

METAANÁLISIS EN EPIDEMIOLOGIA (Primera parte): CARACTERÍSTICAS GENERALES

Miguel Delgado Rodríguez*/** / María Sillero Arenas*** / Ramón Gálvez Vargas*

* Cátedra de Medicina Preventiva. Facultad de Medicina. Granada ** Hospital Virgen de las Nieves. Granada. *** Delegación Provincial de Salud y Consumo. Jaén.

Resumen

En el presente trabajo se resumen las principales características del metaanálisis en salud pública, así como sus objetivos y utilidades. También se establecen las directrices generales en la realización de este tipo de técnicas. A lo largo de la revisión se detallan las distintas etapas en la ejecución de un metaanálisis, comentándose con ejemplos las ventajas e inconvenientes de las distintas opciones que se presentan en el diseño.

Palabras clave: Metaanálisis. Métodos epidemiológicos. Revisiones.

META-ANALYSIS IN EPIDEMIOLOGY: GENERAL CHARACTERISTICS (First part)

Summary

In the present work the general characteristics, aims and applications of meta-analysis in public health are described. The general rules and the relevant steps for the development of this type of studies are considered. These points are illustrated with a number of real examples and the advantages and limitations of the different methodological options are commented.

Key words: Meta-analysis. Epidemiological Methods. Reviews.

Introducción

El término metaanálisis (MA) fue acuñado por vez primera por Glass en 1976¹, para describir la interpretación estadística de datos procedentes de varias experiencias similares. Nosotros lo entendemos como un análisis sistemático del principio de consistencia en las diferentes investigaciones que han hecho una valoración de una determinada relación causal. En el MA la unidad de estudio es cada investigación realizada con anterioridad sobre el tópico que nos interesa. En el terreno de la medicina, los primeros autores que hemos encontrado que utilizaran técnicas de metaanálisis fueron MacMahon y Hutchison², que combinaron resultados de exposición intrauterina a radiaciones y subsiguiente riesgo de leucemia. En el momento presente asistimos a un aumento creciente de realización y publicación de MA en la literatura médica. A pesar de

ello su uso aún no es recomendado de manera rutinaria para la valoración global de un hecho o una asociación.³

Los científicos piensan que la investigación reciente es mejor que la antigua por una razón bastante sencilla: la investigación moderna aprende de las lecciones enseñadas por la del pasado. Pero los últimos hallazgos sólo tienen sentido cuando se comparan con los producidos con anterioridad. En este sentido, para que la ciencia se acumule se necesita un escalón intermediario entre el pasado y el futuro. Las «Revisiones» son por ello necesarias. No sólo cumplen una función de iniciación en la investigación de un tema, también tienen una indiscutible utilidad en la enseñanza. Sin embargo, las «revisiones clásicas» tienen una serie de defectos ya que intentan integrar de manera no sistemática estudios realizados de muy diversas formas. Los problemas principales que presentan este tipo de revisiones son⁴:

— Su carácter subjetivo, ya que el proceso de revisión prácticamente carece de cualquier regla formal y es poco probable que dos estudiosos independientes del mismo tema coincidan a la hora de entresacar de un mismo universo dos revisiones «clásicas».

— No tienen carácter científico. Con gran frecuencia, cuando los autores encuentran a lo largo de su búsqueda bibliográfica estudios con resultados contradictorios, recurren a un «recuento de votos» (tantos a favor, tantos en contra, tantos nulos). De esta manera, los aspectos metodológicos de cada investigación son ignorados.

— Las revisiones tradicionales ofrecen un camino poco eficaz en la extracción de información útil. Esto se produce sobre todo cuando el número de estudios originales es grande.

La forma de obviar estos inconvenientes es la utilización de una metodología que permita combinar cuantitativamente

resultados procedentes de varios estudios. Esta es la razón de ser del metaanálisis.

Definición

El metaanálisis consiste en un conjunto de técnicas que se utilizan para cuantificar la información contenida en estudios similares. Glass¹ lo definió como «el análisis estadístico de una colección amplia de resultados procedentes de varios experimentos con el propósito de integrar los hallazgos». Der Simonian y Laird⁵ lo han definido de manera similar: «el análisis estadístico de una colección de resultados analíticos con el propósito de integrar los hallazgos». Preferimos la definición de Der Simonian y Laird ya que no limitan a priori su ámbito de aplicación; para Glass el metaanálisis sólo es utilizable en estudios experimentales y no en otro tipo de diseños. Con Glass coinciden otros autores⁶⁻⁷, aunque irónicamente algunos de ellos, tras su primera opinión⁷ hayan realizado metaanálisis sobre estudios observacionales⁸.

El metaanálisis se lleva a cabo para aumentar el poder de una prueba estadística y para obtener información que no puede ser derivada de manera fiable de los experimentos individuales. Esto es adecuado para fenómenos de baja frecuencia o cuando se valoran efectos pequeños, pero nunca es un sustituto para un estudio bien diseñado, realizado y analizado con un tamaño muestral adecuado.

Objetivos

Los objetivos principales de un metaanálisis son⁴⁻⁹⁻¹⁰: 1) Investigar la relación entre dos variables que pueden mantener relaciones de causalidad (un factor de riesgo y una enfermedad, un tratamiento y la curación de una enfermedad, etc.) y su magnitud; y 2) aumentar la precisión de las estimaciones del efecto que se estudia. Esta metodología permite además la consecución de otros objetivos concretos: 1) valoración de la consistencia entre diferentes estudios que analizan las mismas variables, y de la consistencia entre distintos niveles de la variable exposición (un factor de riesgo, un tratamiento,

etc.); 2) obtención un mejor estimador global de la relación entre las variables que se analizan; 3) identificación de subgrupos que son particularmente susceptibles a la variable exposición que se estudia; 4) valoración de la calidad de la metodología empleada en cada una de las investigaciones individuales, y elaboración, tras el análisis de la consistencia entre los distintos resultados, de propuestas metodológicas a seguir en futuras investigaciones.

Utilidades

Las aplicaciones del metaanálisis se derivan de los objetivos que acabamos de exponer.

1. Descripción de un determinado fenómeno:

a) Valoración de errores de metodología en la investigación de una asociación. Ejemplos de ellos son la valoración de errores y de los patrones de publicación en la asociación de anticonceptivos orales con los tumores de mama y cuello uterino¹¹⁻¹² o la valoración metodológica que Tirosh y Rabino¹³ realizaron sobre los estudios que analizaron el valor de la fisioterapia en los niños con parálisis cerebral.

b) Descripción simple de lo que ocurre en un cierto colectivo, como por ejemplo, la descripción de los problemas encontrados en las personas que atienden a enfermos dementes¹⁴.

c) Aumento de la precisión de un estimador, como el cálculo de la prevalencia de un determinado proceso en la colectividad. Ejemplo de ello sería la valoración de la prevalencia de infección por el VIH en ciertos grupos de riesgo de la población española¹⁵⁻¹⁶.

2. Análisis de presuntos factores de riesgo frente a una determinada enfermedad. Ejemplos de ellos son las valoraciones de los anticonceptivos orales y el riesgo de cáncer de mama o de cuello uterino¹⁷⁻¹⁸.

3. Valoración de pruebas diagnósticas, como el realizado por Gianrossi et al¹⁹ al valorar la eficacia de la depresión del segmento ST en el electrocardiograma realizado durante el ejercicio frente a la coronariografía en el diagnóstico de la enfermedad coronaria.

4. Valoración de tratamientos. Este ha sido una de sus aplicaciones más clásicas, ya que la mayoría de ellas han necesitado de ensayos clínicos para su puesta en práctica. Ejemplos de ellos tenemos el ya clásico metaanálisis de Yusuf et al²⁰ sobre betabloqueantes en la terapia de infarto agudo de miocardio o la valoración de las soluciones coloidales frente a las cristaloides en la resucitación²¹.

5. Identificación de subgrupos especiales (para un factor de riesgo, una prueba diagnóstica o un tratamiento). Como ejemplo, en el análisis de los anticonceptivos orales como factor de riesgo del cáncer de mama, el riesgo parece recaer especialmente en las mujeres que consumieron el preparado a edades jóvenes (antes del primer embarazo o antes de los 25 años de edad) y que luego desarrollaron un tumor premenopáusic¹⁷.

6. Valoración de la heterogeneidad (consistencia) de diferentes investigaciones. Se pretende con ello justificar el por qué distintas investigaciones han producido diferentes resultados. Un ejemplo de ello lo constituye el metaanálisis de Ewertz et al²², en el que se analizan las inconsistencias encontradas en 8 estudios realizados en los países nórdicos con respecto a considerar a la edad al primer parto y a la paridad como predictores independientes de cáncer de mama. Curiosamente los autores concluyen que una razón puede ser la falta de potencia estadística sin ofrecer ninguna estimación de la misma, cuando ello es bastante factible²³⁻²⁴.

Protocolo

Un metaanálisis debe realizarse con rigor científico (que no siempre se cumple). En la tabla 1 se resumen las características generales que todo metaanálisis debería reunir. Se debe constatar en él la existencia de los siguientes aspectos:

1. Una hipótesis de trabajo.
2. Selección de la población de estudio: Criterios de búsqueda de la información, especificando el tipo de fuentes y el método de búsqueda (faltan, por ejemplo, en el análisis realizado por Prentice y Thomas²⁵ de los anticonceptivos ora-

Tabla 1. Protocolo de un metaanálisis

1. Hipótesis de trabajo
- SELECCION DE LA POBLACION DE ESTUDIO
2. Búsqueda de la información
 - Tipo de fuentes
 - Métodos de búsqueda
 3. Criterios de inclusión
 - Idioma
 - Tipo de efecto
 - Tipo de diseño
 - Tipo de datos (ej. morbilidad o mortalidad)
 - ¿Toda la información o sólo la metodológicamente aceptable?
- METAANALISIS CUALITATIVO
4. Cuestionario de evaluación de los estudios
 - Valoración de la calidad
 - Extracción de los datos cuantitativos pertinentes
- METAANALISIS CUANTITATIVO
5. Valoración del sesgo de publicación
 6. Métodos estadísticos para la combinación de resultados
 7. Análisis de la heterogeneidad

les con numerosos efectos; Criterios de inclusión de una investigación, con declaración expresa de la exposición y efecto que se investiga, entre otros puntos.

3. Metaanálisis cualitativo: Método seguido para la extracción de la información relevante y pertinente de cada investigación (falta, por ejemplo, en un metaanálisis de los anticonceptivos orales y el riesgo de malformaciones congénitas en la descendencia²⁵).

4. Metaanálisis cuantitativo: Valoración del sesgo de publicación (no existe, por ejemplo, en las referencias 8 y 25); Métodos estadísticos empleados para la combinación de resultados (faltan, por ejemplo, en el análisis de Fasoli et al²⁶, sobre anticoncepción postcoital); Metodología empleada para tratar el problema de la heterogeneidad (ausente en numerosos estudios; sirvan de ejemplo los dos últimos comentados²⁵⁻²⁶).

La no constatación de los anteriores hechos dificulta el que la evidencia pueda ser valorada correctamente por otros³. A continuación se describe más ampliamente los distintos puntos enumerados.

Origen de las hipótesis

1. Estudios exploratorios

El primer punto que debería dejar claro un investigador en el metaanálisis es

si tiene una hipótesis de trabajo, o por el contrario, intenta explorar el conocimiento existente sobre una asociación. En el primer caso, el análisis de la información disponible intenta comprobar una idea previamente elaborada. En el segundo, sucede lo opuesto. Un ejemplo de ello es la revisión crítica de la literatura realizada por Herrera y Lynch²⁷ para valorar si la calvicie es un factor de riesgo para la cardiopatía isquémica. Si se observa una asociación tras la combinación de resultados, se elabora una hipótesis que los justifique. En este caso, el análisis no puede concluir nada, ya que no es científico utilizar los mismos datos que sugieren una hipótesis para corroborarla. En este caso, y dado que se va a utilizar toda la información disponible, una estrategia sensata sería utilizar la mitad de los estudios para generar una hipótesis y el resto para confirmarla, una postura similar resulta adecuada cuando la hipótesis de trabajo ha sido sugerida tras la revisión no cuantitativa de toda la información⁹.

El problema se agudiza cuando se busca la asociación entre un efecto concreto y variables múltiples. En esta situación se examinan múltiples asociaciones a lo largo de varios estudios. La probabilidad de encontrar relaciones significativas ($p < 0,05$) entre una exposición y el efecto aumenta en gran medida, y con ella la

probabilidad de encontrar falsos positivos (asociaciones significativas que en realidad se deben al azar). Aparte de poder dividir los estudios a combinar en dos muestras, como en el caso anterior, en esta situación, el investigador podría ser más estricto con el nivel de significación para así disminuir la proporción de falsos positivos. Esto se puede alcanzar mediante el uso de análisis multivariantes. También se puede renunciar a la estadística inferencial. Se usan entonces medidas descriptivas (medias, coeficientes de correlación, etc.) que sugieran una posible hipótesis. A partir de ese momento se aplican las pruebas estadísticas.

2. Estudios con hipótesis de trabajo

Las hipótesis de trabajo suelen tener tres orígenes principales: a) Ya ha sido investigada en todos o en algunos de los estudios que serán combinados; b) Se deriva de otro tipo de investigaciones. Por ejemplo, resultados a nivel de animales de experimentación; c) Se deriva de una revisión de los datos de cada uno de los estudios que componen el estudio.

La situación primera es la más frecuente y no viola uno de los principios del razonamiento científico: las hipótesis deben ser formuladas antes de examinar los datos que serán utilizados para comprobarlas. La segunda es similar a la primera, aunque a veces sea difícil de establecer la independencia de origen de la hipótesis con respecto a los estudios incluidos. La tercera situación es inadecuada. Los datos no comprueban hipótesis, sino que las sugieren. Habrá que realizar otras investigaciones que confirmen esa hipótesis. En esta situación, si no se tiene una hipótesis, se puede utilizar una mitad aleatoria de los estudios para generarla, y la mitad restante para comprobarla.

Las hipótesis de trabajo de un metaanálisis suelen incluir, dependiendo de los objetivos del trabajo, enunciados como los siguientes: ¿Existe una asociación y cuál es su magnitud entre una exposición y un efecto?; ¿Qué tipo de la exposición es la que se asocia con el efecto?; ¿Cuáles son los subgrupos o poblaciones en los que particularmente se asocia la exposición con el efecto?

Búsqueda de la información

1. *Tipo de fuentes:* Existen varias opciones:

a) Intentar buscar toda la información disponible (informes publicados y no publicados de un estudio, tesis de licenciatura y tesis doctorales, informes de becas y contratos de investigación, etc.): Su principal desventaja está en el tiempo (—y el dinero—) necesario para localizar toda esta información y en la calidad, que puede ser muy variable, no siendo la misma la de los informes no publicados que la de los estudios publicados⁴⁻²⁸.

b) Utilizar sólo estudios publicados: Ahorra tiempo y dinero con respecto a la anterior, ya que la información es más accesible. El problema que presenta es el sesgo de publicación^{4,28-30}. Este error se produce por la tendencia a aparecer en la literatura estudios con resultados significativos. En la literatura se han realizado varios intentos de cuantificar este sesgo. Por ejemplo, Koren et al³¹ estudiaron la asociación de consumo de cocaína durante el embarazo y riesgos en el producto de la concepción, a lo largo de los diferentes resúmenes enviados a Society of Pediatric Research durante la década de los ochenta. Estos autores encontraron 58 resúmenes en los que se detallaba un efecto sobre el feto, de los que 9 no encontraron asociación, mientras que los restantes sí la mostraron. De los 9 que no encontraron nada sólo 1 (11%) fue aceptado. Por el contrario, 28 de los 49 (57%) que sí hallaron un efecto adverso en el feto por el consumo de cocaína fueron aceptados. Los estudios negativos rechazados tuvieron un mayor tamaño muestral, verificaron mucho más el consumo de cocaína y tenían con mayor frecuencia un grupo control, que los estudios positivos rechazados. Por lo tanto no se puede aducir que los resúmenes con resultados negativos sean de peor calidad que los positivos. El sesgo de publicación es evidente. Este tipo de fuente es la que con más frecuencia se utiliza en el metaanálisis. Dado el problema del sesgo de publicación mencionado en el párrafo anterior, los autores han de tenerlo siempre en cuenta.

c) Utilizar las bases originales de datos de los estudios que se combinan: Es

sin duda la mejor fuente, porque posibilita la realización del mejor análisis al tener información a nivel de cada sujeto participante en cada estudio. Su inconveniente radica en que es la menos factible (las bases de datos han de existir y sus autores han de ser proclives a permitir su uso) y la que requiere más tiempo (puesta en contacto con los autores, etc.). Hasta el momento presente sólo hemos localizado un estudio en la literatura que haya utilizado este sistema y que además no fue exhaustivo: 12 grupos de investigación se unieron para examinar la hipótesis de dieta y cáncer de mama a través de sus 12 estudios de casos y controles³².

2. *Métodos de búsqueda:* Existen diferentes alternativas y se sugiere que se empleen al menos dos de ellas.

a) Sistemas informatizados: El prototipo de ellos es el MEDLINE, que recoge las mismas revistas que el Index Medicus más un repertorio de revistas de enfermería. Tiene las ventajas de su rapidez y el que nos ofrece un resumen de los artículos. Su calidad va en función de las palabras clave que los autores establecieron para su trabajo. No suele haber problemas cuando en el informe de un estudio se analiza una sola exposición (con toda seguridad se incluirá entre las palabras clave), pero sí cuando, y es muy frecuente, los autores analizan más de una exposición. En este caso, si los resultados encontrados con la exposición diana no son muy concluyentes, no será incluida dentro de las palabras clave. El posible error introducido es de difícil cuantificación. A modo de ejemplo sirva el que hallamos 6 estudios epidemiológicos generales (de los 51 hallados en total) que ni en su título, ni en sus palabras clave (de acuerdo con MEDLINE) se incluían el término «anticonceptivos orales»¹⁸, dentro del metaanálisis de la relación anticonceptivos orales-cáncer de cuello uterino.

b) Búsqueda manual en repertorios: A nivel mundial los dos más conocidos son el Index Medicus y Excerpta Médica. El primero incluye unas 2.400 revistas y la segunda unas 3.400, ofreciendo además un resumen de los mismos. Existen también registros específicos como el Registro de Ensayos de Cáncer o el de Oxford

de Ensayos Perinatales, con búsqueda manual o informatizada. Este tipo de búsqueda tiene los mismos inconvenientes que la anterior, con la desventaja adicional de un consumo mayor de tiempo³³⁻³⁴.

c) Consulta de la bibliografía de cada uno de los artículos localizados por el procedimiento anterior: Siempre ha de hacerse, como complementaria a cualquiera de los métodos anteriores. No sólo han de revisarse las referencias al final de los estudios originales, sino las de cualquier fuente de información: revisiones, libros, etc. Con ello se tienen en cuenta las revisiones hechas por los demás autores que han investigado en el tema. Esto facilita la localización de artículos e informes publicados en revistas no indexadas en los repertorios de consulta. Igualmente, permite también el que se puedan localizar estudios no indexados por las palabras clave de interés, aunque contengan información sobre la asociación que se investiga. Una cuantificación de este tipo de búsqueda la ofrecen Ozminkowski et al³⁵, en el que por este mecanismo se localizaron 11 artículos adicionales a los 57 inicialmente localizados mediante una búsqueda informatizada.

d) La consulta con investigadores (uso de paneles de expertos) que hayan estudiado el problema puede ayudar a la identificación de informes y estudios no publicados. Presenta un inconveniente y es que la información proporcionada por ellos puede ser subjetiva, por lo que se recomienda combinarla con otro tipo de búsqueda.

e) Búsqueda en racimo tomando como punto de partida artículos clave: En esta opción, se parte de las referencias bibliográficas de artículos publicados en revistas de gran difusión y recientemente, y se va retrocediendo. Por ejemplo, fue usado por Shinton y Beevers³⁶ en su metaanálisis de cigarrillos y accidentes cerebrovasculares. Se suele usar este procedimiento en combinación con otros tipos de búsqueda ya mencionados.

f) Revisión de los índices de las revistas: Los métodos de búsqueda anteriores adolecen de un cierto retraso. La experiencia anterior de búsqueda suele reflejar que el número de revistas en las que se publican informes relacionados con el

tópico de interés es limitado. Si el tema es de interés, como suele ser la norma, esto hace aconsejable que los investigadores consulten el contenido de las revistas en las que es presumible que se publique un nuevo estudio.

g) Otras fuentes: No es desdeñable el consultar los libros de actas de congresos y reuniones científicas, los repertorios de tesis doctorales, etc. Con ello se intenta aumentar la calidad de la búsqueda y, además, posibilita el que se pueda cuantificar el sesgo de publicación, cuando se tiene sólo en cuenta la información proporcionada por los estudios publicados.

En un metaanálisis es necesario indicar las fuentes de donde se obtiene la información, de donde se va a obtener la población de estudio de este tipo de revisiones. A pesar de que esto es un criterio de calidad aceptado por todos los autores de metaanálisis, aún es posible encontrar trabajos en los que no se cita el método de búsqueda de artículos. Ejemplo de ello es el de Bracken²⁵, en el que se metaanaliza la relación entre consumo de anticonceptivos orales y malformaciones en la descendencia.

Criterios de inclusión

Los estudios reunidos a través de una búsqueda sistemática constituyen la serie inicial de estudios. No todos ellos serán incluidos en el análisis. Es necesario determinar mediante una serie de criterios de inclusión los que permanecerán y los que serán desechados. Por ejemplo, la búsqueda de la asociación entre contraceptivos orales y cáncer de mama arrojó más de 400 citas, procedentes de, o referentes a, más de 70 estudios distintos. Los criterios de inclusión establecían que el efecto sería el cáncer de mama (en el momento de su diagnóstico), la exposición los contraceptivos orales, que estuvieran basados en datos de morbilidad y no de mortalidad, que fueran diseños observacionales sin datos agregados (no estudios ecológicos), y publicados en inglés, francés, español, italiano, alemán o portugués. Esto redujo la serie a 47 investigaciones (una en francés, otra en español y las restantes en inglés)¹⁷. Las restantes analizaban el efecto pronóstico

de los contraceptivos sobre el cáncer de mama o su relación con los receptores estrogénicos, o su influencia en la mortalidad por cáncer, etc. La selección de estudios a incluir se puede estratificar por algunas características de los estudios. En algunas situaciones, especialmente cuando el número de estudios es muy amplio, puede ser suficiente tener una muestra de la información existente, con arreglo a algunas características esenciales de la investigación. Esto podría ser útil cuando las características del diseño influyen en el resultado final.

La selección además debe hacerse en función de la calidad de la investigación realizada. No todos los autores realizan un análisis de la calidad, para decidir con posterioridad si el estudio ha de ser incluido o no (aunque existen algunas dudas al respecto, creemos con la mayoría que los estudios con sesgos no deberían ser incluidos). Es necesaria pues la confección, dentro del protocolo del estudio, de un cuestionario de evaluación de la investigación. Este cuestionario debe ser aplicado por investigadores «ciegos», que no conozcan ciertas características del estudio a evaluar, y así no se influyen por ciertos aspectos del mismo. Los datos que no deben conocer son: la institución ejecutora y el nombre de los autores (su prestigio puede ejercer un cierto grado de presión), y el sentido de la asociación (si hay relación positiva, negativa o nula entre la exposición y el efecto), ya que su conocimiento podría influenciar también la selección. Esto se consigue si se suprime por otros investigadores esos datos identificadores. Normalmente se fotocopian de los artículos las secciones de «Material y Métodos» y la de «Resultados», tachando en esta última los epígrafes indicativos de los sujetos expuestos y no expuestos. Los evaluadores no deben estar involucrados en el proceso de recogida de la información. Se considera que es suficiente que sean dos los evaluadores, pero que realicen su tarea con independencia. Una vez concluida, examinarán sus conclusiones de manera conjunta. Si existen discrepancias deberán intentar solucionarlas entre ellos y si se mantienen la decisión final será tomada por un tercer investigador o consultor.

El cuestionario puede tener una triple función. En primer lugar, una cualitativa, al discriminar si el resultado final de un estudio está sesgado o no. Esta función puede ser esencial en el análisis. En un metaanálisis de la relación anticonceptivos orales-cáncer de cuello uterino, de los 51 estudios recogidos, sólo 18 fueron considerados metodológicamente aceptables¹⁸. En segundo lugar, tiene una función cuantitativa al poder resumir la calidad general de un estudio en un número (generalmente como un tanto por ciento, siendo el máximo 100). Esto permite relacionar la calidad del estudio con los resultados encontrados en el mismo. Así lo hicieron, por ejemplo, Berlin y Colditz³⁷ al estudiar el efecto de la actividad física en la prevención de la enfermedad coronaria. Encontraron que los estudios metodológicamente más correctos tienden a mostrar un mayor beneficio de la actividad física. En tercer lugar, el cuestionario permite detectar los errores más frecuentes en el tema a estudiar. Por ejemplo, en el metaanálisis del consumo de anticonceptivos orales y los cánceres de mama y cuello uterino, nuestro grupo realizó un análisis de los sesgos más frecuentes¹¹⁻¹², que desembocaron, por ejemplo, en el cáncer de mama, con unas recomendaciones generales para realizar en ese terreno una investigación con una menor posibilidad de estar sesgada³⁸.

Los cuestionarios de evaluación tienen un inconveniente, su validación. No existen cuestionarios ideales para ser aplicados en toda situación. El que ha alcanzado más difusión es el de Chalmers et al³⁹. Hay otros, tan prolijos como el anterior, que se han elaborado más recientemente⁴⁰. Una valoración más simple que la de estos cuestionarios se puede encontrar en Riet et al⁴¹, que estudian la eficacia de la acupuntura en el tratamiento del dolor crónico. En epidemiología observacional no existe ningún cuestionario maestro, ya que éste debe ser adaptado para el problema a estudiar. No obstante, en la literatura existen algunas muestras de los principios a seguir para la evaluación de estudios observacionales. El texto de Elwood⁴² nos ofrece de manera detallada las etapas a seguir en esta evaluación. De manera más sucinta, el metaanálisis de Realini y

Goldzieher⁴³ ofrece un ejemplo de valoración cualitativa de estudios observacionales y resume los criterios desarrollados por Feinstein para evaluar este tipo de estudios. Algo similar sucede con el artículo de Longnecker et al⁸, que nos ofrece una valoración de estudios observacionales, dentro de un metaanálisis de consumo de alcohol y cáncer de mama.

Análisis de datos

1. Identificación de problemas

La aplicación de técnicas estadísticas al MA genera varios problemas que merecen una atención especial y que conducen a una serie de dificultades aún no resueltas. Así, las opiniones sobre la importancia relativa de las pruebas de significación y el uso de los intervalos de confianza varían, aunque es obvio que las pruebas sin estimadores no tienen gran utilidad y que los intervalos de confianza implican términos de probabilidad. La elección del modelo usado para expresar la medida de asociación entre la exposición y el efecto tiene repercusión sobre el valor global de significación. También es relevante el problema de la inferencia estadística múltiple ya que requiere más precaución incluso cuando se realiza un metaanálisis, ya que se suelen analizar varias subseries de datos. Igualmente debe tenerse presente que todos los estudios incluidos en un metaanálisis no son formalmente «independientes» ya que los últimos pueden ser iniciados tras conocer los resultados de los primeros.

2. Elección del modelo estadístico

La descripción detallada de los modelos estadísticos y de los métodos cuantitativos, incluyendo la valoración del sesgo de publicación, aplicados al metaanálisis se realiza en otro apartado⁴⁴. El modelo básico que se emplea es la relación lineal entre las tasas o riesgo de los grupos índice (t) y de comparación (c): $r_{it} = a + b \cdot r_{ic}$ (donde i es el número de estudio). Existen dos casos particulares de especial interés: si $a=0$, resulta que b es igual al riesgo relativo. Se dice que el modelo es multiplicativo. Si $b=1$, resulta que a es igual al riesgo atribuible en los expuestos. El modelo es aditivo.

Se puede asumir que el efecto producido por la exposición sea constante de estudio en estudio o, por el contrario, tenga una distribución al azar a lo largo de los distintos estudios. En el último caso, se debería usar el modelo de efectos aleatorio en vez del más usual, el modelo de efectos fijos. Presenta la ventaja de que admite la variabilidad entre estudios y ayuda a solventar el problema de la heterogeneidad entre estudios. La función logística también admite la variabilidad en los parámetros de población⁹.

Significación sanitaria

En la generalización de resultados hay que tener en cuenta dos hechos esenciales: el cómo fueron los estudios seleccionados y el cómo los sujetos participantes de cada estudio fueron incorporados. La forma en que los estudios fueron seleccionados puede ayudar a eliminar un sesgo de publicación. Si los estudios incorporados no son una muestra representativa del total existentes, el resultado final de la inferencia podría estar sesgado. Un ejemplo de ello fue ofrecido por Vandembroucke⁴⁵. Este autor observó que los estudios sobre un efecto deletéreo de la inhalación pasiva del humo del tabaco en los varones no eran representativos del total existente, ya que apreció una tendencia decreciente del riesgo relativo conforme al tamaño muestral del estudio (medido indirectamente por el error estándar del riesgo relativo) aumentaba.

Para poder extrapolar resultados a un universo es necesario conocer las características poblacionales de los sujetos que fueron incorporados en cada investigación individual. Esto también constituye un aspecto básico a tener en cuenta en la generalización de resultados. Si casi todas las muestras están constituidas por sujetos que reúnen una serie de características comunes o pertenecen a cierto grupo, la inferencia sólo podrá hacerse a ese tipo de poblaciones y no al total de la misma.

A pesar de obtener una asociación significativa entre una exposición y un efecto, esto debe ser revisado. Puede suceder que a pesar de una asociación positiva

y significativa, la relevancia del resultado sea pobre. Esto puede ocurrir cuando el efecto global es de baja magnitud; si hay una falta de certeza en qué tipo de subgrupos se produce la asociación; cuando los resultados no son consistentes (hay heterogeneidad) y cuando ninguno de los estudios individuales alcanzó un resultado significativo. La presencia de heterogeneidad es la situación más frecuente y merece que se discutan algunos de sus aspectos.

Estudio de la heterogeneidad

El concepto de heterogeneidad es central en el MA. La síntesis de estudios realizados en diferentes lugares (países), tiempos, en diferentes poblaciones, y con diferentes diseños, con frecuencia produce el que los resultados sean estadísticamente diferentes entre sí. Esta heterogeneidad no puede ser ignorada, aunque en muchos metaanálisis ni siquiera se tenga en cuenta. Supóngase que todos y cada uno de los resultados de los estudios individuales son ciertos, no están sesgados, y existe una marcada heterogeneidad. En esta situación, el cálculo de un parámetro resumen a lo largo de los diferentes estudios no aporta nada (más bien todo lo contrario, confunde) al conocimiento científico de la asociación. El conocimiento de la media entre los diferentes estudios no permite apreciar la riqueza de la variabilidad existente. Lo más acertado en el caso anterior sería investigar cuáles son las razones que subyacen para que se produzcan resultados contradictorios entre sí⁴⁻²⁸. En la tabla 2 se resumen las principales fuentes de heterogeneidad en un metaanálisis.

Las variaciones en los resultados generados por los estudios a combinar han de ser cuidadosamente valoradas. Pueden ser debidas a dos grandes grupos de características:

1. Características del diseño del estudio: De todos es conocido que cada diseño, a nivel general, posee una serie de ventajas e inconvenientes que influyen de manera clara los resultados que se obtienen. Los problemas que produce un estudio de casos y controles son muy diferentes de los que acarrea un estudio

Tabla 2. Fuentes de heterogeneidad entre varias investigaciones

1. Características del diseño de los estudios:
 - Tipo de diseño
 - Metodología de recogida de la información
 - Tipo de análisis empleado
 - Control de sesgos
2. Características de las poblaciones participantes
 - De la exposición
 - Del efecto
 - De modificadores de la relación exposición-efecto

de cohortes, por ejemplo. En nuestra experiencia, al analizar la asociación entre contraceptivos orales y el cáncer de cuello uterino¹⁸, la heterogeneidad se vio muy disminuida cuando los estudios fueron agrupados en función del tipo de diseño. En este ejemplo, los estudios de cohortes ofrecieron como media valores de riesgo relativo bastante más elevados que los estudios de casos y controles. Con independencia del tipo de diseño general de los estudios que se incluyen en un MA, también influyen las características individuales de cada estudio. Todos los estudios de casos y controles, e igual sucede con cualquier otro tipo de diseño, no son idénticos y muchas veces, ni siquiera similares. Pueden variar en el proceso de selección de la muestra, la recogida de información, el análisis empleado, etc. Por ejemplo, en la recogida de información sobre el consumo de contraceptivos orales no es lo mismo realizar una encuesta telefónica que una entrevista personal. La segunda permite utilizar técnicas de reforzamiento de la memoria como, por ejemplo, el uso de fotografías en color de todos y cada uno de los preparados que la mujer ha podido utilizar. Esta variable u otras propias del diseño podrían ser incorporadas en el análisis e intentar justificar la heterogeneidad encontrada.

El hecho de que los estudios de epidemiología observacional muestren una mayor variabilidad en su diseño ha motivado que algunos autores hayan sugerido que el metaanálisis se reserve para los estudios experimentales. Este tipo de diseños son más estrictos en su metodología que los anteriores. Sin embargo, ésta es una actitud cerrada que impide que otros terrenos de la epidemiología, y en definitiva la sociedad, se beneficien de las técnicas del metaanálisis.

2. Características de las poblaciones participantes: Es la segunda gran fuente de heterogeneidad en una asociación. No todas las poblaciones de estudio de cada una de las investigaciones recopiladas tienen porqué ser comparables. Tres son las características generales a nivel de la población participante que pueden resultar en una heterogeneidad: la exposición, representación de ciertos subgrupos de especial riesgo, y el tipo de efecto⁴⁻⁹.

Por último debe destacarse la importancia que pueden tener las características de la exposición, el asentamiento del estudio, y la distribución de presuntos modificadores de efecto en la explicación de la heterogeneidad del efecto. Por ello, son también variables a ser tenidas en cuenta en el análisis de los datos en un metaanálisis. El análisis de la heterogeneidad, en ocasiones, no la relaciona con las características mencionadas con anterioridad, sino que parece ser debida a un error aleatorio, motivado porque una o varias investigaciones encuentran asociaciones que difieren de la media. Son los «fuera de la normalidad» (outliers) sin razón aparente. Esto se observó en el metaanálisis entre anticonceptivos orales y cáncer de mama, en el que dos estudios encontraron un riesgo relativo anormalmente elevado¹⁷. Lo aconsejable en estos casos es suprimir estas investigaciones del análisis conjunto.

Impacto y futuro del Metaanálisis

La contribución del metaanálisis en la investigación epidemiológica podrá apreciarse en los próximos años, a medida que aumente su utilización. Por el momento, parece bien establecido que su uso puede resultar de interés en las siguientes situaciones: mejora de las revisiones;

la planificación de los estudios futuros; decisión sobre la asociación entre una exposición y un efecto; si un MA de varios estudios, de los que al menos uno alcanzó significación estadística, obtiene un resultado estadísticamente significativo, el MA fortalece la evidencia global; Si un MA de varios estudios, de los que ninguno alcanzó significación estadística, obtiene un resultado estadísticamente significativo, no es recomendable actuar en base a su resultado. Quizás sea necesario estudios de mayor tamaño; si un MA de varios estudios, de los que al menos uno alcanzó significación estadística, obtiene un resultado estadísticamente no significativo, habrán de ser reconsiderados los criterios de selección de los distintos estudios o realizar nuevos estudios. Por último no debe olvidarse su posible contribución en la enseñanza de la epidemiología y de las ciencias de la salud.

Bibliografía

1. Glass CV. Primary, secondary, and meta-analysis of research. *Educ Res* 1976; 5: 3-9.
2. MacMahon B, Hutchison GB. Prenatal X-ray and childhood cancer: a review. *Int J Cancer* 1964; 20: 1172-1174.
3. Woolf SH, Battista RN, Anderson GM, Logan AG, Wang E, and the Canadian Task Force on the Periodic Health Examination. Assessing the clinical effectiveness of preventive maneuvers: analytic principles and systematic methods in reviewing evidence and developing clinical practice recommendations. *J Clin Epidemiol* 1990; 43: 891-905.
4. Light RJ, Pillemer DB. Summing up. The science of reviewing research. Cambridge (MA): Harvard University Press, 1984.
5. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177-188.
6. Bulpitt CJ. Meta-analysis. *Lancet* ii: 93-94.
7. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; 316: 150-155.
8. Longnecker MP, Berlin JA, Orza MJ, Chalmers TC. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA* 1988; 260: 652-656.
9. Boissel J-P, Blanchard J, Panak E, Peyrieux J-C, Sacks H. Considerations for the meta-analysis of randomized clinical trials (Summary of a panel discussion). *Controlled Clin Trials* 1989; 10: 254-281.
10. Jenicek M. Meta-analyse en médecine. Québec: Edisem, 1987.
11. Delgado Rodríguez M, Rodríguez-Contreras R, Sillero Arenas M, Lardelli Claret P, Bueno Cavanillas A, Gálvez Vargas R. Contraceptivos orales y cánceres de mama y cuello uterino. Metaanálisis cualitativo. *Tokogin Práct* 1990; 46: 336-344.

12. Sillero Arenas M, Rodríguez-Contreras R, Delgado Rodríguez M, Bueno Cavanillas A, Gálvez Vargas R. Oral contraceptives and cancer of the cervix uteri. Patterns of research. *Acta Obstet Gynecol Scand* 1991 (en prensa).
13. Tirosh, E, Rabino S. Physiotherapy of children with cerebral palsy. *AJDC* 1989; 149: 552-555.
14. Baumgarten M. The health of persons giving care to the demented elderly: a critical review of the literature. *J Clin Epidemiol* 1989; 42: 1137-1148.
15. Gómez Olmedo M, Fernández Sierra MA, Delgado Rodríguez M, Gálvez Vargas R. Infección por VIH en la población española. I. Metaanálisis cualitativo. *Med Clin (Barc)* 1990; 95: 286-291.
16. Fernández Sierra MA, Gómez Olmedo MR, Delgado Rodríguez M, Gálvez Vargas R. Infección por VIH en la población española. II. Metaanálisis de las tendencias temporales y geográficas. *Med Clin (Barc)* 1990; 95: 366-371.
17. Delgado-Rodríguez M, Sillero-Arenas M, Rodríguez-Contreras R, López R, Gálvez-Vargas R. Oral contraceptives and breast cancer. A meta-analysis. *Rev Epidemiol Santé Publ* 1991; 39: 165-181.
18. Delgado-Rodríguez M, Sillero-Arenas M, Rodríguez-Contreras R, Martín-Moreno JM, Gálvez-Vargas, R. Oral contraceptives and cancer of the cervix uteri. A meta-analysis. *Eur J Epidemiol* 1991 (in press).
19. Gianrossi R, Detrano R, Mulvihill DF, Lehmann, K, Dubach P, Colombo A, McArthur D, Froelicher V. Exercise-induce ST depression in the diagnosis of coronary artery disease. A meta-analysis. *Circulation* 1989; 80: 87-98.
20. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovas Dis* 1985; 27: 335-371.
21. Velanovich V. Crystalloid versus colloid fluid resuscitation: a meta-analysis of mortality. *Surgery* 1989; 105: 65-71.
22. Ewertz M, Duffy SW, Adami H-O, Kvale G, Lund E, Meirik O, Møllergaard A, Soini I, Tulinius H. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the nordic countries. *Int J Cancer* 1990; 46: 597-603.
23. Rodríguez-Coronado R, Delgado Rodríguez M, Sillero Arenas M, Lardelli Claret P, Gálvez Vargas R. Contraceptivos orales y cáncer de cuello uterino: análisis del poder de la asociación. *Rev San Hig Públ* 1991 (en prensa).
24. Delgado Rodríguez M, Sillero Arenas M, Rodríguez-Contreras R, Bueno Cavanillas A, Martín Moreno JM, Gálvez Vargas R. Anticonceptivos orales y cáncer de mama: análisis de la potencia estadística de asociación. *Gac Sanit* 1991; 5: 59-67.
25. Bracken MB. Oral contraception and congenital malformations in offspring: a review and meta-analysis of the prospective studies. *Obstet Gynecol*. 76: 552-557.
26. Fasoli M, Parazzini F, Cecchetti G, La Vecchia C. Post-coital contraception: an overview of published studies. *Contraception* 1989; 39: 459-468.
27. Herrera CR, Lynch C. Is baldness a risk factor for coronary artery disease? A review of the literature. *J Clin Epidemiol* 1990; 43: 1255-1260.
28. Light RJ. Accumulating evidence from independent studies: what we can win and what we can lose? *Stat Med* 1987; 6: 221-228.
29. Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. *Stat Med* 1987; 6: 315-326.
30. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst* 1989; 81: 107-115.
31. Koren G, Graham K, Shear H, Einarson T. Bias against the null hypothesis: the reproductive hazards of cocaine. *Lancet* 1989; ii: 1440-1442.
32. Howe GR, Hirohata T, Hislop TG, Iscovich JM, Yuan J-M, Katsouyanni K, Lubin F, Marubini E, Modan B, Rohan T, Toniolo P, Schunzhang Y. Dietary factors and risk of breast cancer: combined analysis of 12 case-control studies. *JNCI* 1990; 82: 561-569.
33. Dickersin K, Chalmers TC, Simes RJ, Boissel JP, Meinert CL. Report from the panel on the case for registers of clinical trials. *Controlled Clin Trials* 1988; 9: 76-81.
34. Dickersin K, Higgins K, Meinert CL. Identification of meta-analyses. The need for standard terminology. *Controlled Clin Trials* 1990; 11: 52-66.
35. Ozminkowski RJ, Wortman PM. Inborn/outborn status and neonatal survival: a meta-analysis of non-randomized studies. *Stat Med* 1988; 7: 1207-1221.
36. Shinton R, Beevers G. Meta-analysis of relation between cigarette smoking and stroke. *Br Med J* 1989; 298: 798-794.
37. Berlin JA, Colditz GA. A Meta-analysis of physical activity in the prevention of coronary heart disease. *Am J Epidemiol* 1990; 132: 612-628.
38. Delgado Rodríguez M, Sillero Arenas M, Rodríguez-Contreras R, Gálvez Vargas R. Pertinencia y normas para la realización de un estudio sobre anticonceptivos orales y cáncer de mama en España. *Toko-Gin Práct* 1991 (en prensa).
39. Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981; 2: 31-49.
40. Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics* 1989; 84: 815-827.
41. Riet G ter, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol* 1990; 43: 1191-1199.
42. Elwood JM. Casual relationships in medicine: a practical appraisal. Oxford: Oxford University Press, 1988.
43. Realini JP, Goldzieher JW. Oral contraceptives and cardiovascular disease: a critique of the epidemiologic studies. *Am J Obstet Gynecol* 1985; 152: 729-798.
44. Delgado Rodríguez M, Sillero Arenas M, Rodríguez-Contreras R, Gálvez Vargas R. Metaanálisis en epidemiología. II. Métodos cuantitativos. *Gac Sanit* (en prensa).
45. Vandenbroucke JP. Passive smoking and lung cancer: a publication bias? *Br Med J* 1988; 296: 391-392.

