una vez se ha cumplido la fase de investigación clínica precomercialización, las RAM, poco frecuentes pero clínicamente importantes no han podido ser detectadas, si no es por azar, entre otras razones por cuestión de números. Como se ha comentado anteriormente, el número de pacientes expuestos a un fármaco en la fase de estudios clínicos precomercialización puede oscilar entre 2.000 y 5.000. Por otra parte, para detectar con un

nivel de confianza del 95% una RAM que se produzca con una frecuencia de 1/10.000, necesitamos estudiar al menos 30.000 individuos expuestos. La Tabla 5 muestra los números de pacientes necesarios para detectar RAM de distintas frecuencias. Además, ya se ha dicho que los pacientes enrolados en los estudios precomercialización son seleccionados de forma muy sesgada en relación a los pacientes que recibirán el fármaco posteriormente.

Los sistemas de vigilancia postcomercialización, para alcanzar estos objetivos se han desarrollado distintas estrategias y metodologías, de las cuales ninguna es suficientemente completa, pero entre las que existe cierto grado de complementariedad. De hecho, el método paradigmático es el ensayo clínico controlado, pero pocas veces es factible. Como los métodos de farmacovigilancia ya son descritos en otro capítulo, nos limitaremos en éste a hacer algunas consideraciones de tipo epidemiológico sobre los mismos.

Evidentemente, los métodos de farmacovigilancia están orientados a resolver dos tipos de cuestiones distintas pero complementarias. Por una parte, se trata de disponer de un sistema de alerta que facilite la detección rápida de las RAM no detectadas en la experimentación clínica precomercialización, es decir que sea capaz de detectar dichas RAM en los pacientes que reciben el fármaco, en las fases iniciales de su comercialización. Por otra parte, son necesarios métodos que sean capaces de confirmar la relación causal de una RAM, así como de su probabilidad de aparición. En general, a mayor capacidad de detección de un método, menor capacidad de comprobación etiológica y de proporcionar datos de frecuencia válidos.

La comunicación presenta intrínsecamente ciertas limitaciones que tienen que ver con la homogeneidad de criterios en la obtención de la información. Para ello se han publicado unas normas de la información mínima que se deben aportar en las comunicaciones en la literatura médica⁽⁴⁶⁾, así como se han establecido métodos que aseguren la aplicación de los mismos criterios y la misma secuencia de razonamiento (ver capítulo de Farmacovigilancia). Con tales métodos –algoritmos– se puede reducir la variabilidad interobservador es decir: mejorar la reproductibilidad. Pero de todas formas, estos métodos no mejoran mucho su validez –concordancia con la realidad–. Algunos autores⁽⁴⁷⁾ están desarrollando la aplicación del teorema de Bayes en la atribución causalidad de una sospecha de RAM para

Tabla 5. Probabilidad de observar una RAM, según su frecuencia real y el número de pacientes del estudio.

Frecuencia real	1/100	1/500	1/1.000	1/5.000	1/10.000	1/50.000
n° de pacientes	Probabilidad de observar la RAM					
100	0,63	0,18	0,10	0,02	0,01	0,002
200	0,86	0,33	0,18	0,04	0,02	0,004
500	0,99	0,63	0,39	0,10	0,05	0,01
1.000	0,99	0,86	0,63	0,18	0,10	0,02
2.000	0,99	0,98	0,86	0,33	0,18	0,04
5.000	0,99	0,99	0,99	0,63	0,39	0,10
10.000	0,99	0,99	0,99	0,86	0,63	0,18
N° necesario de pacientes para observar una RAM	300	1.500	3.000	15.000	30.000	150.000

(De Sackett DL.⁽⁴⁵⁾)

(Nota: como ejemplos de RAM con frecuencias como las citadas, el linfoma por azatriopina se produce con una frecuencia de 1 de cada 500, la toxicidad ocular por practolol en 1 de cada 1.000, el IAM en mujeres por anticonceptivos orales en 1 cada 5.000, la anafilaxis por penicilina en 1 cada 10.000 y la anemia aplásica por cloramfenicol en 1 cada 50.000).

mejorar este aspecto.

El sistema PEM –Monitorización de Acontecimientos ligados a la Prescripción–, es un sistema adecuado de detección de RAM y especialmente útil en los fármacos recientemente comercializados aunque es poco eficaz en la detección de eventos raros (1/10.000). Como en realidad el PEM es un estudio de cohortes en el que se pueden medir numeradores y denominadores, aunque con algunas limitaciones de validez, puede aportar cierta información cuantitativa con mucha más facilidad y validez que otros sistemas de detección.

Otro método que se puede considerar como una derivación de un estudio de cohortes es la conexión de registros de diagnóstico y prescripción o conexión de archivos clínicos (record-linkage), que tiene ciertas ventajas teóricas pero que no ha sido, hasta el momento, suficientemente eficaz en la detección de RAM y presenta ciertas debilidades como

son la inexistencia de información sobre determinados factores potenciales de confusión –por ejemplo, a través de uno de estos sistemas se encontró una relación entre el uso de metronidazol y el aborto⁽⁴⁸⁾, información potencialmente poco válida de los diagnósticos, no información sobre exposición durante los ingresos hospitalarios, etc. En realidad, ha existido una cierta polémica sobre la utilidad y debilidades de estos métodos, aunque la calidad de algunos estudios realizados mediante, la utilización de estas bases de datos ha mejorado, especialmente en aquellos estudios que han utilizado la Base de Datos de Saskatchewan en Canadá que permiten formar cohortes de expuestos y no expuestos e incluso una adecuada selección randomizada de controles de la misma población y también realizar estudios de casos, y controles con casos prevalentes o incidentes⁽⁴⁹⁻⁵¹⁾.

Los registros de estadísticas vitales y de morbi-

lidad (cáncer, malformaciones, etc.) pueden aportar señales de alerta, en cuanto que permite observar los cambios de incidencia o prevalencia de ciertos problemas de salud, que conduzcan a la sospecha de causalidad atribuible a la terapia farmacológica. Uno de los peligros de estos estudios de correlación es la probabilidad de que se observen asociaciones en grupos que no sean ciertas para los individuos –falacia ecológica–, sino que se deban, por ejemplo, al azar, ya que la información sobre la exposición no se ha obtenido a partir de los individuos, sino a partir de datos crudos de los grupos –cifras de ventas globales–. Una forma de resolver esta cuestión es la de utilizar los datos individuales de los registros para realizar estudios de casos y controles ad-hoc, con casos prevalentes y, de esta forma, extender la utilidad de alerta de estos registros a la posibilidad de comprobar hipótesis etiológicas. Este método, por ejemplo, se utilizó en el caso de la enfermedad tromboembólica y el uso de anticonceptivos orales, mencionado anteriormente⁽⁵²⁾.

El estudio postcomercialización de la cimetidina llevado a cabo en el Reino Unido por Colin-Jones *et al.*⁽⁵³⁾, es un ejemplo de la debilidad de este tipo de estudios frente a los estudios experimentales. Los autores encontraron un aumento de la mortalidad en los pacientes que recibían cimetidina, pero este fármaco tendía a ser usado selectivamente en pacientes con enfermedades subyacentes (angor) que aumentaban de por sí el riesgo de muerte, o por pacientes con supuestos diagnósticos de úlcera u otros trastornos GI menores que, en realidad, sufrían los primeros síntomas de un cáncer gastrointestinal. Un ensayo clínico randomizado hubiera sido necesario para controlar estos factores de confusión y sesgos de selección (clasificación).

En general, el diseño de estudios de cohortes, que presenta la ventaja de poder estudiar fármacos con bajas tasas de exposición y de poder facilitar datos de incidencia y riesgo relativo, presenta grandes inconvenientes por su coste especialmente cuando se necesitan largos periodos de observación –y por su relativa eficiencia estadística. Existen algunas excepciones, como la ya comentada del sistema inglés del PEM y como los estudios de cohortes aplicadas a determinadas situaciones en las que el uso de fármacos en una proporción importante y específica de la población, cohortes el caso de los anticonceptivos orales, en que se han podido

estudiar con relativa facilidad cohortes extensas de mujeres expuestas^(54,55).

Los estudios intensivos en pacientes hospitalizados, a pesar de las ventajas teóricas de la información clínica disponible, de la posibilidad de obtener estimaciones de incidencia y de poder realizar estudios de casos y controles ad-hoc, han proporcionado pocos resultados, por lo costoso del sistema. Además, existen ciertas dudas sobre los potenciales factores de confusión y sesgos de información, de la misma manera que debe considerarse el grado de representatividad de la población estudiada⁽⁵⁶⁾.

La aplicación del diseño de estudios de casos y controles a la comprobación de hipótesis etiológicas en la investigación de RAM se ha mostrado eficiente, por lo que se ha intentado adaptar y mejorar su aplicabilidad a este campo. Las ventajas y limitaciones de este tipo de diseños ya se han discutido anteriormente. Ejemplo de su eficiencia es el clásico estudio, pequeño pero definitivo, de 8 casos y 32 controles para establecer la relación entre el tratamiento de mujeres con dietiestilbesterol y el adenocarcinoma de vagina en sus hijas⁽⁵⁷⁾. Estos estudios no sólo son eficaces en comprobar hipótesis, sino que han, demostrado su utilidad en la detección de nuevas RAM.

BIBLIOGRAFÍA

1. Spitzer WO. Clinical Epidemiology (editorial). *J. Chron' Dis*, 1988; 39:41 I-S.
2. Spitzer WO. Drugs as determinants of health and disease; in the population. An orientation to the bridge science of pharmacoepidemiology. *Clin Epidemiol*, 1991; 44:823-30.
3. Porta H, Hartsema AG. The contribution of epidemiology A contribution to the study of drugs. *Drug InteR Clin Pharm*. 1987; 21:121-4.
4. Álvarez-Dardet C, Bolúmar F, Porta Serró H. La medición de la frecuencia de la enfermedad. *Med Clin (Barc)*, 1987; 88:287-291.
5. Barker DJP, Rose GA. Epidemiología en la práctica médica. Salvat. Barcelona, 1983.
6. Frick MH, Elo O, Haapa H, *et al.* Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged with dislipdemia: safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N Engl J Med*, 1987; 317:1237-45.
7. Brett AS. Traeting Hypercholesterolemia: How should practicing physicians interpret the published data for patients *N Engl J Med*, 1989; 321:676-80.

8. Porta Serra M, Álvarez-Dardet C, Bolúmar F, et al. La calidad de la información clínica: validez. *Med Clin* 1988; 89:741-7.
9. Álvarez-Dardet C, Bolúmar F, Porta Serra M. Tipos de estudios *Med Clin* 1987; 89:296-301.
10. Lilienfeld AM, Lilienfeld DE. *Fundamentos de epidemiología*. México: Fondo Educativo Interamericano, 1983.
11. Miettinen OS. Estimability and estimation in case-referent studies. *Am J. Epidemiol.* 1976; 103:226-35.
12. Rothman KJ. *Epidemiología Moderna*. págs. 160-97. Díaz de Santos. Madrid, 1987.
13. Kleinbaum DG, Kupper LL, Morgensten H. *Epidemiological Research. Principles and Quantitative Methods*. Belmont CA. Lifetime Learning Publications, 1982.
14. Altimiras J, Borrás JM. Conceptos de epidemiología. Algunas aplicaciones al campo de los medicamentos (III). Calidad de los estudios: validez. *Farm Clin*, 1989; 6:158-70.
15. Guallar EM. Alternativas al diseño tradicional de los estudios de cohortes: los estudios de casos y controles híbridos y los estudios de cohortes y casos. En: Álvarez-Dardet C, Porta M eds. *Revisiones en Salud Pública*. vol. 2. Masson. Barcelona, 1991.
16. Tognoni G. *Drug Use and Monitoring*, en *Evaluation of Health Care*. WW Holland edit. Oxford: Oxford University Press, 1983.
17. Laporte JR, Tognoni G. *Principios de Epidemiología del Medicamento*. 2ª ed. Masson-Salvat. Barcelona, 1993.
18. Cannon P. Prevalence of angina as assessed by a survey of prescriptions for nitrates. *The Lancet*, 1988; 30:979-81.
19. Grupo EUM de la SEFH (coordinador: J. Altimiras). Estimación de la dosis diaria prescrita de antiinfecciosos en hospitales españoles. *Farm Hosp* 1994; 18:15-21.
20. Cuña B, González M, Simó RM, Villares MC. Utilización de antibióticos expresados en dosis diarias definidas por 100 camas día en 24 hospitales durante el año 1984. *Rev AEFH*, 1987; XI:221-36.
21. Jiménez V, Recalde JM, Arias A, et al. Utilización de antibióticos en 17 hospitales andaluces. *Farm Clin*, 1987; 4:15968.
22. Gallastegui C, Montoro JB, Oliveras M. Análisis de utilización de medicamentos en cinco hospitales generales de Barcelona. *Farm Clin*, 1987; 4:442-53.
23. Cano SM, Lalueza MP, San Juan N, et al. Análisis de utilización de medicamentos en cinco hospitales generales de Barcelona (II): antibióticos beta-lactámicos. *Farm Clin*, 1987; 4:546-52.
24. Carreras MJ, Codina C, Pumarola D, et al. Análisis de utilización de medicamentos en cinco hospitales generales de Barcelona (III): antibióticos no betalactámicos de uso sistémico. *Farm Clin*, 1987; 4:644-50.
25. García B, Gea E, Gachs M, et al. Análisis de utilización de medicamentos en cinco hospitales generales de Barcelona (psicofármacos). *Farm Clin*, 1987; 4:73343.
26. Anónimo. *Guidelines for ATC Classification*. Oslo: Nordic Collaborating Centre for Drug Statistics Methodology, 1991.
27. Anónimo. *Guidelines for DDD*. Oslo: Nordic Council on Medicines. WHO Collaborating Centre for Drug Statistics Methodology, 1991.
28. Rams N, Artal A, Carretero M, et al. Consumo de antiinfecciosos 86-89. Influencia en los resultados de la utilización de la DDD o de la DDR. *Farmacia Hospitalaria*, 1990; 14:230-35.
29. Laporte JR, Porta M, Capella D. Drug utilization studies: a tool for determining the effectiveness of drug use. *Br J. Clin Pharmacol*, 1983; 16:301-7.
30. Altimiras J, Duque A, Pastor E. Pharmacoepidemiological approach to improve drug therapy: Use of ranitidine in prophylaxis of gastroduodenal damage. *Proceedings of European Meeting on Pharmacoepidemiology*. Verona: European Society of Clinical Pharmacy, 1992.
31. Meylers AS. Adverse events: identification and attribution. *Drug Intell Clin Pharm*, 1987; 21:915-20.
32. Porta Serra M. Los ensayos clínicos en fase IV: por qué, cuándo y cómo. En: García Alonso F, Bakke OM, eds. *Metodología del Ensayo Clínico*. Fundación Dr. Antonio Esteve/Ediciones Doyma. Barcelona, 1991; 31-39.
33. Peto R. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. (I, II) *Br J. Cancer*; 1976; 34:585-612, 1977; 35:1-39.
34. Yusuf S, Collins R, Peto R. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results of mortality reinfarction and side effects from 33 randomized controlled trials. *Europ Heart J*, 1985; 6:556-85.
35. GISSI. Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *The Lan-*

- cet, 1986; 1:397-401.
36. Jeniek M. Meta-Analysis in Medicine. *J Clin Epidemiol*, 1989; 42:35-44.
 37. Sacks HS, Berrier J, Rietman D, et al. Meta-Analyses of randomized Controlled Trials *N Engl J Med*, 1987; 316:450-455.
 38. Einarsson TR, Leeder JS, Koren G. A Method for MetaAnalysis of epidemiologic studies *D Intell Clin Phann*, 1989; 22:813-823.
 39. De Simonian R, Laird N. Meta-Analysis in Clinical Trials. *Contr. Clin Trials*, 1986; 7:177-88.
 40. Collins R, Yusuf S, Peto R. Overview of Randomized Trials of Diuretics in Pregnancy. *B M J*, 1985; 290:17-23.
 41. Fernández-Ballart JD, Vobecky J, Martí-Henneberg C. Metaanálisis: síntesis e integración de los resultados de estudios independientes en medicina. *Med Clin* 1991; 96:382-7.
 42. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *The Lancet*, 1991; 338: I 127-30.
 43. Kasilo O, Devoto MA, Marchioli R, et al. Very large scale trials and surveillance networks: towards unified strategies. En: *post-marketing Surveillance of Drugs*. Garatini S, ed. Milan: Wicting Editore, 1987.
 44. Strom BL, Miettinen OL, Melmon KL. Post-marketing studies on drug efficacy: when must they be randomized? *Clin Pharmacol and Therap*, 1983; 94:1-7.
 45. Sakett DL. Compliance. En: *Inman WHW, ed. Monitoring for Drug Safety 6th edition*. Lancaster: MRT Press Limited, 478. 1985.
 46. Abrutin E. Better reporting of adverse drug reactions: *Aran Intern Med*, 1985; 102:264-65.
 47. Lane DA, Hutchinson TA. The future of adverse drug reaction diagnosis: computers, clinical judgement and the logic of uncertainty. *Proceedings of DIA Workshop. Drug Infor* 1, 1986; 20:455-533.
 48. Kegg DCG, Doll R. Record-linkage for drug monitoring. *J. Epidemio Comm Health*, 1981; 35:25-31.
 49. Strom BL, ed. 1989, opp. cit. 149-243.
 50. Hernández M, Walker AM, Jick H. Use of replacement estrogens and the risk of myocardial infarction. *Epidemiology*. 1990; 1:128-33.
 51. Moskowitz MA, Jick SS, Burnside S, et al. The relationship of oral contraceptive use to rheumatoid arthritis. *Epidemiology*, 1990; 1:153-6.
 52. Jick F. Venous thromboembolic disease and ABO blood type. *The Lancet*, 1979; 1:539-42.
 53. Colin-Jones DG. Postmarketing surveillance of the safety of cimetidine: mortality during second, third and fourth years of follow-up. *B M J*, 1989; 291: 1084-88.
 54. Royal College of General Practitioners. *Oral contraceptives and health*. London: Pitman Medical. 1974.
 55. Vessey M, Doll R, Peto R, et al. A long-term follow-up women using different methods of contraception, an interrim report. *J. Bio Sci*, 1976; 8:373-427.
 56. Ibáñez L, Laporte JR, Carné X. Adverse drug reactions leading to hospital admission. *Drug Safety*, 1991; 6:450-59.
 57. Herbst AL, Ulfeldertt R, Postranzer DC. Adenocarcinoma of vagina. Association of maternal stilbestrol therapy and tumor appearance in young women. *N Engl J Med*, 1997; 284:878-81.

Referencias generales

1. Epidemiología

- Last JM. *Diccionario de epidemiología*. Salvat. Barcelona, 1989.
- Kramer MS. *Clinical Epidemiology arad Biostatistics*. Springer-Verlag, Springer-Verlag, 1988.
- Feinstein AR. *Clinical Epidemiology. The architecture of clinical research*. Philadelphia PA. WB Saunders, 1985.
- Fleiss JL. *Statistical methods for restes arad proportions*. John Wiley. New York, 1981.
- Fletcher RH, Fletcher SW, Wagner EH. *Epidemiología clínica*. Ediciones Consulta. Barcelona. 1989.
- Jenicek M, Cleroux R. *Epidemiología. Principios, técnicas, aplicaciones*. Salvat. Barcelona, 1987.
- Sackett DL, Haynes B, Tugwell P. *Epidemiología Clínica. Una ciencia básica para la medicina clínica*. Díaz de Santos. Madrid. 1989.

2. Farmacoepidemiología: textos generales

- Colombo F, Shapiro S, Slone D, Tognoni G, eds. *Epidemiological Evaluation of Drugs*. Amsterdam: Elsevier] North-Holland Biomedical Press, 1977.
- Tilson HH, ed. *Progress in Pharmacoepidemiology*. Arn J; *Prev Med*, 1986; 4: part 2/supl.
- Miettinen O. Efficacy on therapeutic practice: will epidemiology give the answers? En: *Drug Therapeutics. Concepts for Physicians*. Churchill-Livingstone. Edimburgh, 1981.

- Strom BL, Miettinen OS, Melmon KL. Postmarketing studies of drug efficacy: When must they be randomized? *Clin Pharmacol Therap*, 1983; 34:1-7.
 - Strom BL, Miettinen OS, Melmon KL. Postmarketing studies of drug efficacy: Why? *Am J Med*, 1985; 78:475-80.
 - Strom B, Miettinen OS, Melmon KL. Postmarketing studies of drug efficacy: How? *Am J Med*, 1984; 77: 703-8.
 - Kewitz H, Roots I, Voigt K, eds. *Epidemiological Concepts in Clinical Pharmacology*. Springer-Verlag: Springer-Verlag, 1987.
 - LLunde PKM, Agenás I, eds. *Drug Utilization in Relation to Morbidity*. Stockholm: Acta Medica Scandinavica, 1988.
 - Strom BL. ed. *Pharmacoepidemiology*. New York: Churchill Livingstone, 1989.
 - Hartzema AG, Porta MS, Tilson HH, eds. *Pharmacoepidemiology. An Introduction*, 2nd edition. Cincinnati OH. Harvey Whitney Books, 1991.
 - Altimiras J, Figueras M, Salvat S, et al. *Curso Modular de Farmacoepidemiología y Farmacoeconomía para Farmacéuticos de Hospital (1-5)*. Editores Médicos, S.A. 1997.
 - Monte Boquet E. *Farmacoepidemiología en internet*. *El Farmacéutico de Hospitales* 2001; 123: 44-48.
 - Hidalgo E. *Farmacoepidemiología. Últimas novedades bibliográficas*. *El Farmacéutico de Hospitales*. Jul 2001; 123.
3. *Interpretación de los estudios en la toma de decisiones*
- Jiménez Villa J. *Lectura crítica de la literatura científica II. Evaluación de los resultados*. FMC Formación Médica Continuada en Atención Primaria. 2000, 7, 5: 283-296.
 - Jovell AJ, Navarro-Rubio MD. *Evaluación de la evidencia científica*. *Med Clin (Barc)* 1995; 105:740-43.
 - Ioannidis JPA, Lau J. *Can quality of clinical trials and meta-analyses be quantified?* *Lancet* 1998; 352: 590-91.
 - Berlin JA, Rennie D. *Measuring the quality of trials. The quality of quality scales*. *JAMA* 1999, 282, 11: 1083-84.
 - Juni P et al. *The hazards of scoring the quality of clinical trials for meta-analysis*. *JAMA* 1999; 282: 1054-60.
 - Barton S. *Which clinical studies provide the best evidence*. *BMJ* 2000; 321:255-56.
 - Anónimo. *El cuestionario CONSORT: requisitos para la publicación de ensayos clínicos*. *Investigación Clínica y Bioética*. 1997; 22:5-8.
 - Sacristán JA, Soto J, Galende I. *Evaluación crítica de ensayos clínicos*. *Med Clin Barc* 1993; 100: 780-787.
 - Guyatt GH, Sackett DL, Cook DJ. *User's guides to the medical literature. II How to use and article about therapy and prevention. A. Are the results of the study valid?* *JAMA* 1993, 270,21: 2598-2601.
 - Guyatt GH, Sackett DL, Cook DJ. *User's guides to the medical literature. II How to use and article about therapy and prevention. B. What were the results and will they help me in caring for my patients?* *JAMA* 1994, 271,1:59-63.
 - Altman DG. *Better reporting of randomized controlled trials: the COONSORT statement*. *BMJ* 1996; 313: 570-571.
 - Puigventós F. *Novedades terapéuticas año 2000: los retos del médico y del gestor*. *Med Clin* 2001, 116, 12:465-469 1997; 315:672-675.
 - Mc Alister FA, Straus SF, Guyatt GH, Brian Haynes R. *User's guides to the medical literature. Integrating research evidence with the care of the individual patient*. *JAMA*, 2000; 283:2829-2836.
 - Sackett DL, Scott Richardson W, Rosenberg W, Brian Haynes R. *Medicina basada en la evidencia, cómo ejercer y enseñar MBE*. Churchill Communications Europe España ed. Madrid, 1997.
 - Oxman Ad, Cook DJ, Guyatt AH. *User's guides to the medical literature. VI How to use an overview*. *JAMA*, 1994; 272, 17: 1367-1371.
 - Smeeth L, Haines A, Ebrahim S. *Numbers needed tot treat derived from meta-analyses-sometimes informative, usually misleading*. *BMJ* 1999; 318:1548-51.
 - López-Arrieta JM, Qizilbash N: *La medicina basada en pruebas: revisiones sistemáticas*. La Colaboración Cochrane. *Med Clin* 1996; 107:581-585.
 - Greenhalgh T. *How to read a paper: papers that summarise other papers*, *BMJ* 1997; 315:540-3.
 - Cook RJ, Sackett DL. *The number needed to treat: a clinically useful measure of treatment effect*. *BMJ* 1995; 310:452-454.
 - Martín J, Martínez M, Cantero P. *El número de pacientes que será necesario tratar: una herramienta de interés para el clínico*. *Med Clin* 1999; 113: 156-7.
 - Smeeth L, Haines A, Ebrahim S. *Numbers needed to treat derived from meta-analyses sometimes infor-*

- mative, usually misleading. *BMJ* 1999; 318: 1548-1551.
- Heffner JE. Does evidence-based medicine help the development of clinical practice guidelines? *Chest* 1998; 113:172s-178s.
 - Feder G, Eccles M, Grol R, Griffiths C, et al. Using clinical guidelines. *BMJ* 1999; 318:728-730.
 - Sheldon TA, Guyatt GH, Haines A. Getting research findings into practice, When to act on the evidence. *BMJ* 1998; 317:139-142.
 - Woolf SH, Grol R, Hutchinson A, et al. Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999; 318:327-330.
 - Norheim OF. Healthcare rationing-are additional criteria needed for assessing evidence based clinical practice guidelines? *BMJ* 1999; 319:1426-1419.
4. *Estudios de utilización de medicamentos*
- Arnau JM. Estudios de utilización de medicamentos. En: *Farmacoepidemiología*, dirigida por L Matos. Santiago de Compostela: Xunta de Galicia, 1995: 191-204.
 - García R. Introducción a la metodología de los estudios de utilización de medicamentos. En: *Farmacoepidemiología*, dirigida por L Matos. Santiago de Compostela: Xunta de Galicia, 1995: 205-221.
 - García B, Gallastegui C, Altamiras J, et al. Prescripción de medicamentos no incluidos en la Guía Farmacoterapéutica, como indicador de control de calidad. *Revista de la AEFH*, XII: 19-24. 1988.
 - Grupo EUM de la SEFH (coordinador: J. Altamiras). Estimación de la dosis diaria prescrita de antiinfecciosos en hospitales españoles. *Farm Hosp* 1994; 18:15-21.
 - Laporte JR, Tognoni G. Estudios de utilización de medicamentos y de farmacovigilancia. En: *Principios de epidemiología de medicamentos*. 2ª ed., dirigida por JR Laporte y G Tognoni. Masson-Salvat. Barcelona, 1993; 1-15.
 - Capellà D, Laporte JR. Métodos aplicados en estudios descriptivos de utilización de medicamentos. En: *Principios de epidemiología de medicamentos*. 2ª ed., dirigida por JR Laporte y G Tognoni. Masson-Salvat. Barcelona, 1993; 67-87.
 - Dukes MNG eds. *Drug utilization studies: Methods and uses*. WHO Regional Publications. European Series; n° 45, 1993.
 - Spooner CH, Pickard AS, Menon D. Edmonton Quality Assessment Tool for Drug Utilization Reviews. EQUATDUR-2: the development of a scale to assess the methodological quality of a drug utilization review. *Med Care* 2000 Sep; 38(9):948-58.
 - Bergman U, Popa C, Tomson Y, et al. Drug utilization A simple method for assessing the quality of drug prescribing. *Eur J Clin Pharmacol* 1998 Apr; 54(2):113-8.
 - Ministerio de Sanidad y Consumo. INSALUD. Sistema de codificación de principios activos y dosis diarias definidas del insalud. Madrid, enero 2001).
5. *Farmacoepidemiología, textos y artículos específicos*
- Light RJ, Pillmer DB. *Summing up. The science of reviewing research*. Cambridge: Harvard University Press, 1984.
 - Peto R, Gray R, Collins R. Randomized Trial of Prophylactic Aspirin in British Male Doctors. *BMJ*, 1988; 296:313-316.
 - UK-TIA Study Group UK-TIA aspirin trial: interim results. *BMJ* 1988; 96:316-320.
 - NHS Centre for reviews and dissemination: Undertaking systematic reviews of reseach on effectiveness, CRD report numebre 4 (2ª ed.), 2001.
 - Egger M, Davey G, Phillips AN: Meta-analysis: Principles and procedures. *BMJ*. 1997; 315:1533-37.
 - Pladevall M, Arana A, Varas C. Revisión sistemáticas y metaanálisis en cardiopatía isquémica. *Cardiovascular risk factors*. 2000; 9,4:270-281.
 - Delgado-Rodríguez M, Sillero-Arenas Miguel. Sesgos en el metaanálisis. *Med Clin* 1999; 112 (Supl 1):43-50.
 - Moher D, Cook DJ, Eastwood S et al. Mejora de la calidad de los informes de los metaanálisis de los ensayos clínicos controlados. El acuerdo QUORUM. *Revista española de Salud Pública*. 2000, 74, 2.
 - Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. *Lancet* 1998; 35:47-52.
 - Etmninan ME, Levine M. Interpreting meta-analyses of pharmacologic interventions: the pitfalls and how to identify them. *Pharmacotherapy* 1999; 19 (6):741-745.
 - Charlton BG: The uses and abuses of meta-analyses. *Family Practice* 1996; 13,4:397-401.
 - Cappelletri Jc, et al. Large trials vs meta-analysis of smaller trials. How do their results compare. *JAMA*, 1996, 276,16:1332-1338.

- Jadad AR et al Systematic reviews and met-analyses on treatment of asthma, critical evaluation BMJ 2000; 320:537-540.
- Medical Research Council Working Party: Stroke and Coronary Heart Disease in Mild Hypertension: risk factors and the use of treatment. BMJ, 1988; 296: 1565-70.
- Aspirin Myocardial Infarction Study Research Group: a Randomized Controlled Trial of Aspirin in Persons Recovered from Myocardial Infarction. JAMA, 1980; 243:661-9.
- Multiple Risk Factor Intervention Trial Research Group. Risk Factor Changes and Mortality Results. JAMA. 1987; 248:1465.
- Jones JC, Herman RL. The Future on Adverse Drug Reaction Diagnosis. Computers Clinical Judgement and the Logic of Uncertainty. Drug Inform 1986; 20,4.
- Irman WHW, ed. Monitoring for Drug Safety second ed. Lancaster: MTP Press Limited, 1985.
- H, Vessey MP. Case-Control Studies in the Evaluation Drug Induced Illness. Am J Epidemiol, 1978; 107:1-7.
- Vening GR. Identification of Adverse Reactions to New Drugs (I-IV) B. M. J., 286, 199-202, 289-292, 365-68, 458-60, 544.
- Shapiro S. The Epidemiological evaluation of Drugs Acta Scand, 1983; Supp 1683:23-27.
- Skegg DCG. Potential for Bias in Case-Control Studies of Oral Contraceptives and Breast Cancer. Am J Epidemiol, 1988; 127:205-211.
- Stolley PD, Strom BL. Evaluating and Monitoring the Safety and Efficacy of Drug Therapy and Surgery. J Chron Dis, 1988; 39:1145-55.
- Edlavitch SA. Postmarketing Surveillance Methodologies. Drug Intell Clin Pharm, 1988; 22:68-78.
- Smith-Rogers A, Porta M, Tilson HH. Guidelines for decision making in postmarketing surveillance of drugs. J Clin Res Pharmacop, 1990; 4:241-51.

