

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/342723653>

Metaanálisis: una forma básica de entender e interpretar su evidencia

Article in *Revista de Senología y Patología Mamaria* · July 2020

DOI: 10.1016/j.senol.2020.05.007

CITATIONS

0

READS

213

4 authors, including:



Antoni Lluca

Hospital Universitari General de Castellón

65 PUBLICATIONS 212 CITATIONS

[SEE PROFILE](#)



Laura Granel Villach

Hospital Universitari General de Castellón

26 PUBLICATIONS 122 CITATIONS

[SEE PROFILE](#)



Manuel Bellver Oliver

Hospital Universitario General de Castellón

31 PUBLICATIONS 172 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:

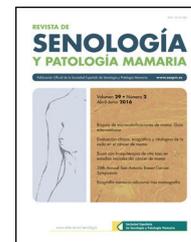


Advanced ovarian cancer outcomes [View project](#)



Revista de Senología y Patología Mamaria

www.elsevier.es/senologia



ARTÍCULO DOCENTE

Metaanálisis: una forma básica de entender e interpretar su evidencia

Vicente Javier Escrig Sos^{a,*}, José Antonio Lluca Abella^b, Laura Granel Villach^a y Manuel Bellver Oliver^a

^a Servicio de Cirugía General y Digestiva, Hospital General Universitario, Castellón, España

^b Servicio de Ginecología y Obstetricia, Hospital General Universitario. Universitat Jaume I, Castellón, España

Recibido el 18 de abril de 2020; aceptado el 11 de mayo de 2020

PALABRAS CLAVE

Metaanálisis;
Revisión sistemática;
Sesgo de publicación;
Heterogeneidad

Resumen El metaanálisis es un método sistemático para sintetizar resultados de diferentes estudios empíricos sobre el efecto de una variable independiente, sea de intervención o tratamiento, en un resultado final preciso. Desarrollado principalmente en la investigación médica y psicológica como una herramienta para sintetizar información empírica sobre los resultados de un tratamiento, ahora el metaanálisis se usa cada vez más en las ciencias de la salud, formando parte de las revisiones sistemáticas, como una herramienta de inferencia estadística. Sin embargo, a pesar de sus indiscutibles propiedades para la síntesis de la información dispersa en la literatura y para la resolución de controversias con un alto grado de evidencia, adolece de distintos problemas en la consecución de estos objetivos. Es necesario conocer su entramado metodológico, aunque sea a un nivel básico, para valorar su validez en el logro de dichos objetivos.

© 2020 SESPM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

KEYWORDS

Meta-analysis;
Systematic review;
Publication bias;
Heterogeneity

Meta-Analysis: A Basic Way To Understand And Interpret Your Evidence

Abstract Meta-analysis is a systematic method to synthesise results from different empirical studies on the effect of an independent variable, whether an intervention or treatment, on a precise outcome. These studies were mainly developed in medical and psychological research as a tool to synthesise empirical information on the results of treatment. They are currently increasingly used in health sciences, forming part of systematic reviews, as a tool for statistical inference. However, despite their undoubted usefulness for synthesising information scattered in the literature and for resolving controversies with a high grade of evidence, these studies also show various problems in achieving their aims. Familiarity with their methodological framework is essential, even at a basic level, to assess their validity in achieving their objectives.

© 2020 SESPM. Published by Elsevier España, S.L.U. All rights reserved.

* Autor para correspondencia.

Correo electrónico: javierescrig@icloud.com (V.J. Escrig Sos).

<https://doi.org/10.1016/j.senol.2020.05.007>

0214-1582/© 2020 SESPM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introducción

El metaanálisis es una técnica estadística que combina y resume los resultados de varios estudios individuales. En medicina clínica lo que resume es la diferencia de efecto entre dos intervenciones médicas, generalmente algún tratamiento reciente y potencialmente útil frente a otro tratamiento ya contrastado y asentado en la práctica médica habitual. Al combinar varios estudios se produce un notable aumento del tamaño de la muestra, lo cual lleva a obtener más fácilmente un resultado final estadísticamente significativo y más preciso: su banda de fluctuación, o Intervalo de Confianza poblacional, será más estrecha. Desde varios estudios, incluso estudios no concluyentes, se puede obtener así un resultado estadísticamente concluyente y que, al menos teóricamente en cuanto a su estimación puntual, estaría muy cercano al verdadero valor existente en la población en el aspecto que se analiza.

La definición moderna de metaanálisis proviene de la psicometría y data del trabajo de Glass¹ publicado en 1976, es decir, ya arrastra una historia y una evolución de más de 40 años². En aquel artículo se manifiesta la imperiosa necesidad existente para la comunidad científica de poder resumir la ingente cantidad de información dispersa en la literatura. Para ser fidedigno, dicho resumen se debe hacer de la forma más sistemática y reglada posible. Por ello, el metaanálisis forma parte de la llamada revisión sistemática de la literatura³. Se podría decir que es la parte estadística inferencial de la revisión sistemática (fig. 1), aquella que resume y extrapola a la población general los resultados extraídos de la muestra analizada. Por ello, la calidad de un metaanálisis está directamente ligada a la calidad

metodológica de la revisión sistemática, a la calidad de la selección de estudios que ella produce, y a la calidad de los propios estudios que la integran, generalmente ensayos clínicos aleatorizados, en los que se analiza la diferencia de efecto producida entre dos intervenciones médicas del tipo que sea. Un buen metaanálisis ocupa pues el escalón más elevado en la Medicina basada en la evidencia.

Básicamente, existen dos grandes tipos de metaanálisis convencional: el metaanálisis de datos agregados y el metaanálisis de datos individuales, cada uno de ellos con diversas variantes en cuanto a su tratamiento estadístico. El metaanálisis de datos individuales es el más indicado, pero es también el que ostenta en la práctica enormes dificultades en su realización, especialmente en lo referente a la recogida de datos de pacientes individuales de múltiple origen. Por ese motivo el metaanálisis de datos agregados es el más prevalente en la literatura y se le puede considerar como el «clásico» para los clínicos. Parte de recoger las medidas estadísticas de resumen de cada uno de los estudios que combina (medias, proporciones, etc.), que deben estar debidamente plasmadas como variables en una tabla de datos (fig. 2) a la que se aplicarán las técnicas de análisis estadístico que las combinarán y resumirán en un resultado final. Este es el tipo de metaanálisis que nos ocupa en este artículo de divulgación.

Información, fortalezas y limitaciones

No se debe cometer la inconsciencia de asumir sin más que un metaanálisis ya equivale de por sí a evidencia definitiva. El metaanálisis tiene diversos problemas que generalmente proceden de los factores de confusión que puedan contener

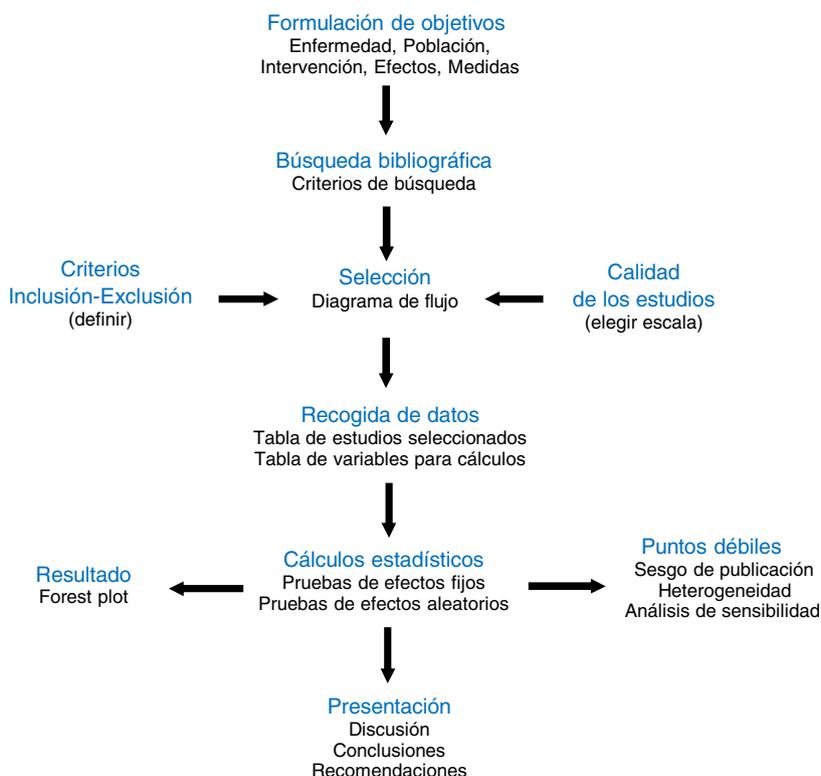


Figura 1 Etapas de una Revisión sistemática y Metaanálisis.

A

Estudio	Año	(Tratamiento Experimental)			(Tratamiento Control)		
		Total casos	Media	D.E.	Total casos	Media	D.E.
Autor 1	1999	50	12	8	60	6	10
Autor 2	2001	75	20	6	85	4	15
Autor 3	2005	90	10	4	120	12	6

El resultado se puede reportar como:

- Diferencia media entre Experimental y Control
- Diferencia media estandarizada (Tamaño del Efecto: Cohen, Hedges, Glass)

B

Estudio	Año	(Tratamiento Experimental)		(Tratamiento Control)	
		Sucesos	No Sucesos	Sucesos	No Sucesos
Autor 1	1999	30	125	25	130
Autor 2	2001	48	225	66	245
Autor 3	2005	44	135	55	139

El resultado se puede reportar como:

- Diferencia de proporciones entre Experimental y Control.
- Odds Ratio
- Riesgo Relativo

Figura 2 Disposición habitual de las variables en una tabla de datos para metaanálisis
 A: ejemplo de tabla básica para analizar diferencias de efectos de tipo cuantitativo.
 B: ejemplo de tabla básica para analizar diferencias de efectos de tipo cualitativo.
 D.E.: desviación estándar.

los estudios que combina y resume, factores que si existen se trasladan siempre al metaanálisis, de la mejor o peor selección de estudios que se hace en la revisión sistemática y los derivados del propio análisis estadístico. Se trata de sesgos que de alguna manera se deben identificar, porque si existen, lo que hacen es alejar el resultado final del verdadero resultado existente en la población⁴. No obstante, aunque tal aproximación al verdadero resultado en la población se logre en mayor o menor medida, un buen metaanálisis puede despejar incógnitas y siempre constituirá una sinopsis que ahorra largas horas de búsqueda y lectura de la literatura. Dentro de la valoración de la calidad de un metaanálisis también es muy importante comprobar que cumpla con unas determinadas normas de publicación⁵⁻⁷ como la mejor manera de que el lector pueda identificar sus fortalezas y limitaciones.

Entre los diversos problemas que pueden aparecer hay dos que son capitales, y que de forma constante se analizan en todos los artículos. Las normas de publicación obligan. Se trata del Sesgo de publicación⁸ y de la heterogeneidad⁹ del resultado final, que afectan tanto a la validez interna como a la validez externa del metaanálisis.

El sesgo de publicación

Puede producirse cuando en la selección de artículos se recogen mayoritariamente grandes ensayos con resultados concluyentes, es decir, estadísticamente significativos, y se excluyen ensayos no concluyentes que frecuentemente presentan resultados de escasa magnitud, por lo que no suelen ser publicados, al menos en revistas importantes de lengua inglesa. Las políticas editoriales de las revistas científicas y las políticas de las empresas farmacéuticas favorecen este fenómeno, así como la búsqueda de artículos en sólo una o pocas bases de datos bibliográficas. Todo ello influye en que este tipo de ensayos tengan una menor probabilidad de

ser seleccionados en la revisión sistemática. Se cree que en algunos metaanálisis la no inclusión de estudios no concluyentes puede incrementar artificialmente hasta seis veces el tamaño de las diferencias finales⁴. Para aminorar el sesgo de publicación en los resultados de un metaanálisis se debe hacer un esfuerzo adicional para localizar estudios no publicados y estudios no concluyentes. Identificar estudios no publicados es ahora algo más fácil, gracias a una mejor comunicación entre los investigadores de todo el mundo, y gracias a los registros en los que constan todos los estudios de una determinada enfermedad o tratamiento, independientemente del resultado⁴.

La identificación de un sesgo de publicación se hace habitualmente sobre unas gráficas en forma de embudo o *funnel plot*^{10,11}. Son unas gráficas de dispersión entre alguna medida relacionada con el error estándar (tamaño de la muestra) y los resultados de cada ensayo incluido que aparecen como puntos o círculos alrededor del resultado del metaanálisis (fig. 3).

La figura 3A es un ejemplo de ausencia de sesgo de publicación. Tanto los puntos correspondientes a los ensayos con gran muestra, como a aquellos con muestras menores, se distribuyen homogéneamente dentro del triángulo y a cada lado de la línea de puntos que marca el resultado del metaanálisis. La figura 3B es un ejemplo de la existencia de un importante sesgo de publicación. Los puntos correspondientes al resultado de los distintos ensayos incluidos no se distribuyen homogéneamente dentro del triángulo ni alrededor del resultado del metaanálisis. En este ejemplo se puede observar que no han sido incluidos en el metaanálisis ensayos con menor muestra y con diferencias de signo positivo, seguramente no concluyentes debido a lo escaso de su muestra. Esto hace que el resultado final se vea atraído sesgadamente hacia unas diferencias de signo negativo.

La interpretación gráfica del sesgo de publicación es una buena manera de analizar su existencia. Todo lo que

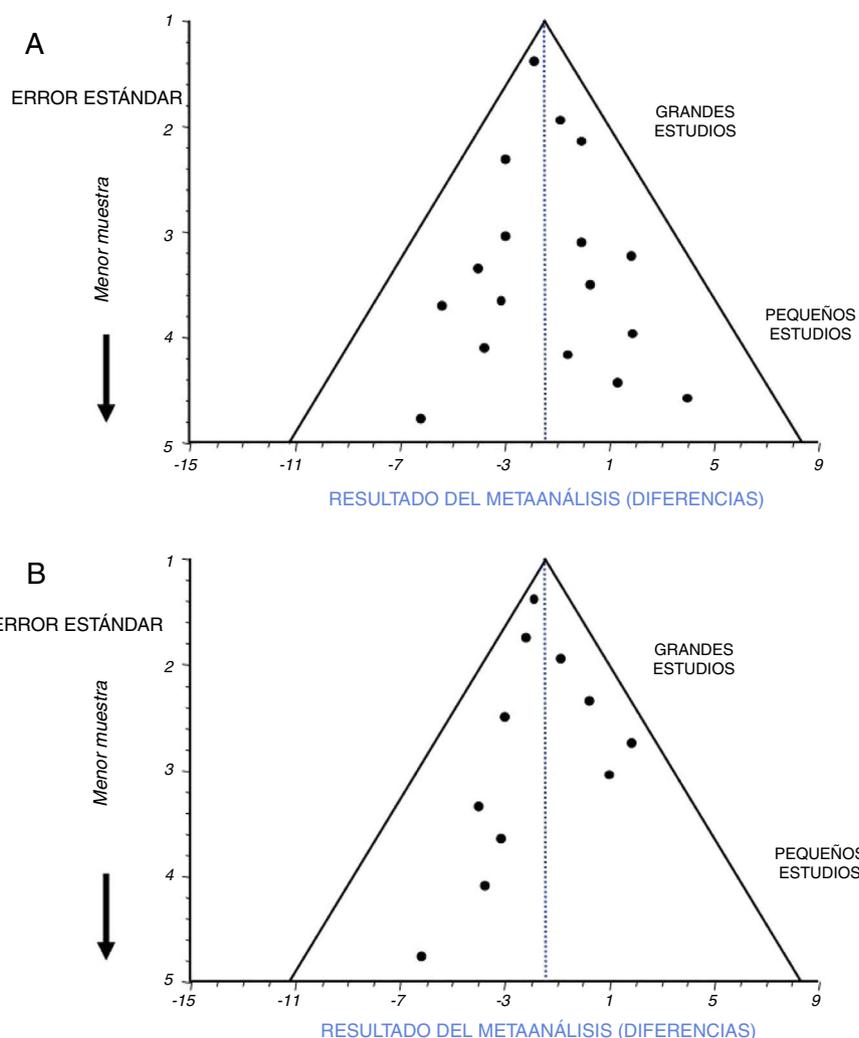


Figura 3 Funnel plot para escrutar la existencia de sesgo de publicación.

no se vea como en la figura 3A indica sesgo de publicación en mayor o menor medida. Junto con la gráfica Funnel se suele reportar alguna prueba de hipótesis orientada a este sesgo, con un valor p que indicará sesgo de publicación cuando sea estadísticamente significativo¹². La incertidumbre aparece cuando ese valor p no es significativo. Entonces, no se puede confirmar este sesgo, pero tampoco se confirma que esté descartado. Es en este caso cuando más útil resulta una interpretación visual de la gráfica Funnel. Sin embargo, cuando en un metaanálisis se hayan incluido pocos ensayos la gráfica Funnel mostrará pocos puntos, lo cual dificultará su interpretación, y también será más probable que las pruebas de hipótesis ofrezcan valores p no significativos, de modo que podemos volver a quedarnos en la incertidumbre de saber si este sesgo está realmente presente. En estos casos más dudosos pueden ser de ayuda los llamados métodos «Trim and Fill», y los Índices de tolerancia^{13,14} que indican el hipotético número de ensayos pequeños o no concluyentes que sería necesario introducir en el metaanálisis para que hubiese un cambio notable en el valor del resultado final y en su significación estadística. Lógicamente, si ese número es pequeño nos estará señalando un posible sesgo de publicación. Así pues,

puede ocurrir que en algunas ocasiones sea complicado valorar la presencia y el alcance de este importante sesgo.

La heterogeneidad

Se refiere al grado en que los resultados individuales de los ensayos incluidos en el metaanálisis difieren entre sí. Cuando es elevada, tal variabilidad lo que produce de nuevo es la total inseguridad de que el resultado del metaanálisis se corresponda bien con el verdadero valor existente en la población, porque nos viene a decir que posiblemente en el metaanálisis hemos incluido realmente diferentes cuestiones de investigación: «estamos mezclando peras con manzanas». Se trata pues de un aspecto fundamental al valorar la validez del resultado del metaanálisis.

Actualmente, su existencia se suele escrutar mediante la obtención de un estadístico llamado I-cuadrado (I^2) que se reporta en forma de porcentaje, de 0 a 100%, y se acompaña de un valor p. Valores progresivamente superiores a un tercio (33%) ya nos van indicando un grado progresivo de heterogeneidad a considerar, que se verá confirmada si supera el 50% y si además su valor p es estadísticamente significativo.

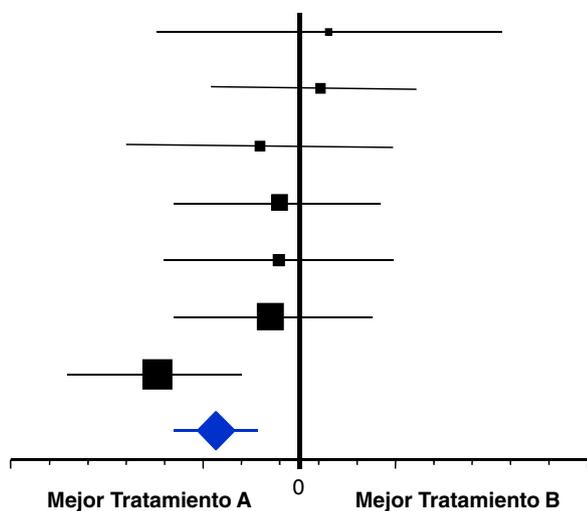


Figura 4 Gráfica Forest.

Al igual que con el sesgo de publicación, la heterogeneidad es posible percibirla desde los gráficos que se utilizan para mostrar el resultado del metaanálisis, las gráficas «en árbol» o *forest plot* (fig. 4). Una gráfica Forest contiene diversos elementos que deben conocerse al detalle.

Consta de un eje horizontal que marca la gama de posibles diferencias de efecto del metaanálisis, así como a qué tipo de intervención favorecen. El eje vertical situado centralmente es la línea de diferencias inexistentes, es decir, señala la hipótesis nula. Se marca con un 0 cuando las diferencias se refieren a medias o porcentajes, y con un 1 cuando corresponden a medidas relativas del riesgo (Odds ratio, Riesgo relativo, o Hazard ratio). Los cuadrados negros, en su centro, señalan la estimación puntual de las diferencias halladas en cada uno de los ensayos incluidos en el metaanálisis. Estos cuadrados se acompañan de una línea horizontal que corresponde a su intervalo de confianza. La mayor o menor dimensión de los cuadrados nos indica el peso que tiene el resultado de cada ensayo sobre el resultado final del metaanálisis, que será mayor cuanto mayor sea su tamaño de muestra y menor sea su variabilidad. Cuando estas líneas cruzan el eje vertical del «empate» nos indica que el resultado de tal ensayo no es concluyente, es decir, no es estadísticamente significativo, y viceversa. En la parte más inferior se representa un rombo o diamante (azul en el ejemplo) que indica el valor de la estimación puntual del resultado del metaanálisis, acompañado también de su intervalo de confianza que nos servirá para constatar su significación estadística siempre que no cruce el eje vertical.

La mayor o menor heterogeneidad de un metaanálisis puede ser escrutada visualmente a través de estas gráficas Forest. Denotan heterogeneidad cuando existen ensayos cuyos resultados puntuales (cuadrados) favorecen a una intervención y otros favorecen a la contrapuesta. También cuando la estimación puntual obtenida en el metaanálisis (diamante) difiere significativamente del resultado particular de uno o varios ensayos incluidos (fig. 5).

La figura 5A es un ejemplo de ausencia de una heterogeneidad que sea importante o a tener en cuenta. Todos los resultados de los ensayos incluidos favorecen a una de las intervenciones. Además, si trazamos una línea imaginaria (la

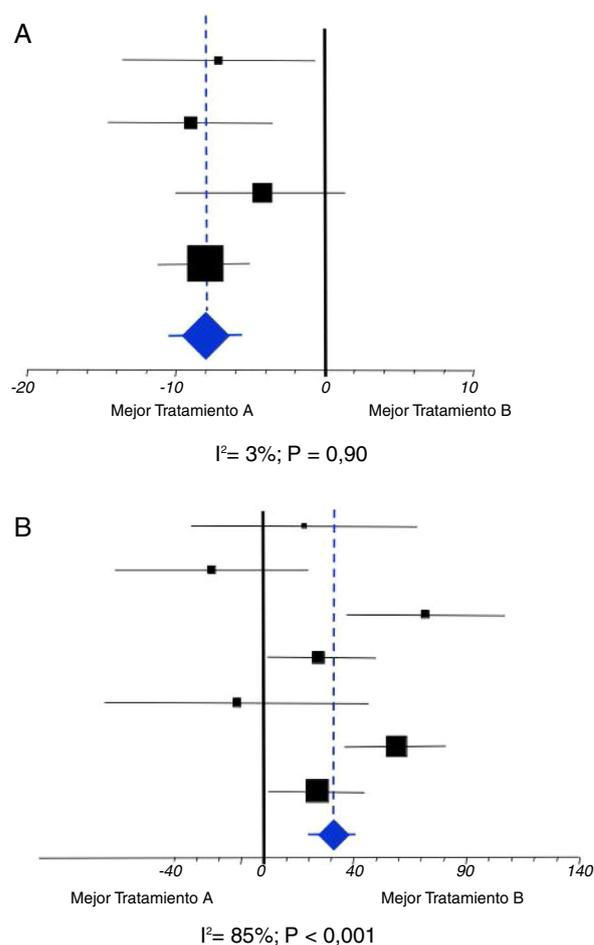


Figura 5 Escrutinio de la heterogeneidad a través de las gráficas Forest.

línea azul discontinua) desde el resultado final del metaanálisis se puede observar que dicho resultado no difiere estadísticamente del resultado de ningún ensayo individual puesto que cruza todos los intervalos de confianza de los ensayos incluidos. Por otra parte, la estimación puntual de I^2 es muy baja y sin significación estadística: es compatible con el valor 0 de I^2 . La figura 5B es un ejemplo de todo lo contrario, de una enorme heterogeneidad. Se puede observar que hay ensayos incluidos en el metaanálisis cuyos resultados favorecen a una u otra intervención, y que incluso hay tres ensayos cuyo resultado difiere significativamente del resultado final del metaanálisis dado que la imaginaria línea azul no cruza sus intervalos de confianza. Lógicamente, aquí la estimación puntual de I^2 será muy elevada, y estadísticamente significativa: I^2 difiere de 0.

La heterogeneidad se produce cuando entre los ensayos incluidos en el metaanálisis hay discordancias metodológicas: cuando no tienen exactamente el mismo objetivo aunque parezca similar; cuando las poblaciones de pacientes y la enfermedad no son exactamente iguales en sus características; cuando las variables de resultado utilizadas no están definidas exactamente igual, ni están medidas exactamente de la misma manera; cuando la intervenciones aplicadas a los pacientes no son exactamente las mismas aunque lleven el mismo nombre y cuando alguno de los ensayos incluidos presentan algún sesgo en su resultado.

También se ve favorecida cuando se incluyen ensayos no aleatorizados cuyos resultados pueden estar lastrados fácilmente por diversos sesgos.

En ocasiones, cuando la gravedad de la enfermedad no es la misma en todos los ensayos incluidos, conviene hacer un metaanálisis estratificado para cada nivel de gravedad del proceso estudiado, con el fin de eliminar el alto grado de heterogeneidad que proporcionaría en este escenario un metaanálisis global de todos los ensayos. Para la estratificación como herramienta de control de la heterogeneidad también se suele usar el tamaño de los ensayos, su nivel de calidad, como, por ejemplo, ensayos aleatorizados frente a no aleatorizados, su año de publicación, etc., es decir, aquellos factores metodológicos que puedan no ser homogéneos en los ensayos incluidos.

El análisis de sensibilidad, no siempre incluido en los artículos, es otra herramienta de control de la heterogeneidad. Consiste en realizar sucesivos metaanálisis excluyendo cada vez uno de los ensayos seleccionados. Si alguno de los resultados desentona fuertemente con el resto se habrá identificado el ensayo que aporta un mayor grado de heterogeneidad, ensayo que tal vez debería ser excluido del metaanálisis. Es un mecanismo similar en cierto modo al que utiliza el llamado metaanálisis acumulado que consiste en calcular el resultado añadiendo cada vez un ensayo de los seleccionados, pero siguiendo un orden que suele ser el año de publicación, así hasta obtener el resultado final cuando se hayan analizado todos los ensayos seleccionados. La metarregresión es otra técnica estadística en la que este control se ejerce mediante un ajuste por una o varias variables concretas, no por ensayos.

Una parte de la heterogeneidad puede controlarse a través del propio análisis estadístico efectuado para combinar los resultados de los ensayos individuales que se resumen en el resultado final del metaanálisis. En este sentido, hay dos grupos principales de pruebas estadísticas para el tratamiento estadístico de los datos: las Pruebas de Efectos Fijos, y las Pruebas de Efectos Aleatorios, cada cual con distintas variantes a elegir. Las pruebas de efectos fijos calculan el resultado final del metaanálisis promediando el resultado individual de cada ensayo con ponderaciones únicamente basadas en su tamaño de muestra y en la variabilidad de su propio resultado. Las pruebas de efectos aleatorios calculan el resultado final ponderando además por la variabilidad existente entre los resultados de todos los ensayos incluidos. Estas pruebas o modelos pueden proporcionar un resultado final no coincidente si existe cierta heterogeneidad, pero siempre será más preciso (intervalo de confianza menos amplio) en las pruebas de efectos fijos que las pruebas de efectos aleatorios.

En principio, las pruebas de cálculo de efectos fijos se deberían usar preferentemente cuando los ensayos incluidos son metodológicamente muy similares, mientras que las pruebas de efectos aleatorios tendrían una mayor indicación cuando los ensayos incluidos muestren las discrepancias metodológicas antes mencionadas. Los artículos sobre metaanálisis deben incluir alguna tabla que resuma las características de las variables de cada uno de los ensayos y la calidad metodológica de los mismos, como información descriptiva y como justificación en la elección del método de cálculo estadístico empleado. No obstante, en la práctica lo que se suele hacer muchas veces es obtener el resultado

final aplicando ambos tipos de pruebas. En el caso de que ambos resultados difieran mucho en su magnitud se debería tener en cuenta el obtenido mediante las pruebas de efectos aleatorios, aunque la diferencia final esté sujeta a mayor fluctuación poblacional. Este proceder, en el fondo, es como una especie de barrera de defensa ante la obtención de un resultado heterogéneo.

Una mención especial merece los metaanálisis relacionados con resultados quirúrgicos. En los metaanálisis de cirugía la heterogeneidad suele estar muy presente, quizás porque los ensayos que se incluyen son propensos a esconder sesgos metodológicos de diversa naturaleza. Entre otras cosas, en ellos es difícil el cegamiento y muchas de sus variables no tienen una escala de medida objetiva y uniforme. También puede faltar el fundamental principio de incertidumbre, que es una especie de conflicto de intereses, es decir, el cirujano tiene de entrada preferencia por favorecer determinados tipos de técnicas, como actualmente podrían ser las relacionadas con la laparoscopia. Esto influye tanto en el desarrollo del estudio como en el sentido y magnitud del resultado. Además, los ensayos clínicos en cirugía son escasos, con frecuencia tienen una pequeña muestra que hace que la aleatorización no controle perfectamente la existencia de factores de confusión, y no destacan excesivamente en cuanto a su calidad metodológica global¹⁵. Por lo tanto, y en general, los metaanálisis sobre temas quirúrgicos merecen siempre una valoración crítica muy exhaustiva.

Pasos fundamentales en la interpretación de un metaanálisis

Son dos las grandes preguntas que nos debemos hacer cuando nos enfrentamos a la lectura de un artículo de metaanálisis: ¿tiene problemas metodológicos, es decir, su validez interna es lo suficientemente correcta? ¿tiene para nosotros validez externa, es decir, lo que nos informa es aplicable a nuestros pacientes?

Para la primera pregunta, ante todo, hemos de valorar si la información que aporta el metaanálisis está correctamente expuesta y estructurada. Para ello, es bueno disponer de las normas de publicación existentes⁷ y comprobar que se cumplen todas, es decir, que se aporta toda la información necesaria para su valoración. Hay que desconfiar de la calidad metodológica del metaanálisis si esto no se da. Después, hemos de fijarnos en la tabla de las características de los estudios incluidos y de su calidad, para que no surja un posible sesgo de selección de ensayos que sólo favorecen a uno de los tratamientos confrontados. Para ello, es útil escrutar el diagrama de flujo de la búsqueda bibliográfica que nos debe informar de los criterios de búsqueda y de los rasgos y motivos de inclusión y exclusión de los estudios encontrados en la bibliografía. Por último, se debe comprobar la posible presencia del sesgo de publicación según las directrices antes expuestas que, de existir, podría haber alterado la magnitud y la dirección del resultado.

Si la primera pregunta tiene una respuesta favorable, ya se puede pasar a valorar la siguiente, la que se refiere a la validez externa. Aquí, lo que suele mirarse ante todo es la presencia de heterogeneidad. Es lo más importante, porque si existe y es elevada el resultado no será fiable en absoluto para nosotros, ni para nadie. Puede ser útil leer

qué dicen los autores sobre ello en la discusión del artículo, si es que comentan las posibles causas o fuentes de dicha heterogeneidad y los puntos débiles del estudio.

Suponiendo que la heterogeneidad sea inexistente o asu- miblemente baja, ya sólo queda ver el resultado, y aquí es más conveniente fijarse en su magnitud y en la importan- cia clínica que conlleva que en su significación estadística. La consideración de la importancia clínica de un resultado no es algo que nos resuelva la estadística sino que depende de nuestro propio criterio y conocimiento sobre el problema que se analiza, y se basa en valorar, dada la magnitud de la diferencia de efectos, si estaría justificado (balance riesgo- beneficios) cambiar nuestra intervención habitual por la que se ve favorecida en el metaanálisis, o persistir en el uso de esta si es la que se ha utilizado. El mejor hallazgo sería observar un resultado no heterogéneo que consideremos clínicamente importante y que sea estadísticamente signifi- cativo, ofrecido por un metaanálisis metodológicamente correcto en líneas generales.

Un último paso, pero no menos importante, para la valo- ración de la validez externa del metaanálisis es retornar al escrutinio de los rasgos de los pacientes incluidos y de la enfermedad. Ningún resultado nos servirá si los rasgos de nuestros pacientes difieren de los que se muestran en el metaanálisis, ni si el grado habitual de la enfermedad que tratamos es distinta, ni tampoco por añadidura si la aten- ción médica que en general aplicamos a nuestros casos y sus resultados pudieran ser distintos al contexto del metaanáli- sis. Si se dan estas discrepancias nunca podremos reproducir el resultado del metaanálisis en nuestro entorno, y lo más probable es que nuestro resultado sea peor. Esta debilidad en la validez externa de un metaanálisis es extensible a los ensayos clínicos, y no es infrecuente que se pase por alto cuando se leen artículos de este tipo.

En definitiva y como ya se ha mencionado antes, un metaanálisis no equivale sin más discusión a «controversia resuelta» por más que ocupe el primer puesto en la escala de la evidencia, y siempre es necesario aplicarle una mirada crítica, al menos en los aspectos fundamentales que hemos tratado.

Metaanálisis de supervivencia

Cuando se analizan las diferencias en supervivencia de dos intervenciones se ha recurrido tradicionalmente a metaanálisis basados en Odds Ratios. Se debe tener en cuenta que esta medida relativa del riesgo sólo cuenta números de eventos (muertes, recidivas) pero no tiene en conside- ración el mayor o menor tiempo que tardan en producirse ni la influencia de los casos censurados, aquellos casos per- didos en el seguimiento y los que no han sufrido el evento final. No tienen la misma supervivencia tres eventos apa- recidos al año de seguimiento que otros tres eventos que acontecen a los cinco años. Por lo tanto, lo que hace este tipo de metaanálisis es sólo una aproximación a la verdadera supervivencia.

Los Hazard Ratios constituyen la medida relativa del riesgo de sufrir un evento que sí tienen en cuenta los tres aspectos de la supervivencia: número de eventos, tiempo e influencia de los casos censurados. Por lo tanto, un Hazard Ratio es la medida relativa del riesgo que se debería utilizar

para un auténtico metaanálisis de la supervivencia. Si no se han utilizado anteriormente, o actualmente con más frecuencia, es por la dificultad del cálculo estadístico y por la ausencia de software que lo permitiera. Actualmente este problema ya está resuelto.

La condición necesaria para emplearlos es que los ensa- yos clínicos que se incluyan en el metaanálisis reporten el valor del Hazard Ratio y el valor de los dos límites de su intervalo de confianza, o en última instancia, la informa- ción necesaria para calcularlos desde otros datos¹⁶. Desde estas tres cifras se pueden obtener, mediante las oportunas transformaciones matemáticas¹⁶⁻¹⁸, las restantes variables necesarias a incluir en la tabla de datos de los programas estadísticos, como pueden ser Stata[®] o MedCalc[®], entre otros muchos. Básicamente son tres las variables a manejar: el propio Hazard Ratio (HR) tal cual o en escala logarít- mica (logHR), el error estándar del Hazard Ratio en escala logarítmica (EElogHR), o el llamado estadístico V que es la varianza del Hazard Ratio en escala logarítmica. Los prin- cipales métodos de cálculo existentes utilizan LogHR con EElogHR, o bien HR con V. El resultado final (diamante) mos- trará el HR combinado con su intervalo de confianza.

Metaanálisis de pruebas diagnósticas

Es otra variedad de metaanálisis de datos agregados. La introducción de datos de los estudios individuales deriva de una tabla 2x2 que confronta una prueba diagnóstica de referencia (*Gold Standard*) con la prueba a combinar y resumir en el metaanálisis. Hay que disponer pues de cuatro variables imprescindibles: el número de verdaderos y falsos positivos y negativos reportados en cada estudio. El resultado final mostrará la Sensibilidad, Especificidad, y Cocientes de probabilidad combinados de la prueba diag- nóstica a estudio, así como sus índices de rendimiento diagnóstico general, Odds Ratio y Área bajo la curva ROC. En líneas generales, la valoración de la validez de este tipo de metaanálisis es similar a lo expuesto antes, aun- que tiene algunas características particulares, como es la evaluación del efecto del distinto umbral de positividad utilizado en la prueba diagnóstica¹⁹ analizada, que reper- cute en la heterogeneidad del metaanálisis. Actualmente se utilizan complejas rutinas de cálculo multivariantes que todavía están en discusión y en evolución. No obstante, para los iniciados puede ser ejecutado con relativa facilidad en distintos paquetes de software estadístico^{19,20}.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Glass GV. Primary Secondary, and Meta-Analysis of Research. *IJERI*. 1976;5:3-8.
2. Glass GV. Meta-analysis ad middle age: a personal history. *Res Synth Methods*. 2014, <http://dx.doi.org/10.1002/jrsm.1133>.
3. Egger M, Smith GD, Altman DG. Systematic reviews in health care. *Meta-analysis in context*. Ed London: BMJ Books; 2001.

4. Walker E, Hernandez AV, Kattan MW. Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*. 2008;75:431–9.
5. Urrutia G, Tort S, Bonfill X. Metaanálisis (QUOROM). *Med Clin (Barc)*. 2005;125:32–7.
6. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet*. 1999;354:1896–900.
7. Urrutia G, Bonfill X. Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Med Clin (Barc)*. 2010;135:507–11.
8. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol*. 2000;53:207–16.
9. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statis Med*. 2002;21:1539–58.
10. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple graphical test. *BMJ*. 1997;315:629–34.
11. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *J Clin Epidemiol*. 2001;54:1046–55.
12. Sterne JAC, Gavaghan D, Egger M. Publication and related bias in meta-analysis: Power of statistical test and prevalence in the literature. *J Clin Epidemiol*. 2000;53:1119–29.
13. Gleser LJ, Olkin I. Models for estimating the number of unpublished studies. *Stat Med*. 1996;15:2493–507.
14. Rosenthal R. The «file drawer problem» and tolerance for null results. *Psychol Bul*. 1979;86:638–41.
15. Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C. Evaluating Meta-analyses in the General Surgical Literature. A Critical Appraisal. *Ann Surg*. 2005;241:450–9.
16. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analysis of the published literature for survival endpoints. *Statis Med*. 1998;17:2815–34.
17. Tierney JF, Stewart L, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials*. 2007;8:16–22.
18. Primo J, Escrig J. MetaSurv: calculadora Excel para metaanálisis de supervivencia. 2008. <http://www.redcaspe.org/herramientas/calculadoras>. [Consultada 12.4.2020].
19. Zamora J, Plana MN, Abaira V. Estudios de evaluación de la validez de una prueba diagnóstica: revisión sistemática y metaanálisis. *Nefrología*. 2009;29:15–20.
20. Wang J, Leeflang M. Recommended software/packages for meta-analysis of diagnostic accuracy. *J Lab Precis Med*. 2019;4:22–7, <http://dx.doi.org/10.21037/jlpm.06.01>.