

# El uso del método epidemiológico para la evaluación de la calidad de la asistencia (II)

**Javier Moliner**

*Dirección Gerencia. Hospital Clínico Universitario Lozano Blesa. Zaragoza.*

Correspondencia

Dr. Javier Moliner.

Coordinador del Programa de Calidad.

Dirección Gerencia.

Hospital Clínico Universitario Lozano Blesa.

San Juan Bosco, 15. 50009 Zaragoza.

Correo electrónico: esta-moliner@hcu-lblesa.es

## Introducción

En la primera parte de este repaso del método epidemiológico y sus utilidades en la evaluación de calidad nos hemos centrado en los tipos de medidas en epidemiología, tanto de frecuencia como de asociación, y en las aplicaciones e interpretaciones de esas medidas.

La segunda parte la podemos dedicar a echar un vistazo a las estrategias para calcular esas medidas, es decir, los tipos de estudios epidemiológicos. Distinguiremos entre cinco clases de estudios: ecológicos, transversales, de seguimiento, de casos y controles y experimentales.

## Check-points al planificar un estudio

Cualquiera que sea el tipo de estudio elegido, en su diseño se busca conseguir la máxima validez, es decir, el grado en que el estudio mide el objetivo que se propone. La validez intrínseca o interna se logra cuando el estudio elige bien las personas a analizar y mide todos los datos necesarios de forma correcta. La validez interna se considera un requisito para que exista validez externa que, cuando se da, permite la extrapolación de los resultados del estudio a la población que no ha participado en el mismo.

Las principales amenazas para la validez interna son los errores sistemáticos o sesgos, fundamentalmente de tres tipos: de selección (los individuos elegidos o no elegidos para el estudio no son los adecuados), de clasificación o información (los datos de enfermedad o de factores causales no están bien recogidos) y la confusión (existe alguna variable que se interpone e impide cuantificar correctamente la relación real entre causa y enfermedad).

Junto con los sesgos, la otra fuente de error en la interpretación de los resultados de los estudios es la que produce el azar, el error aleatorio. Este error se produce siempre que se trabaja con muestras de casos en lugar de con la población total. La precisión de las medidas epidemiológicas se cuantifica con los intervalos de confianza, que indican, para

una determinada probabilidad de error, entre qué márgenes se encuentra realmente el dato que se ha calculado.

Para juzgar si el diseño de un estudio ha sido adecuado se puede establecer una lista de comprobación, con algunos "puntos críticos" que conviene repasar.

En primer lugar, hay que tener claros los objetivos. La primera pregunta que debe responderse de forma clara es ¿por qué vamos a hacer este estudio? Se debe contar al menos con una hipótesis o una pregunta de investigación formuladas de manera precisa.

Para medir y comparar es fundamental la "definición de caso". El problema que se quiere analizar, la enfermedad o el factor de riesgo deben estar definidos de la forma más precisa posible, buscando la repetibilidad del estudio y la comparabilidad de los resultados. Las definiciones de caso empleadas en los estudios de sucesos adversos hospitalarios son buenos ejemplos en este punto. En el estudio que citábamos en la primera parte, de García Martín et al<sup>1</sup>, se consideraba suceso adverso "un problema de cualquier naturaleza o gravedad experimentado por un paciente durante su estancia en el hospital y potencialmente atribuible al manejo clínico o administrativo, en lugar de a la enfermedad subyacente", y se encontró algún caso en más del 40% de los pacientes. Por el contrario, la definición más estricta de suceso adverso que empleó el Harvard Medical Practice Study<sup>2</sup>, "cualquier daño resultante del manejo médico que produce un alargamiento de la estancia o una discapacidad al alta", encontró estos casos en "sólo" el 4% de los pacientes.

Cabe aquí una referencia a los criterios para elaborar las definiciones. Para realizar mediciones se necesita una definición de enfermedad o de problema de calidad que permita clasificar a cada individuo como enfermo o como no enfermo, como "caso" o como "no caso". Sin embargo, tanto las enfermedades como los problemas de calidad sólo responden, en general, a preguntas del tipo "¿qué grado de enfermedad padece?". Aunque acabamos por fijar esos puntos de corte, umbrales para tratamiento, estándares o niveles de calidad aceptables, conviene no olvidar que con frecuencia estos niveles son sólo arbitrarios o de conveniencia. Otros criterios empleados a veces, como el basado en los datos históricos (por el

que la tasa de infección hospitalaria no debería ser superior al 8% o la de partos por cesárea al 20%) o el de los datos estadísticos o extrapolados (toda cifra de colesterol sanguíneo mayor de 200 mg/dl define a su poseedor como enfermo) no dejan de estar libres de arbitrariedad.

Elegir los instrumentos de medida es otro punto importante. Los métodos deben poder ser usados de manera estandarizada y, en lo posible, deben ser aceptables para las personas que participan, rápidos y baratos.

El entrenamiento de las personas que recogen datos, sobre todo cuando intervienen varias personas, es necesario para lograr la estandarización. Existen técnicas para minimizar el efecto de la participación de varios revisores, como repartir los casos de forma aleatoria, anotar quién ha sido el observador en cada caso y comparar sus resultados o ensayar con una prueba piloto antes de comenzar el estudio. En un epígrafe anterior sobre las medidas de concordancia entre observadores se ha repasado algunas formas de evaluar si el entrenamiento ha sido efectivo.

### **Estrategia de muestreo**

Para seleccionar la muestra más conveniente, normalmente será mejor si el método de selección es aleatorio, es decir si todos los individuos de la población tienen una probabilidad conocida (generalmente la misma) de ser seleccionados. En el muestreo aleatorio simple, cada miembro de la población debe tener un número seriado y se puede utilizar una lista de números aleatorios generada por un ordenador. Si se quiere analizar determinados subgrupos de población e interesa que esos subgrupos estén suficientemente representados, se puede recurrir al muestreo aleatorio estratificado, consistente en un muestreo simple dentro de cada subgrupo. Si por conveniencia se desea centrar un estudio en determinadas áreas (p. ej., geográficas), se puede plantear un muestreo polietápico, seleccionando primero una muestra de áreas y, después, dentro de cada una, realizar otro muestreo aleatorio.

### **Cálculos del tamaño de la muestra**

Con frecuencia, los estudios son más pequeños de lo que el investigador querría y, también con frecuencia, el investigador puede darse cuenta justo al final, cuando descubre que los números no permiten conseguir ni siquiera el objetivo principal del estudio. Para intentar evitarlo, si el objetivo del estudio está formulado en términos precisos, mediante fórmulas matemáticas o con uno de muchos programas informáticos (Epi Info es muy útil), se puede hacer fácilmente los cálculos del tamaño muestral necesario. Al programa informático hay que proporcionarle básicamente dos datos: tener una idea cuantificada de lo que estamos buscando y la cantidad de error que estamos dispuestos a aceptar por el hecho de trabajar con una muestra.

### **Selección de sujetos para el estudio**

En muchos estudios existe el riesgo de que las personas que se desea estudiar rechacen participar o que no se pueda

obtener datos de determinados casos. La respuesta a la pregunta "¿cuál es la tasa de respuesta aceptable?" es "depende", y depende del objetivo del estudio y de la población. Los problemas surgen cuando los no respondedores son atípicos, y entonces interesa saber cómo son de no representativos en relación con la pregunta estudiada. Para estudiar la tasa de respuesta y su relevancia se puede emplear dos enfoques: uno es analizar las variables disponibles sobre los no respondedores y comprobar si existen diferencias con los que sí respondieron, y otra posibilidad es tomar una pequeña muestra aleatoria de los no respondedores y tratar de obtener sus respuestas, para analizar la existencia de sesgos. Una tercera opción es analizar los datos "poniéndose en lo peor"; por ejemplo: asumiendo que todos los que no responden tienen la enfermedad, o que no la tiene ninguno, o que todos los que no han contestado a una encuesta lo hubieran hecho en un mismo sentido.

El sesgo de selección ocurre cuando los sujetos estudiados no son representativos de la población sobre la que se quiere generalizar las conclusiones, porque la forma de elegir los sujetos hace que unos tengan más probabilidades de participar que otros. La importancia de este sesgo debe ser evaluada a la hora de extrapolar los resultados. En ocasiones, el sesgo de selección puede ser provocado por el investigador como sucede, por ejemplo, al eliminar los fines de semana o los días festivos en los estudios de adecuación de estancias hospitalarias.

### **Análisis de los datos**

En la fase de diseño es conveniente tener una previsión de cómo será el análisis final de los datos, para lo que puede ayudar dibujar las tablas (sin datos, claro) del informe final. Con la hipótesis de trabajo y las variables a recoger, se puede predecir cuáles pueden ser los datos esperados en las principales celdas de esas tablas. El análisis estadístico depende del tipo de variables de que se dispone y del número de casos previsto. Esta información debe ser estimada en el diseño del estudio, por lo que el investigador debería indicar desde el principio qué pruebas estadísticas piensa utilizar. No se debe olvidar que consultar a alguien experto en técnicas estadísticas puede ser una buena inversión.

### **Tipos de estudios epidemiológicos**

Distinguiremos entre cinco clases de estudios epidemiológicos: ecológicos, transversales, de seguimiento o longitudinales, de casos y controles y experimentales.

La principal característica que une a los cuatro primeros tipos es que son estudios observacionales, en los que el investigador sólo mira, no influye en la cantidad de riesgo a que está sometido cada paciente. La diferencia con los estudios experimentales es que en éstos cada individuo es sometido a un tipo de exposición u otro, decidida según el diseño del experimento.

Otra forma de agrupar los tipos de estudios es según si su base de análisis son las poblaciones o los individuos. Los estu-

dios ecológicos y una parte de los experimentales se realizan sobre poblaciones. El resto son estudios de base individual.

### **Estudios ecológicos**

Los estudios ecológicos, o de correlación, se caracterizan porque comparan datos de poblaciones, no de individuos. Se miden las tasas de enfermedad en varias poblaciones y se analiza su relación.

Las poblaciones comparadas se pueden definir de varias maneras, aunque los estudios de series temporales (estudiar lo que ocurre en una o más poblaciones a lo largo del tiempo) y los de comparaciones geográficas (varias poblaciones en un solo momento de tiempo) son los más frecuentes.

Los estudios ecológicos informan sobre los patrones de distribución de las enfermedades. Son útiles para estudiar fenómenos ambientales, sociales o culturales en los que la unidad individual no es de interés y también para evaluar la efectividad de medidas de intervención sobre poblaciones.

Se da el caso de que algunas mediciones son más fiables, más válidas cuando se calculan a escala poblacional que a escala individual. Por ejemplo, en el caso del consumo de alcohol puede ser más precisa la medición en un estudio ecológico que en uno individual (se conocen las ventas de bebidas alcohólicas en una determinada zona, mientras que el consumo individual puede ser más difícil de conocer).

Este tipo de estudios debe servir para generar hipótesis que tendrán que comprobarse posteriormente con otro tipo de diseños, ya que una asociación encontrada al comparar poblaciones no garantiza que exista individualmente. Extraer conclusiones inapropiadas sobre causalidad de forma individual a partir de observaciones poblacionales es el error conocido como falacia ecológica.

La principal ventaja de los estudios ecológicos es que resultan baratos y rápidos de realizar (siempre que las fuentes de datos sean completas y estén disponibles), sobre todo si los datos se obtienen de otras estadísticas ya publicadas.

### **Series temporales**

Muchos problemas presentan fluctuaciones importantes con el tiempo. Las unidades ecológicas a comparar se definen utilizando criterios temporales (partos por cesárea cada mes), analizando cómo cambia la frecuencia de la enfermedad o relacionando estos cambios con las variaciones en el nivel medio de exposición a lo largo del tiempo (edad materna, presencia de determinados diagnósticos). Si las tendencias en la evolución de un problema se correlacionan con cambios en los supuestos factores de riesgo, su estudio puede proporcionar pistas para el análisis.

En los problemas "crónicos" resulta útil estudiar las tendencias en períodos de muchos años, aunque en la explicación de los resultados pueden influir muchos factores como el desarrollo tecnológico, los criterios diagnósticos, los cambios poblacionales, etc. Un factor que complica la investigación en los estudios temporales largos es que muchos datos antiguos no puedan ser comprobados ahora.

Un ejemplo de estudio de series temporales con datos de utilización de medicamentos ha sido recientemente publi-

cado por Jabary<sup>3</sup>, comparando el uso de antihipertensivos en España entre 1986 y 1994, y el cambio en el patrón de uso en relación con las recomendaciones científicas existentes. Los autores encuentran en ese tiempo un significativo aumento (del 117%) en el consumo de antihipertensivos (en número de dosis por cada 1.000 habitantes y año), con incrementos relevantes de inhibidores de la ECA y bloqueadores beta. Tras el análisis de las variables recogidas, concluyen que las recomendaciones nacionales e internacionales han tenido un efecto nulo en los patrones de prescripción.

### **Comparaciones geográficas**

En estos estudios se trata de comparar las tasas de enfermedad y de factores de riesgo entre países o entre regiones. Se puede analizar la relación entre pares de variables con fines exploratorios, o también se puede analizar un grupo de variables consideradas como explicativas y comparar su relación con una variable respuesta buscando elaborar un modelo predictivo de un efecto a partir de variables explicativas.

Un inconveniente en el análisis es que los estudios de correlación trabajan con valores medios y en la interpretación de los coeficientes de correlación se corre el riesgo de caer en el sesgo de agregación: al comparar entre varios grupos donde utilizamos el modelo de regresión lineal simple la variable independiente (x) que surge del agrupamiento nos estructura un modelo lineal que difiere del que surgiría si empleásemos la variable independiente individual; por ello, los resultados están sesgados y difieren de los que se podrían obtener en un análisis simple de un estudio de cohortes.

Existen muchos ejemplos recientes de estudios de comparación geográfica. Caicoya et al<sup>4</sup> analizan los valores poblacionales de distintas variables relacionadas con el uso de la diálisis en el territorio del antiguo Insalud. Se encuentra con variaciones importantes en las tasas de utilización entre hospitales y entre provincias y encuentra algunas variables correlacionadas con el uso de diálisis, como las que indican disponibilidad de recursos (especialistas en nefrología o puestos de diálisis), pero no con otras variables como la morbilidad atendida entre zonas geográficas.

En otro estudio, García Sempere y Peiró<sup>5</sup> analizan los datos de 170 zonas básicas de salud de la Comunidad Valenciana, comparando variables demográficas (tasas de defunción, nivel sociocultural), de renta y desarrollo, de oferta y de uso de servicios con el gasto farmacéutico. La correlación de cada variable con el gasto farmacéutico permite encontrar variables asociadas con más gasto, como el porcentaje de personas mayores de 80 años y la tasa de defunciones, o con menos gasto, como el número de turistas (que probablemente indica la existencia de una población más joven y una menor morbilidad). Como indican los autores, en este tipo de estudio deben evitarse las interpretaciones causales; aunque los modelos construidos puedan ser útiles para predecir el gasto farmacéutico (y para asignar presupuestos), ello no significa que las variables explicativas sean factores causales.

Otro ejemplo de datos ecológicos comparados sobre indicadores de calidad relevantes es la información que proporciona el sistema HEDIS (disponible en: [www.ncqa.org](http://www.ncqa.org)) que

publica y compara datos de 245 planes de salud en Estados Unidos, analizando la variabilidad de los resultados.

### **Estudios transversales**

Los estudios transversales se caracterizan porque los datos sobre el efecto y la exposición se miden de forma simultánea. Proporcionan información sobre el estado de salud de una población y sobre la utilización o las necesidades de servicios.

Son estudios útiles para conocer la prevalencia de factores de riesgo, para enfermedades o problemas que producen discapacidades y en fases presintomáticas de enfermedades más graves. La información de estos estudios se puede emplear para sugerir hipótesis causales, pero debe interpretarse con cuidado, ya que pueden existir dificultades para decidir qué es la causa y qué es el efecto.

Una medida de asociación empleada en estos estudios es la razón de prevalencias, que se calcula como el cociente entre la prevalencia de enfermedad en individuos expuestos y no expuestos. Al trabajar con datos de prevalencia hay que tener en cuenta que en la interpretación pueden influir tanto factores causales como factores determinantes de la supervivencia.

Las relaciones entre variables se miden con *odds ratio* (OR), que se interpretan como medidas de asociación, pero como no existe secuencia temporal, no se puede afirmar la relación causa-efecto.

Se puede encontrar muchos ejemplos de estudios transversales. El estudio IBERICA registra los datos de pacientes que han sufrido un infarto agudo de miocardio (IAM) y que han ingresado en distintos hospitales españoles, con el objetivo de describir las características de esos pacientes y el tipo de tratamientos a que son sometidos. En una reciente publicación<sup>6</sup> de datos de este registro (correspondientes al año 1997), con más de 4.400 individuos, se cita que el 18% de los pacientes con un IAM era diabético, el 46% hipertenso o que el 12% presentó síntomas atípicos. Otros datos interesantes del estudio son el bajo porcentaje de uso de bloqueadores beta al alta (44%), o el elevado porcentaje de uso de nitratos (53%), en este caso superior a lo recomendable.

Los estudios que miden la adecuación de días de estancia en el hospital suelen ser estudios de diseño transversal. Desde el trabajo de Gertmann y Restuccia<sup>7</sup>, en que se desarrolló la herramienta más usada en estos estudios, el AEP, muchos estudios han encontrado una asociación entre las estancias inadecuadas y ciertas variables como la edad o el sexo de los pacientes, la duración de la hospitalización o la formación del médico.

### **Estudios de seguimiento**

En este caso, los sujetos son seguidos a lo largo de un período, realizando mediciones de los factores de riesgo (lo más habitual será medir la exposición una vez y suponer que no cambia) y de la aparición de enfermedad.

Los individuos que componen los grupos o cohortes se seleccionan en función de la presencia de una determinada

característica o exposición. El supuesto más simple es el de dos grupos, uno de ellos expuesto a un riesgo y el otro no expuesto. En cada grupo se calculan las tasas de incidencia y se comparan calculando el riesgo atribuible y el riesgo relativo.

Podemos diferenciar dos tipos de estudios de seguimiento: *a)* aquellos en que se selecciona una cohorte de expuestos y se busca una cohorte de comparación (no expuestos), y *b)* aquellos en que se selecciona una cohorte única, en la que se miden diversas exposiciones, cuyos efectos van a ser estudiados.

Los objetivos de los estudios de seguimiento pueden ser: *a)* cuantificar la incidencia de la enfermedad; *b)* comparar el riesgo que tienen los expuestos frente a los no expuestos; *c)* descubrir si existe una relación dosis-respuesta, y *d)* determinar la fracción atribuible, es decir, la proporción de casos que pueden atribuirse al factor de riesgo.

Estos estudios se diferencian por su tamaño y por su complejidad. Pueden variar desde el control a una gran población durante décadas hasta pequeños grupos durante unos días o semanas.

Para reducir el riesgo de algunos factores de confusión se puede seleccionar grupos de comparación con las mismas características en esos factores (apareamiento o *matching*), o medir esos factores en ambos grupos y ajustar después los resultados en el análisis estadístico.

Los estudios de seguimiento pueden ser prospectivos o retrospectivos. En los estudios prospectivos se elige una muestra de individuos que no presentan la enfermedad, se miden los factores que podrían estar relacionados con su aparición, y se efectúa un seguimiento. Son estudios caros, y resultan poco eficientes para enfermedades poco frecuentes y para problemas que tienen un período de inducción largo.

En los estudios retrospectivos la reunión de la cohorte, las mediciones iniciales, el seguimiento y la enfermedad se produjeron en el pasado. Son más baratos y requieren menos tiempo. Todos los individuos que han desarrollado la enfermedad y todos los que no la han presentado proceden de la misma población, con lo que se reduce el riesgo de sesgos de selección.

Se puede combinar los dos tipos de estudio, recogiendo datos retrospectiva y prospectivamente en una misma cohorte. Esta estrategia sirve para exposiciones que tienen efectos a corto y largo plazo como, por ejemplo, las exposiciones ambientales a productos tóxicos.

El seguimiento prospectivo es más costoso y exige más implicación del investigador, pero los resultados de la investigación van a ser mejor valorados.

Junto a los ejemplos descritos en la primera parte al repasar las medidas de asociación<sup>8</sup>, encontramos otros trabajos que realizan estudios de seguimiento útiles en evaluación de calidad. Krumholz et al<sup>8</sup> realizaron un estudio de cohortes retrospectivo para describir el modelo de uso de bloqueadores beta (BB) tras IAM en pacientes de más de 65 años, buscando datos de infrautilización de BB y factores relacionados. Sobre un grupo de 115.015 pacientes que sobrevivieron a una hospitalización con IAM, distinguieron entre los que fueron y los que no fueron tratados con BB como medicación al alta y

analizaron distintas variables, como la mortalidad durante el primer año. Estos autores encontraron que sólo el 50% de los pacientes sin contraindicaciones llevaba un tratamiento al alta con BB y que, con los datos del seguimiento, el uso de BB redujo el riesgo de mortalidad al cabo del año en un 14%.

### Estudios de casos y controles

En los estudios de casos y controles se empieza por identificar a los pacientes en los que se ha producido una enfermedad o suceso y a otros en que no se ha desarrollado, y luego se comprueba y se compara su historia de exposición a factores sospechosos.

Estos estudios permiten evaluar un amplio número de factores que pueden estar relacionados con la enfermedad o el suceso, y estudiar también la interrelación entre estos factores. Son una solución para analizar las enfermedades con largos períodos de latencia y para enfermedades raras, ya que de otra forma se necesitaría hacer un seguimiento de un gran número de personas. En la evaluación de la calidad, vamos a enfrentarnos con frecuencia con "enfermedades" raras y con múltiples y desconocidos factores de riesgo.

La selección de los casos es el punto de inicio. Se requiere una definición de caso adecuada. En general, es mejor usar casos nuevos que casos prevalentes, ya que la prevalencia la determinan tanto el riesgo de enfermar como los factores que influyen en la duración. Además, si la enfermedad empezó hace mucho, la exposición empezó hace más, y puede ser difícil de averiguar (de recordar).

La selección de los controles suele ser más problemática que la de los casos. Los controles deben ser representativos de la misma población base de la que proceden los casos, es decir, gente que no tiene la enfermedad pero que si la tuviera podrían ser elegidos como casos. Otro requisito es que la exposición de los controles debe ser medible con la misma exactitud que la de los casos. Dos fuentes de controles habituales son la población general y los pacientes con otras enfermedades (útiles como controles siempre que sus enfermedades no tengan relación con la enfermedad en estudio ni con sus factores de riesgo).

Una estrategia para aumentar la rentabilidad del estudio es que para cada caso se seleccione uno o varios controles con los mismos valores en algunas variables cuyo efecto se quiere controlar (apareamiento o *matching*). También se puede realizar el apareamiento utilizando grupos en lugar de individuos.

La medida de asociación que se puede calcular es la OR, para analizar la relación entre cada variable de exposición y la enfermedad. Se calcula también una OR para cada grado de exposición cuando la variable explicativa tiene más de dos categorías. Como se discutía al describir las medidas de asociación, con este diseño no se puede calcular el riesgo atribuible, por lo que no se podrá determinar la relación entre factores, el impacto relativo de cada factor.

Dos ejemplos de este tipo de estudios: Bates et al<sup>9</sup> plantean dos estrategias de estudio (un estudio de casos y controles y otro de seguimiento) para analizar la relación entre diversas variables y la aparición de un efecto adverso por

medicamentos en pacientes hospitalizados. En el estudio de casos y controles revisan 190 casos de efectos adversos de medicamentos y eligen un control (pacientes que no han sufrido un efecto adverso) para cada caso, seleccionando el control en la misma sala de hospital y con la misma duración de estancia hasta el momento de aparecer el efecto adverso. Al comparar con esta estrategia distintas variables con el riesgo de desarrollar un efecto adverso, no encuentran relaciones estadísticamente significativas. En la tabla 1 se presentan los resultados de la comparación entre casos y controles según si tomaban o no más de 15 fármacos distintos en las 24 h anteriores al análisis de los datos.

En este estudio los autores describen alguno de los problemas citados más arriba en la selección de casos y controles. Por ejemplo, los sujetos del estudio tienen una estancia hospitalaria más larga y es más probable que ingresen en la UCI que el resto de los pacientes del hospital, lo que deberá tenerse en cuenta al extrapolar los resultados.

El otro ejemplo de casos y controles es el ya citado de García Martín et al<sup>1</sup>, que analizaba la relación entre haber sufrido un efecto adverso durante la hospitalización y el riesgo de muerte. Revisando 524 casos y 524 controles encontraron una asociación significativa entre factor de riesgo y resultado. El 57% de los casos y el 43% de los controles habían presentado algún efecto adverso, con una OR de 1,75 (intervalo de confianza del 95%, 1,37-2,24).

### Estudios experimentales

Los estudios experimentales son el paradigma de la medicina basada en la evidencia, y deberían ser la estrategia de elección para generar conocimiento científico en evaluación de la calidad.

En estos estudios el investigador determina qué pacientes se exponen a un factor y cuáles no. Con ello, se trata de evitar el efecto de los factores de confusión, especialmente cuando la asignación de pacientes a uno u otro grupo (experimental y control) se realiza de forma aleatoria y cuando el número de datos analizado es suficiente.

La gran fortaleza de estos estudios se basa en que la asignación aleatoria produce grupos que deberían ser similares en todo salvo en la intervención que se está probando, lo que evita explicaciones alternativas a los resultados.

La principal utilidad de los estudios experimentales es la evaluación de intervenciones terapéuticas mediante ensa-

Tabla 1. Presentación de resultados de la relación entre exposición y enfermedad

	Casos	Controles	Total
Toma > 15 fármacos	48	38	86
< 15 fármacos	142	152	294
Total	190	190	380
OR = 1,35 (intervalo de confianza del 95%, 0,83-2,19).			



yos con asignación aleatoria y grupo control: cada paciente es asignado a un solo grupo de forma aleatoria, recibe una sola intervención y los resultados obtenidos en cada uno de los grupos se comparan entre sí.

Inicialmente, el diseño del ensayo debe especificar qué criterios han de cumplir los pacientes para ser incluidos en el estudio. Después se pide el consentimiento del paciente para participar en el ensayo. Si la tasa de aceptación es baja debe considerarse el grado de representatividad respecto a la población de los voluntarios que han aceptado. Luego se asignan aleatoriamente de forma automatizada. Los criterios para retirar a un paciente del estudio deben estar escritos de antemano. Es importante medir todos los posibles efectos adversos.

Una estrategia que sirve para mejorar la confianza en los resultados del experimento es "cegar" a los participantes, es decir, impedir que conozcan cuál de las intervenciones del ensayo están padeciendo u observando. Puede ser "ciego" el paciente, que no sabe a qué grupo ha sido asignado, y puede ser "ciego" el que realiza la intervención o el que mide los resultados.

El tamaño del ensayo debe estar decidido de antemano teniendo en cuenta la distribución esperada de las medidas de resultados y la diferencia en resultados que se quiere detectar. Puede que antes del fin previsto ya se haya demostrado las conclusiones, lo que se sabrá realizando análisis a intervalos y estableciendo criterios para interrumpir el ensayo si un tratamiento demuestra ser mejor que otro.

En un ensayo publicado por Drakulovic et al<sup>10</sup> se analizó el efecto que la posición semisedente del paciente tenía en la prevención de la neumonía nosocomial en las unidades de cuidados intensivos. El ensayo se detuvo después del análisis intermedio, cuando se habían asignado aleatoriamente 86 pacientes con ventilación mecánica en dos grupos, 39 a posición semisedente y 47 a decúbito supino. La frecuencia de neumonía de sospecha clínica fue claramente menor en el primer grupo (8%) que en el segundo (34%). El ensayo también sirvió para analizar la influencia de otros factores. En este caso, se mostraron como factores de riesgo independiente para desarrollar una neumonía nosocomial la nutrición enteral (OR = 5,7) y mantener la ventilación mecánica durante 7 días o más (OR = 10,9).

A la hora de realizar el análisis hay que saber que la comparación se debe hacer colocando a cada paciente en el grupo al que fue asignado inicialmente. Es posible que en algún grupo se hayan producido pérdidas en el seguimiento o cambios de pacientes de un grupo a otro. Por ello, y para conservar las ventajas de la asignación al azar, se debe realizar el análisis por "intención de tratar", con todos los pacientes formando parte del denominador al que fueron asignados en el inicio.

También se realizan ensayos para evaluar la eficacia de cambios organizativos. Curley et al<sup>11</sup> realizaron un ensayo para comprobar el efecto de un nuevo sistema de trabajo en la atención a pacientes hospitalizados en un servicio de medicina interna. Se asignó aleatoriamente pacientes a varias unidades o grupos de profesionales, cuando alguna de esas unidades había cambiado su sistema de trabajo para realizar

rondas interdisciplinarias diarias, frente a otras que mantuvieron el sistema tradicional de reuniones semanales. El ensayo encontró que el nuevo sistema permitía reducir la duración de la estancia y el coste, y mejorar la adecuación de determinados tratamientos como las terapias respiratorias y los cuidados nutricionales.

Una estrategia de ensayo usada en la evaluación de medicamentos son los ensayos con entrecruzamiento, en los que cada paciente que participa en el ensayo recibe los dos tratamientos secuencialmente, con un período de "lavado" intermedio. Es un diseño útil en enfermedades crónicas y relativamente estables y cuando los resultados de la intervención desaparecen de forma rápida. Cada paciente sirve como su propio control.

Una variación más son los ensayos de un solo caso ( $n = 1$ ), en los que en un único paciente se prueban dos tratamientos, o un tratamiento activo y un placebo. Se administra un tratamiento en cada subciclo, y se repite la prueba durante varios ciclos. Este método es útil cuando la enfermedad es crónica y estable, y la respuesta al tratamiento es inmediata. Este tipo de estudio está indicado cuando el médico o el paciente dudan de la efectividad de un tratamiento, cuando existen dudas sobre la dosis óptima en este paciente, y como prueba diagnóstica (test estímulo-respuesta).

Los estudios experimentales se pueden realizar sin asignación aleatoria. La característica principal es que el investigador controla el factor de estudio pero no existe asignación al azar del mismo. Las dos variedades son según exista o no grupo control:

– Estudios con grupo control y sin asignación aleatoria. Son útiles cuando la asignación aleatoria no ofrece ventajas; por ejemplo, cuando existen sólo dos unidades de estudio, o cuando resulta difícil aleatorizar, como puede ser el caso de ensayos con intervenciones organizativas. El grupo que recibirá la intervención lo decide el investigador. En estos estudios se asume que ambos grupos son similares en lo que respecta a factores pronóstico. Sin embargo, pueden diferir por una serie de variables que no han sido controladas. Por esta razón, los resultados se han de interpretar con precaución. En un ejemplo de intervención sin asignación al azar entre dos salas de hospitalización, Mozes<sup>12</sup> encontró que la proporción de días de estancia inadecuados pasó del 48 al 15% y la duración media de la estancia de 6,3 a 4,6 días, después de una intervención en una de las salas basada en pedir a los médicos que cada día justificaran el criterio de estancia, y si no lo hacían, dar de alta al paciente. En otra sala adyacente sobre la que no se realizó ninguna intervención también se redujo, aunque de forma no significativa, la proporción de días inadecuados, del 56 al 36%.

– Ensayos sin grupo control, o estudios "antes-después". Se evalúa la respuesta de un grupo de sujetos antes y después de una determinada intervención, actuando cada individuo como su propio control. Con este diseño no se puede asegurar que los cambios aparecidos sean debidos a la propia intervención, a otros factores no controlados, o al simple hecho de que los sujetos se saben estudiados (ejemplo conocido como efecto Hawthorne). El efecto de regresión a la media

—cuando seleccionamos a pacientes por tener valores anormales de una variable, por ejemplo de presión arterial, las sucesivas mediciones en esos mismos pacientes, sin haber hecho nada, tenderán a estar más cercanas al valor medio— debe ser tenido en cuenta en el análisis.

Goldberg et al<sup>13</sup> realizaron un ensayo para probar la eficacia de las técnicas de *academic detailing* (AD) y de los equipos de mejora de calidad (EMC) para aumentar la aceptación de las guías de práctica en la atención primaria de la hipertensión arterial y la depresión. Quince equipos de profesionales sanitarios de cuatro clínicas fueron distribuidos en bloques en uno de los tres grupos del ensayo: AD solamente, AD con EMC y grupo control. En el grupo de AD, dos médicos de cada clínica fueron entrenados para realizar esta técnica, con diversos materiales de apoyo. Los médicos asignados a los EMC formaron un grupo en cada clínica y para cada enfermedad, con la ayuda de un asesor para utilizar las técnicas del ciclo de Shewhart. Ninguno de los dos grupos de intervención logró mejorar los patrones de prescripción ni los índices de resultados clínicos. Los autores sugieren como limitaciones el distinto nivel de esfuerzo y de aceptación de las técnicas en cada uno de los cuatro centros de atención primaria participantes. No se descarta la existencia de un efecto “Hawthorne” en el grupo control, así como la existencia de cambios debidos a efectos “seculares” producidos por la aparición de nuevos fármacos o de noticias sobre efectos adversos que hacen variar los patrones de consumo.

Otra variante de los estudios experimentales son los ensayos con poblaciones o ensayos de tipo comunitario, que pueden ser una estrategia muy útil en la investigación en mejora de calidad. En este caso, los asignados aleatoriamente no son individuos sino poblaciones. Si el número de poblaciones es pequeño, la aleatorización puede no tener mucho sentido, puede ser razonable asignar de manera deliberada para asegurar la posibilidad de comparación. Son estudios caros, pero proporcionan una gran evidencia de las relaciones entre causas y efectos. En caso de existir aleatorización, son las comunidades enteras y no los sujetos las que se distribuyen en el grupo experimental y el control. El riesgo de sesgos en la formación de los grupos experimental y control sigue existiendo, ya que resulta difícil que la distribución de todos los factores que puedan influir negativamente en el estudio sea similar entre los grupos comunitarios.

Soumerai et al<sup>14</sup>, en un ensayo con grupo control y asignación aleatoria, con el hospital como unidad de análisis, repartieron 37 hospitales del Estado de Minnesota en dos grupos, de 20 el grupo experimental y 17 el control. Los hospitales se estratificaron por tamaño y los de la misma ciudad se asignaron al mismo grupo. Después de distribuir guías de práctica sobre el manejo del infarto de miocardio (IAM), en el grupo experimental se identificó en cada hospital a los médicos expertos en IAM, estos expertos diseñaron material informativo y participaron en actividades educativas dirigidas a los colegas de sus hospitales, sobre todo a través de sesiones con grupos pequeños. La intervención incrementó de forma significativa el uso de aspirina (17% más de casos bien tratados) y de bloqueadores beta (63%). Descendió el uso

inadecuado de lidocaína (del 50%) tanto en el grupo experimental como en el control y no se incrementó el uso de la terapia trombolítica por encima del 70% inicial. Entre otras conclusiones, los autores indican que la existencia del grupo control evitó atribuir a la intervención el descenso en el uso de lidocaína.

Con otro estudio de diseño similar, Marrie et al<sup>15</sup> evalúan los resultados de una vía clínica en el tratamiento de la neumonía comunitaria. Para realizar el ensayo se separan dos grupos de hospitales de forma aleatoria, con 10 centros en un grupo y 9 en el otro, y en el grupo experimental se introduce la metodología de la vía clínica especificando los criterios de ingreso y la pauta de tratamiento antibiótico (con levofloxacina). Los resultados del ensayo, medidos en proporción de complicaciones, reingresos y mortalidad no son distintos entre los dos grupos. En el grupo de la vía clínica se redujo la duración de la estancia en el hospital y mejoró la adecuación de los ingresos. La asignación aleatoria de hospitales no impidió que los pacientes en el grupo experimental tuvieran al principio del ensayo puntuaciones de gravedad mayores que los del grupo control.

## Estudio de los brotes epidémicos

Por definición, una epidemia es la aparición de un número de casos superior al esperado. Aunque las enfermedades transmisibles se están reduciendo en las sociedades industrializadas, muchos problemas se siguen manifestando en forma de epidemias, de brotes.

Para la investigación de las causas de una epidemia existe una rutina, un método que sirve tanto para las enfermedades transmisibles como las no transmisibles. Las fases de este método son:

1. Verificar la existencia del brote. Existe una agrupación de casos, que pueden definirse sólo por unos pocos signos o características comunes, y que dan la impresión de compartir la causa.
2. Conocer la frecuencia previa de la enfermedad para decidir si se está ante un brote. Las enfermedades contagiosas siguen una curva gradual de aparición de casos, mientras que en las que se producen por la acción única de un agente todos los casos pueden aparecer de forma brusca.
3. Para identificar las características de la epidemia, hay que realizar una historia, una recogida retrospectiva de datos, de todos los casos.
4. Realizar una búsqueda adicional de casos.
5. Definir una población a riesgo, con un mapa de la extensión geográfica de la epidemia.
6. Identificar una experiencia que haya sido vivida por los afectados y no por los no afectados, y con ello elaborar una hipótesis sobre la fuente y la evolución de la epidemia.
7. Calcular tasas de ataque. Es una medida de incidencia en la que el numerador son los casos de una enfermedad y el denominador la población a riesgo durante un período limitado de exposición (número de casos que aparecen entre quienes tomaron un alimento sospechoso). Comparar las ta-

sas de ataque de distintos factores puede dar las claves para la investigación.

8. Analizar los datos, buscar asociaciones estadísticas. La OR suele ser la medida de asociación entre cada variable y el riesgo de haber padecido el problema epidémico.

Un ejemplo muy reciente de análisis de un brote lo publican Jones et al<sup>16</sup>, describiendo un brote de complicaciones neurológicas asociado a errores en el manejo de catéteres intratecales. La perfusión intratecal de medicación produjo una agrupación de casos de complicaciones neurológicas graves, que se analizaron mediante un estudio de cohortes en una planta de neurocirugía. Se siguió a 61 pacientes con bombas para analgesia, 19 de los cuales se trataban con morfina. Ninguno de los 42 que no llevaban morfina desarrolló una complicación, mientras que 8 de los 13 que eran tratados con morfina padecieron complicaciones en un período de 4 semanas. En preparaciones de una farmacia se encontró que aunque estaban etiquetadas como morfina pura, contenían también metadona, etanol o metanol. La existencia de errores asociados a la farmacia extrahospitalaria fue la causa encontrada como responsable de las complicaciones.

### Agradecimientos

A Susana Lorenzo, por su confianza y su paciencia. A Soledad Márquez, a Voro Peiró y al resto de profesores del IVESP por su buen trabajo.

### Bibliografía

- García M, Lardelli P, Bueno A, Luna J, Espigares M, Gálvez R. Proportion of hospital deaths associated with adverse events. *J Clin Epidemiol* 1997;50:1319-26.
- Brennan T, Leape L, Laird N. Incidence of adverse events and negligence in hospitalised patients: results of the Harvard Medical Practice Study 1. *N Engl J Med* 1991;324:370-6.
- Jabary NS, Herrero AM, Gonzalez JA. The use of antihypertensive therapy in Spain (1986-1994). *Am J Hypertens* 2000;13: 607-10.
- Caicoya M, Natal C, Alonso P, Moral L. Análisis de la variabilidad en el uso de diálisis en el territorio Insalud. *Rev Calidad Asistencial* 2001;16:101-6.
- García Sempere A, Peiró S. Gasto farmacéutico en atención primaria: variables asociadas y asignación de presupuestos de farmacia por zona de salud. *Gac Sanit* 2001;15:32-40.
- Fiol M, Cabadés A, Sala C, Marrugat J, Elosua R, Vega G, et al. Variabilidad en el manejo hospitalario del infarto agudo de miocardio en España. Estudio IBERICA (Investigación, Búsqueda Específica y Registro de Isquemia Coronaria Aguda). *Rev Esp Cardiol* 2001;54:443-52.
- Gertmann P, Restuccia J. The Appropriateness Evaluation Protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981;19:855-71.
- Krumholz HM, Radford MJ, Wang Y, Chen J, Heiat A, Marciniak TA. National use and effectiveness of beta-blockers for the treatment of elderly patients after acute myocardial infarction: National Cooperative Cardiovascular Project. *JAMA* 1998;280:623-9.
- Bates D, Miller E, Cullen D, Burdich L, Williams L, Laird N, et al. Patient risk factors for adverse drug events in hospitalized patients. *Arch Intern Med* 1999;159:2553-60.
- Drakulovic MB, Torres A, Bauer TT, Nicolas JM, Nogue S, Ferrer M. Supine body position as a risk factor for nosocomial pneumonia in mechanically ventilated patients: a randomised trial. *Lancet* 1999;354:1851-8.
- Curley C, McEachern JE, Speroff T. A firm trial of interdisciplinary rounds on the inpatient medical wards. An intervention designed using continuous quality improvement. *Med Care* 1998;36(Suppl 8):AS4-12.
- Mozes B. Reduction of redundant hospital stay through controlled intervention. *Lancet* 1987;25:968-9.
- Goldberg H, Wagner E, Fihn S, Martin D, Horowitz C, Christensen D, et al. A randomized controlled trial of CQI teams and academic detailing: Can they alter compliance with guidelines? *Jt Comm J Qual Improv* 1998;24:130-42.
- Soumerai SB, McLaughlin TJ, Gurwitz JH, Guadagnoli E, Hauptman PJ, Borbas C, et al. Effect of Local Medical Opinion Leaders on Quality of Care for Acute Myocardial Infarction. A Randomized Controlled Trial. *JAMA* 1998;279:1358-63.
- Marrie T, Lau C, Wheeler S, Wong C, Vandervoort M, Feagan B. A controlled trial of a critical pathway for treatment of community acquired pneumonia. *JAMA* 2000;283:749-55.
- Jones TF, Feler CA, Simmons BP, Melton K, Craig AS, Moore WL, et al. Neurologic complications including paralysis after a medication error involving implanted intrathecal catheters. *Am J Med* 2002;112:31-6.