
Megaensayos o metaanálisis: modas o necesidades

Víctor Moreno Aguado

Servei d'Epidemiologia i Registre del Càncer. Institut Català d'Oncologia
Laboratori de Bioestadística i Epidemiologia. Facultat de Medicina
Universitat Autònoma de Barcelona.

Ensayo clínico y evidencia científica

La práctica médica cambia orientada por una serie de estímulos entre los que destacan los resultados obtenidos en ensayos clínicos controlados. Desgraciadamente, un gran número de ensayos clínicos, aun bien realizados desde el punto de vista metodológico, no permiten obtener conclusiones válidas debido al escaso número de pacientes incluidos. El tamaño de la muestra es uno de los aspectos importantes en el diseño de un ensayo. Este tamaño depende fundamentalmente, cuando la variable de interés es dicotómica, de la tasa de eventos en el grupo de referencia y de la magnitud del efecto a detectar. En muchas afecciones frecuentes estos parámetros son pequeños, por lo que el número necesario es enorme. Por ejemplo, la mortalidad aguda por infarto de miocardio, que está alrededor de un 10%, precisa un ensayo con más de 4.000 pacientes para demostrar una reducción relativa de un 25%. La decisión de realizar un ensayo de esta envergadura es difícil, por lo que muchas veces se recurre a estudios de menor tamaño, que emplean variables subrogadas y son más asequibles.

En otras patologías el tamaño de muestra necesario puede ser menor, del orden de varios centenares de pacientes. El problema es que, aunque un nuevo tratamiento demuestre eficacia en un ensayo de tamaño medio ($300 < n < 1.000$), la aceptación por la comunidad científica de que el tratamiento debe ser implantado en la práctica habitual exige la replicación de los resultados en diferentes grupos de enfermos y que un número importante de pacientes haya sido incluido en ensayos controlados. Esta acumulación de evidencia puede obtenerse por dos procedimientos complementarios: los megaensayos y los metaanálisis.

Megaensayos

Un megaensayo es un ensayo clínico controlado, que incluye varios miles de pacientes (en general más de 10.000) y está diseñado con el objetivo de medir el efecto de un tratamiento sobre un resultado clínico relevante, como la mortalidad asociada a una enfermedad¹. En la actualidad, el beneficio añadido que generalmente puede aportar un nuevo tratamiento es de pequeña magnitud, pero si se trata de una enfermedad de elevada incidencia, el número de personas susceptibles de beneficiarse puede ser muy grande. Este aspecto hace necesario el diseño de ensayos clínicos capaces de medir efectos pequeños sin problemas de sesgos ni errores aleatorios. Los sesgos sólo pueden evitarse mediante una adecuada aleatorización y un adecuado análisis estadístico. Para evitar los errores aleatorios es necesario un tamaño de muestra suficiente².

Estos grandes ensayos, para facilitar la inclusión del mayor número posible de pacientes, suelen emplear protocolos muy sencillos, que apenas se apartan de la práctica general empleada en el tratamiento de la patología de interés. Su análisis estadístico también suele ser muy sencillo, pues las hipótesis a contrastar son pocas y bien definidas. El análisis de subgrupos o variables secundarias muchas veces se realiza sólo de forma descriptiva, sin pruebas de significación estadística, para resaltar la importancia de que la evidencia sobre el papel beneficioso de un tratamiento sólo puede establecerse mediante ensayos diseñados específicamente para cada objetivo.

Los megaensayos, sin embargo, no pueden realizarse para probar la eficacia de cada tratamiento candidato. Son muy costosos en recursos humanos y económicos, y necesariamente son lentos por la necesidad de incluir, de manera coordinada, a varios miles de pa-

cientes. Por otra parte, una vez un fármaco demuestra eficacia en una indicación, subsiguientes ensayos clínicos en la misma indicación no pueden realizarse en comparación con placebo, sino con el tratamiento que demostró eficacia. En este caso, las diferencias de efecto que cabe esperar son mucho menores y ello supone la necesidad de emplear muchísimos más pacientes. Por ejemplo, es difícil que se pueda repetir megaensayos con análogos de fármacos que hayan demostrado eficacia a no ser que se anticipe claramente la posibilidad de que el análogo pueda superar las ventajas del primero³. Para resolver este aspecto a menudo se emplean objetivos subrogados o diseños para demostrar la equivalencia terapéutica entre fármacos de la misma clase.

Con frecuencia, por falta de recursos en general, los ensayos clínicos realizados se basan en un tamaño de muestra relativamente pequeño. Si bien este aspecto no impide que los resultados obtenidos sean significativos, en caso de que el efecto del tratamiento de interés sea claramente superior al de referencia, la posibilidad de generalizar los resultados es limitada, pues la población incluida en el estudio probablemente no sea representativa del conjunto de la población susceptible de ser tratada. Con mayor frecuencia, un ensayo clínico de tamaño pequeño resulta en un estimador del efecto de una magnitud no despreciable, pero carente de significación estadística. Esta situación, que en términos estadísticos se denomina falta de poder estadístico, deriva de un diseño incorrecto, bien por ignorar la necesidad de calcular un tamaño de muestra adecuado, o por sobrestimar *a priori* la diferencia de interés entre tratamientos.

Metaanálisis

Un metaanálisis es una revisión sistemática de un conjunto de ensayos* que emplea técnicas estadísticas cuantitativas para combinar y resumir los resultados. Su objetivo principal es aportar evidencia sobre la eficacia del tratamiento ensayado en la patología de interés. Otros objetivos adicionales son: aumentar el poder estadístico y la precisión de los estimadores del efecto, resolver posibles resultados

aparentemente contradictorios entre los ensayos, explicar la variación en los tamaños de los efectos, explorar el efecto en diferentes grupos, orientar sobre nuevas preguntas que no se habían planteado en los estudios originales por falta de poder estadístico y generar hipótesis⁴.

Los metaanálisis son la mejor herramienta para resumir los resultados de un conjunto de ensayos ya realizados, pues intenta sistematizar el proceso y evitar la subjetividad. Son baratos, pues se basan en estudios ya finalizados y no precisan grandes recursos a excepción de unas buenas fuentes de documentación para localizar los estudios y el esfuerzo de los investigadores para seguir una metodología rigurosa⁵.

En ausencia de megaensayos, los resultados de un metaanálisis constituyen la mejor fuente de información para tomar decisiones clínicas y para el diseño de futuros estudios en aspectos como qué nuevas hipótesis deben evaluarse y qué tamaño de muestra es necesario. Se ha propuesto que antes de la realización de un nuevo ensayo clínico debería realizarse un metaanálisis a partir de la revisión bibliográfica de los estudios similares ya realizados. Si la evidencia acumulada es suficiente, la realización de otro ensayo más puede no estar justificada a no ser que se pretenda confirmar el resultado. Si existen dudas o hay controversia, el metaanálisis puede ayudar a identificar los aspectos problemáticos y proporcionar ideas para mejorar el diseño del nuevo ensayo.

Con la intención de acelerar el desarrollo de terapias eficaces, recientemente se ha propuesto la realización sistemática de metaanálisis que actualicen sus resultados cada vez que un ensayo finaliza y aporta datos nuevos. A este método se le llama metaanálisis acumulativo^{6,7}. Se ha reconocido que en algunas situaciones los metaanálisis habían concluido el efecto beneficioso de una terapia mucho antes de que un megaensayo lo demostrara claramente. En estas situaciones se plantea el dilema ético de si es necesario exponer a un número importante de pacientes a un placebo o a un tratamiento inferior. Se ha discutido mucho de cómo los metaanálisis acumulativos pueden modificar el desarrollo de los megaensayos. Chalmers sugiere que los resultados de un metaanálisis previo al diseño de un nuevo ensayo deben afectar al cálculo del tamaño de muestra necesario, subrayando la idea de que el nuevo ensayo será uno más en aportar información al metaanálisis acumulativo. Así es

*La técnica del metaanálisis también puede aplicarse para combinar estudios observacionales (casos y controles o cohortes). Los aspectos metodológicos en este tipo de estudios son más complicados y no se comentan en esta revisión.

posible que el número de pacientes necesario sea relativamente reducido si se tiene en cuenta la información previa acumulada.

Si se aceptan los resultados de un metaanálisis acumulativo como concluyentes, es posible que grandes ensayos en marcha deban ser detenidos prematuramente, con los problemas logísticos y estadísticos que ello comporta. Se ha solicitado que el diseño de los megaensayos tenga en cuenta la posibilidad de que la monitorización interina de los resultados pueda verse modificada por estudios externos, en concreto por metaanálisis^{8,9}. Otros investigadores, ante los problemas de validez que se presentan en la interpretación de los metaanálisis, prefieren basar la evidencia en los resultados de los megaensayos e impedir que se introduzcan modificaciones en los protocolos por este concepto¹⁰.

El mensaje de que el metaanálisis es la panacea para generar evidencia científica es peligroso. Puede hacer pensar que no es importante ni el tamaño ni la calidad de un ensayo clínico, pues una vez finalizado siempre puede incorporarse a un metaanálisis. Claramente, un metaanálisis no debe consistir en un intento de conseguir un resultado significativo a partir de sumar muchos estudios no significativos con efectos irrelevantes. Por otra parte, incluso cuando los estudios a combinar son ensayos clínicos controlados, las características de los metaanálisis son más cercanas a los estudios observacionales que a los experimentales¹¹ y la interpretación de los resultados debe tener en cuenta una serie de aspectos metodológicos que se revisarán a continuación.

Selección de estudios a incluir

Este aspecto está relacionado con el objetivo del metaanálisis. En función de la pregunta específica que se quiera contestar, se deberán seleccionar unos ensayos u otros. Un buen metaanálisis debe basarse en un protocolo que defina claramente los objetivos y la metodología a seguir¹². Se deben especificar las fuentes de información para localizar metaanálisis y los criterios de inclusión y exclusión. Un aspecto muy controvertido es qué hacer con los ensayos de dudosa calidad. Hay quien prefiere incluir todos los estudios, independientemente de su calidad, y quien exige unos mínimos requerimientos para su inclusión. En general, la calidad de un estudio sólo puede juzgarse parcialmente a partir de los aspectos mencionados en la publicación, aunque se

han diseñado criterios estandarizados^{14,15}. Más adelante, en el apartado de heterogeneidad se comentarán otros aspectos relacionados con la calidad de los ensayos.

Sesgo de publicación

Idealmente un metaanálisis debería incluir todos los ensayos que cumplan los criterios, tanto los publicados como los no publicados. Es bien conocido el fenómeno por el cual los estudios que obtienen un resultado negativo, es decir, no alcanzan el nivel de significación estadística necesario, tienen una menor probabilidad de ser publicados¹⁶. En este fenómeno intervienen tanto los investigadores, que pierden el interés por publicar el estudio, como los promotores, que pueden no estar interesados en la publicación de un resultado que les es desfavorable, y los editores de las revistas científicas, que amparados en la "falta de interés" de un resultado negativo, o en aspectos metodológicos, pues un resultado negativo con frecuencia se debe a una falta de poder estadístico, rechazan la publicación. La consecuencia del sesgo de publicación es que los metaanálisis tienden a sobrestimar el efecto real del tratamiento. La identificación de este problema no es sencilla. Los estudios de gran tamaño suelen publicarse todos, sean positivos o negativos. Son los estudios de pequeño tamaño los que tienden a caer en el olvido. Un análisis gráfico de tipo "funnel-plot" que representa el efecto observado en los estudios frente a la precisión del estimador permite sospechar la existencia de sesgo de publicación¹⁷. También se han derivado fórmulas para estimar el número de ensayos de pequeño tamaño con resultado negativo que deberían existir para revertir el efecto observado¹⁸. Si este número no es muy grande, la validez del metaanálisis queda en entredicho. Para evitar el sesgo de publicación se ha propuesto la creación de registros internacionales de ensayos¹⁹.

Sesgo de idioma o "Torre de Babel"

Recientemente se ha resaltado la importancia que tiene el resultado de un ensayo para elegir el idioma de su publicación en países de lengua no inglesa²⁰. Tras estudiar las publicaciones de una serie de investigadores alemanes, se pudo comprobar cómo los estudios con resultados positivos con mayor frecuencia se publicaban en revistas de lengua inglesa, mientras que los estudios negativos aparecían

más a menudo en revistas alemanas. Es frecuente que metaanálisis realizados en países de lengua inglesa excluyan las publicaciones en otras lenguas. Este fenómeno genera un efecto similar al sesgo de publicación^{21,22}.

Independencia de los ensayos

Los métodos estadísticos empleados para combinar ensayos suponen independencia entre los mismos. A menudo los datos de un mismo ensayo se publican varias veces, en general aspectos parciales de los mismos, otras veces hay claras duplicaciones. Como los artículos suelen aparecer en revistas diferentes y a veces incluso con autores diferentes, no es difícil incluir el mismo ensayo más de una vez en un metaanálisis²³.

Heterogeneidad entre estudios

Los efectos observados en los diferentes ensayos a combinar pueden resultar heterogéneos. Es decir, es posible que el efecto sea claramente a favor del nuevo tratamiento en un subconjunto de los ensayos, pero que otros ensayos muestren un efecto nulo o incluso a favor del tratamiento estándar. En estas situaciones, más frecuentes de lo que sería deseable, se cuestiona la legitimidad de combinar efectos que apoyan hipótesis contrapuestas. Para detectar la existencia de heterogeneidad pueden realizarse tests estadísticos y es muy útil la gráfica de Galbraith²⁴, que dibuja cada ensayo con un punto con coordenadas basadas en el inverso de la precisión y el efecto estandarizado. En esta gráfica se pueden dibujar unas bandas de confianza, de manera que los ensayos que quedan por fuera de ellas son responsables de la heterogeneidad. En caso de detectar heterogeneidad, debe investigarse la causa de la misma²⁵, que suele estar relacionada con diferencias de diseño y calidad de los ensayos (tabla I). De hecho, los resultados de un metaanálisis sólo pueden ser interpretados si la heterogeneidad se debe a diferencias en la metodología de los ensayos²⁶.

Algunos autores han intentado cuantificar de qué factores depende la magnitud del efecto estimado²⁷. Sus conclusiones son que los defectos metodológicos, especialmente los relacionados con la asignación aleatoria de los pacientes, son la causa principal de sesgo en la estimación del efecto.

Para resolver el problema de la heterogeneidad se han propuesto varias aproximaciones.

TABLA I
ASPECTOS DEL DISEÑO DEL ENSAYO QUE PUEDEN CONDUCIR A HETEROGENEIDAD

Definición de la enfermedad. Criterios diagnósticos
Características del tratamiento nuevo y del tratamiento control (placebo o activo)
Definición del resultado o de la medida de efecto
Población tratada: edad, sexo, raza
Duración del estudio
Nivel de enmascaramiento: abierto, ciego simple, doble ciego
Tasa de abandonos y exclusiones no justificadas
Tipo de randomización y efectividad de la misma
Tipo de análisis: explicativo o por intención de tratar

Pueden excluirse los ensayos de dudosa calidad que contribuyan de manera importante a la heterogeneidad del efecto. También se ha sugerido emplear un factor de ponderación que disminuya el peso de los ensayos de peor calidad en su contribución al resumen final del efecto¹⁵. Sin embargo, estas ponderaciones son subjetivas, por lo que no suelen ser bien aceptadas²⁸. Si la causa de la heterogeneidad no puede ser explicada, o se debe a aspectos del diseño que no se basan en la calidad y el investigador desea emplear todos los estudios, se ha propuesto el uso de modelos estadísticos con efectos aleatorios para generar un estimador de efecto combinado que contemple la heterogeneidad^{29,30}. Estos modelos suponen que la magnitud del efecto no es constante en todas las situaciones, sino que factores no medidos modifican el efecto y hacen que los investigadores observen una importante variabilidad. Algunos investigadores piensan que el uso de modelos con efectos aleatorios difícilmente está justificado. Peto y cols. opinan que estos modelos suponen que la muestra de ensayos incluida en el metaanálisis es “representativa” del conjunto de posibles estudios existentes². Ya que en general los estudios incluidos son todos los existentes, esta asunción no está justificada y es preferible su método de análisis que él llama “libre de asunciones” más que de efectos fijos. Greenland, a su vez, considera que estos modelos sólo pueden ser el último recurso tras investigar la causa de heterogeneidad y no poder ser explicada²⁸.

Análisis de subgrupos

Un metaanálisis puede enfocarse para analizar resultados que no fueron el objetivo principal de los ensayos incluidos. También puede ser que el interés se centre en el estudio del efecto en subgrupos de población con características especiales como ancianos, fumadores, etc. La validez de las conclusiones del metaanálisis en estas situaciones es dudosa. La aleatorización aplicada en cada estudio no puede garantizar la homogeneidad de los pacientes comparados respecto a otros posibles factores confusores, y el gran tamaño de muestra global del metaanálisis no resuelve este problema. Un análisis combinado de los estudios empleando datos individualizados con información sobre otras variables pronósticas puede permitir un ajuste de estos posibles confusores, pero el interés de la aleatorización se pierde en muchos de estos casos y el estudio adquiere características observacionales más que experimentales, con las complejidades añadidas que ello supone.

Obviamente, el tamaño de un metaanálisis en cuanto al número de ensayos y pacientes que incorpora es un factor importante de cara a valorar la validez del mismo. No es lo mismo un metaanálisis que combina cuatro ensayos pequeños con varios cientos de pacientes que un metaanálisis que combina más de cien ensayos con varias decenas de miles de pacientes. Un tamaño grande orienta a que se ha realizado un gran esfuerzo para evitar el sesgo de publicación. En el mejor de los casos se puede solicitar a los investigadores de los ensayos los datos originales y crear un grupo de reanálisis combinado. Los ejemplos que existen de estas iniciativas han proporcionado resultados de validez similar a la de los megaensayos^{2,31}.

Discordancias entre los resultados de metaanálisis y megaensayos

El metaanálisis, a pesar de los problemas mencionados, ha alcanzado gran prestigio en los últimos tiempos gracias al esfuerzo de algunos grupos por normalizar su metodología y generar informes periódicos con la evidencia científica acumulada en diversos campos de la terapéutica³². Recientemente, sin embargo, existen ejemplos de conclusiones discordantes entre los resultados de metaanálisis y megaensayos. Tal vez el más notorio ha sido el resultado del ensayo ISIS-4 respecto al papel del

magnesio en el tratamiento agudo del infarto de miocardio. Este megaensayo, con 58.050 pacientes, proporcionó un resultado negativo cuando un metaanálisis previo había estimado un gran efecto, que había sido confirmado por otro ensayo relativamente grande, el LIMIT-2 con 2.316 pacientes. Ha habido gran discusión en la literatura para intentar explicar las discordancias entre estos resultados. Los creyentes en el efecto beneficioso del magnesio atribuyen la falta de eficacia en el ISIS-4 a un problema con el tiempo de administración, pero es difícil pensar que si realmente existe este efecto no haya sido apreciado por un ensayo de tal magnitud. Existen otros ejemplos similares en la literatura³³, lo que ha generado interés en explorar de qué factores dependen las discordancias y evaluar la capacidad predictiva de los metaanálisis respecto al efecto demostrado por los megaensayos. En general, se consideran los resultados de los megaensayos como el patrón de referencia con validez absoluta, y cualquier discrepancia entre éstos y los metaanálisis se achaca a deficiencias metodológicas de éstos últimos, en especial al sesgo de publicación¹⁷.

Villar y cols. han comparado el resultado de 30 metaanálisis realizados en diversas áreas terapéuticas en medicina perinatal, tras excluir el ensayo de mayor tamaño ($n > 1.000$) con el resultado de éste último³⁴. Los resultados mostraron un buen acuerdo general en cuanto a la dirección del efecto (80%), pero sólo un 60% coincidieron en dirección y significación estadística. El valor predictivo positivo, es decir, la probabilidad de que un metaanálisis positivo sea confirmado por un megaensayo fue del 50 al 67% según la magnitud del efecto observado. Los autores notaron que la mayoría de los metaanálisis proporcionaron un estimador del efecto mayor que los megaensayos de referencia. Este fenómeno puede explicarse por una mayor probabilidad de un error de tipo I en los ensayos pequeños incluidos, por sesgos de publicación y por diferencias en el enfoque de los ensayos, que en general es más pragmático en los grandes y más explicativo en los pequeños³⁵.

Cappelleri y cols. reexaminaron de manera similar los resultados de 79 metaanálisis y compararon el resultado de la combinación de los estudios pequeños con la combinación de los grandes ($n > 1.000$)³⁶. En su análisis, en un 90% de las ocasiones no había diferencias estadísticamente significativas entre los estimadores del efecto del metaanálisis y de los en-

sayos grandes. Sin embargo, las discrepancias se duplicaban si se empleaba un modelo de efectos fijos para realizar la combinación de los estudios. Los autores encontraron explicaciones a 10 de los 15 ensayos con resultados discordantes. Los motivos más frecuentes fueron diferencias en la tasa de eventos en el grupo control entre los ensayos pequeños y los grandes y ligeras diferencias en el protocolo. En un metaanálisis se sospechó la existencia de sesgo de publicación.

Recientemente, LeLorier y cols. han comparado los resultados de megaensayos ($n > 1.000$) publicados en 4 revistas prestigiosas de medicina (NEJM, Lancet, Ann Int Med y JAMA), con metaanálisis sobre los mismos temas publicados previamente³⁷. Estos autores no han reanalizado los metaanálisis como hacían los anteriores, sino que compararon los resultados de diferentes publicaciones (metaanálisis y megaensayos) desde el punto de vista de significación estadística. Tras identificar 12 megaensayos y 19 metaanálisis con 40 resultados de interés, el acuerdo global fue moderado, con un valor predictivo positivo del 68% y un valor predictivo negativo del 67%. Los autores señalan que el dicotomizar el resultado en positivo-negativo según un criterio estadístico resalta las discordancias, pero es el criterio que suele emplearse para decidir si una nueva terapia debe incorporarse a la práctica rutinaria. Según sus resultados, si no se hubieran realizado megaensayos, en un 32% de las ocasiones se habría adoptado el uso de pautas terapéuticas ineficaces y se habrían desechado tratamientos eficaces en un 33% de las ocasiones. La revisión de motivos de discrepancia es similar a la de los autores anteriores.

En conclusión, metaanálisis y megaensayos son estudios complementarios. Los metaanálisis son herramientas muy útiles para sintetizar la información y permiten generar hipótesis, pero no son adecuados para contrastarlas con seguridad absoluta. Por sus características observacionales están expuestos a sesgos, en especial los de publicación, y los estudios de concordancia de sus resultados con los de los megaensayos no son todo lo tranquilizadores que sería deseable. En general, no deberían usarse para recomendar pautas de tratamiento en ausencia de ensayos de suficiente tamaño. Los megaensayos continúan siendo necesarios para establecer la evidencia de la superioridad de un tratamiento cuando el efecto es de pequeña magnitud. Sin embargo, los megaensayos no pueden resolver todas las preguntas,

como qué efecto tiene una dosis diferente de la ensayada, o un régimen diferente, o en pacientes que no cumplieran los criterios de inclusión. En estas situaciones, la información aportada por los metaanálisis bien realizados será la de mejor calidad disponible.

BIBLIOGRAFÍA

1. Yusuf S, Collins R, Peto R. Why do we need large, simple randomized trials? *Stat Med* 1984; 3: 409-420.
2. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *J Clin Epidemiol* 1995; 48: 23-40.
3. Lubsen J. Mega-trials: is meta-analysis an alternative?. *Eur J Clin Pharmacol* 1996; 49 (Suppl 1): S29-S33.
4. Victor N. "The challenge of meta-analysis": discussion. Indications and contra-indications for meta-analysis. *J Clin Epidemiol* 1995; 48: 5-8.
5. Guallar E, Damián J, Martín-Moreno JM. Metaanálisis y revisiones sistemáticas en cardiología. *Rev Esp Cardiol* 1997; 50: 345-354.
6. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327: 248-254.
7. Lau J, Schmid CH, Chalmers TC. Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *J Clin Epidemiol* 1995; 48: 45-57.
8. Chalmers TC, Lau J. Changes in clinical trials mandated by the advent of meta-analysis. *Stat Med* 1996; 15: 1263-1268.
9. Chalmers TC, Lau J. Meta-analytic stimulus for changes in clinical trials. *Stat Methods Med Res* 1993; 2: 161-172.
10. Pocock SJ. The role of external evidence in data monitoring of a clinical trial. *Stat Med* 1996; 15: 1285-1293.
11. Anello C, Fleiss JL. Exploratory or analytic meta-analysis: should we distinguish between them? *J Clin Epidemiol* 1995; 48: 109-116.
12. D'Agostino RB, Weintraub M. Meta-analysis: a method for synthesizing research. *Clin Pharmacol Ther* 1995; 58: 605-616.
13. Light RJ. Accumulating evidence from independent studies: what we can win and what we can lose. *Stat Med* 1987; 6: 221-231.
14. Chalmers TC, Smith HJR, Blackburn B, et al. A method for assessing the quality of a randomized control trial. *Controlled Clin Trials* 1981; 2: 31-49.
15. Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, Abbe KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol* 1992; 45: 255-265.
16. Felson DT. Bias in meta-analytic research. *J Clin Epidemiol* 1992; 45: 885-892.

17. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; 315: 629-634.
18. Gleser LJ, Olkin I. Models for estimating the number of unpublished studies. *Stat Med* 1996; 15: 2493-2507.
19. Meinert CL. Toward prospective registration of clinical trials. *Controlled Clin Trials* 1988; 9: 1-5.
20. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997; 350: 326-329.
21. Gregoire G, Derderian F, Le LJ. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? *J Clin Epidemiol* 1995; 48: 159-163.
22. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, LeLorier J, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996; 347: 363-366.
23. Tramèr MR, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *Br Med J* 1997; 315: 635-640.
24. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med* 1988; 7: 889-894.
25. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994; 309: 1351-1355.
26. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet* 1991; 338: 1127-1130.
27. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. *JAMA* 1995; 273: 408-412.
28. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994; 140: 290-296.
29. DerSimonian R, Laird NM. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986; 7: 177-188.
30. Stram DO. Meta-analysis of published data using a linear mixed-effects model. *Biometrics* 1996; 52: 536-544.
31. Sandercock P. Collaborative worldwide overviews of randomized trials. *Ann NY Acad Sci* 1993; 703: 149-154.
32. Chalmers I. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993; 703: 156-163.
33. Borzak S, Ridker PM. Discordance between meta-analyses and large-scale randomized, controlled trials. Examples from the management of acute myocardial infarction. *Ann Int Med* 1995; 123: 873-877.
34. Villar J, Carroli G, Belizan JM. Predictive ability of meta-analyses of randomised controlled trials. *Lancet* 1995; 345: 772-776.
35. Villar J, Piaggio G, Carroli G, Donner A. Factors affecting the comparability of meta-analyses and largest trials results in perinatology. *J Clin Epidemiol* 1997; 50: 997-1002.
36. Cappelleri JC, Ioannidis JP, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, et al. Large trials vs meta-analysis of smaller trials: how do their results compare? *JAMA* 1996; 276: 1332-1338.
37. LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997; 337: 536-542.

DISCUSIÓN

E. GUALLAR: En primer lugar convendría matizar que las limitaciones mencionadas de los metaanálisis, como el sesgo de publicación, los problemas en la calidad, la heterogeneidad, etc., constituyen también limitaciones en general para toda la literatura científica y no son exclusivas del metaanálisis. Cualquier ensayo clínico muy bien diseñado y perfectamente analizado que aparece publicado en una revista está afectado por el sesgo de publicación, porque cuando lo leemos no estamos viendo los otros artículos sobre este tema con resultados nulos y que no se publican. Tal vez una de las grandes ventajas del metaanálisis es que ha puesto de manifiesto la existencia de unas limitaciones que deben tenerse siempre en cuenta. En referencia con algunos comentarios sobre la aproximación bayesiana a algunos de estos proble-

mas, no puedo resistir la tentación de preguntarme sobre el hecho de que cuando nosotros disponemos de un metaanálisis no podemos ignorar que no tenemos nada de información. Por tanto, desde el punto de vista ético y desde el punto de vista terapéutico, ¿cómo debemos tratar esta información previa de que disponemos, y hasta qué punto podemos ignorar la evidencia que hay acumulada hasta el momento a la hora de diseñar o plantear un nuevo ensayo?

V. MORENO: Creo que, claramente, no se puede ignorar la información previa disponible. Cuando sobre el tema de las interacciones entre metaanálisis y megaensayos he dicho que no estaba de acuerdo en tres puntos y uno de ellos era precisamente éste, estaba pensando en la manera como se planteaban en la publicación de Chalmers en la *Statistics*

in *Medicine*^a. En concreto, no estoy de acuerdo en la conclusión de los autores de que no hace falta realizar más ensayos cuando el resultado de un metaanálisis es concluyente sobre la existencia o no de efecto. Los resultados contrapuestos entre metaanálisis y megaensayos avalan este punto. Por otro lado, en mi opinión, es lícito diseñar un nuevo ensayo clínico apurando el tamaño de muestra por falta de recursos, en base a aprovechar toda la información previa, pero siempre debe reconocerse que únicamente se aportará una información más y que nunca llevará a una conclusión definitiva sobre el tema.

G. URRÚTIA: En esta y otras presentaciones de hoy se ha comentado el tema del sesgo de publicación, es decir, la dificultad de publicar los estudios con resultados negativos, hecho atribuible tanto a los intereses obvios de la industria, como a la censura que ejercen los editores de las revistas y/o a la inflavaloración de la relevancia científica de dichos resultados negativos por parte de los propios investigadores. Como representante aquí de la Colaboración Cochrane me gustaría dar a conocer la existencia de un proyecto financiado por el Fondo de Investigación Clínica cuyo objetivo es la identificación y la descripción de la totalidad de ensayos clínicos realizados en España desde el año 1971, tanto si han sido publicados como si no. De forma retrospectiva, y prospectivamente a partir de ahora (como parte de la estrategia de la Colaboración Cochrane), dicha identificación y el posterior seguimiento de los estudios podrá aportar información valiosa que pueda ser incorporada a revisiones sistemáticas de la eficacia de las distintas intervenciones sanitarias.

R. PLANAS: Sobre el tema de los resultados negativos de los estudios y el sesgo que se genera por su falta de publicación me gustaría añadir algunos comentarios. En una reunión de consenso sobre hipertensión portal, se llegó al acuerdo de que para el caso de los ensayos clínicos sería recomendable exigir en el contrato la obligación de publicación de

los resultados generados por el estudio incluso cuando éstos fueran negativos. Ello evitaría problemas muy habituales que sufrimos los investigadores a la hora de difundir nuestros resultados y contribuiría de forma importante a la divulgación de datos que difícilmente pueden ser incluidos en estudios de revisión.

J.M. MIRÓ: Quisiera únicamente incluir muy brevemente algunos comentarios generales. En primer lugar, es cierto que los contratos con la industria farmacéutica limitan de manera muy importante la divulgación de los resultados de los ensayos clínicos que se realizan. Por otro lado, considero muy interesante la iniciativa de crear una base de datos con todos los ensayos clínicos que se realizan en nuestro país. Esta lista, que debería estar al alcance de los grupos de investigadores, podría ser una fuente de información muy interesante que, entre otras ventajas, evitaría la innecesaria repetición de algunos estudios. Por último, creo que no existe la infraestructura necesaria en nuestro país para poder realizar una monitorización activa de aquellos ensayos clínicos en los que el promotor es un individuo o una institución y no es la industria farmacéutica. En mi opinión y como una posibilidad más para mejorar la calidad de la investigación clínica en los centros hospitalarios, los Comités Éticos de Investigación Clínica deberían disponer de medios para desarrollar el correcto seguimiento de los ensayos clínicos.

X. CARNÉ: Recientemente y por parte de los editores de diversas publicaciones científicas, como Richard Horton de *The Lancet*^b, se ha anunciado lo que se denomina una "amnistía editorial" que promueve la divulgación de todos los ensayos clínicos que se conozcan, con resultados positivos o no, hayan o no sido publicados, ya sea en forma de resumen o de artículo original completo. Ello puede contribuir a la creación de bases de datos mucho más completas sobre los ensayos clínicos, lo que puede resultar de gran interés científico por los motivos que se han comentado anteriormente.

^aCorresponde a la referencia bibliográfica número 8 de este capítulo de la monografía.

^bHorton, R. Medical editors trial amnesty. *Lancet* 1997; 350:756 [Commentary]