



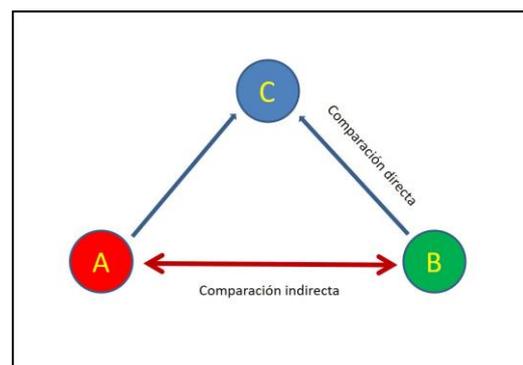
anónimo se inventó la propiedad transitiva, que dice más o menos así: si un número  $x$  se relaciona con  $y$ , e  $y$  se relaciona con  $z$ , habrá transitividad si  $x$  se relaciona con  $z$ . Otra vez a la anatomía: si mi pierna es mía y mi pie es de mi pierna, mi pie también es mío. Después vinieron más propiedades derivadas de estas tres, pero aquí lo vamos a dejar, porque hoy vamos a utilizar el poder de la propiedad transitiva para saber cuál de dos cosas que en realidad no hemos llegado a comparar es la mejor de las dos. Pensad, por ejemplo, en una turba enloquecida que entra corriendo en un centro comercial el primer día de rebajas. Lo miran todo antes de decidir qué comprar, pero no hace falta comparar todos los productos dos a dos para saber cuál nos gusta más.

En medicina pasa algo parecido. Lo habitual es que haya varias opciones para tratar una misma enfermedad (aunque los que ya llevamos uno años en el negocio sabemos que cuantas más existen, más probable que ninguna sirva para nada). Los ensayos clínicos, y los metanálisis de ensayos clínicos, solo comparan dos a dos y puede ocurrir que nadie haya comparado los dos que nosotros tenemos a nuestra disposición o que queramos saber cuál es, en teoría, el mejor de todos los disponibles.

Pues bien, para eso se ha inventado un diseño metodológico llamado metanálisis en red (MAR), también llamado metanálisis con comparaciones múltiples o metanálisis con comparaciones mixtas. Y en este último término, comparaciones mixtas, está el quid de la cuestión, porque resulta que hay varios tipos de comparaciones. Veámoslas.

Vamos a suponer que tenemos tres posibles tratamientos que, tras una honda reflexión, he decidido llamar A, B y C. La situación más sencilla es

comparar dos de ellos, A y B, por ejemplo, con un ensayo clínico convencional. Estaríamos haciendo una comparación directa entre las dos intervenciones. Pero puede ocurrir que no tengamos ningún ensayo que compare directamente A y B, pero sí dos ensayos diferentes que comparen las intervenciones con otra misma intervención, C (podéis verlo en la figura adjunta). En este caso podemos recurrir al poder de la propiedad transitiva y hacer una comparación indirecta entre A y B en función de su eficacia relativa frente a C. Por ejemplo, si A reduce la mortalidad un 100% frente a C y B la reduce un 50% frente a C, podremos decir que B reduce un 50% frente a A. Claro que, para poder hacer esto, tiene que cumplirse lo de la transitividad, cosa que no podemos dar por supuesta. Por ejemplo, si a mí me gusta el cerdo y al cerdo le gusta rebozarse por el barro, eso no quiere decir que a mí me guste rebozarme por el barro. La transitividad no se cumple en este supuesto (creo).



Pues bien, un MAR no es más que una serie de comparaciones directas, indirectas y mixtas que permiten comparar los efectos relativos de varias intervenciones. Lo habitual es representar las múltiples comparaciones con un diagrama en forma de red en el que podemos ver las comparaciones directas, indirectas y mixtas. Cada nodo de la red, que puede variar en tamaño en función de su peso específico, es uno de los estudios primarios de la revisión, mientras que las líneas que unen los

nodos representan las comparaciones. La red completa representará todas las comparaciones de tratamientos identificadas a partir de los estudios primarios de la revisión que incorpora nuestro MAR.

Al igual que ocurre con los otros tipos de metanálisis aparejados a una revisión sistemática, la validez del MAR dependerá de la validez de los estudios primarios, de la heterogeneidad entre los mismos y de los posibles sesgos de información existentes, factores que condicionarán la calidad de las comparaciones directas.

Además, las comparaciones indirectas se consideran de carácter observacional y requieren, como ya hemos comentado, que el investigador emita el dictamen de transitividad de las intervenciones basándose en sus conocimientos sobre las mismas, sobre la enfermedad y de los diseños de los estudios primarios.

Otro aspecto específico del MAR es el de la coherencia o consistencia, que hace referencia entre el nivel de acuerdo entre la evidencia procedente de comparaciones directas e indirectas. Este nivel de acuerdo, que puede medirse con métodos estadísticos específicos, debe ser alto para que la medida resumen de resultado tenga validez. Los resultados de las comparaciones deben tener la misma dirección, no pueden ser divergentes. Cuando esto no se cumpla, probablemente la causa esté en la mala calidad metodológica de los estudios primarios, en su heterogeneidad o en la presencia de sesgos.

Como en otros metanálisis, el resultado del MAR se expresa con una medida resumen de resultado que puede ser una odds ratio, una diferencia de medias, un riesgo relativo, etc. Esta estimación puntual se acompaña de un intervalo que nos da información sobre la

precisión de esta estimación. El análisis estadístico del MAR puede emplear métodos frecuentistas (el que solemos ver en los ensayos clínicos habituales) o métodos bayesianos. Estos últimos se basan en la asignación de una probabilidad del efecto del tratamiento previa al análisis de los datos para después asignar una probabilidad a posteriori tras el análisis. Para lo que nos interesa aquí, los métodos frecuentistas valorarán la precisión de la estimación puntual mediante los conocidos intervalos de confianza (habitualmente al 95%), mientras que los bayesianos proporcionarán sus intervalos de credibilidad (también al 95%), de significado similar.

Con todos estos datos obtendremos una relación ordenada de los tratamientos comparados, con el primero encabezando la lista. Pero no os confiéis demasiado, hay que mirar estos rangos con cuidado por varias razones. Primera, el mejor tratamiento en una situación puede no serlo en otra diferente. Segunda, hay que tener en cuenta otros factores como coste, disponibilidad, conocimiento del clínico, etc. Tercero, estas listas ordenadas no tienen en cuenta la magnitud de las diferencias entre los diferentes elementos. Y cuarta, el azar puede jugar nos malas pasadas y colocar en buena posición tratamiento que, en realidad, no sean tan buenos como pueda parecer.

Una vez vistas, muy por encima, las peculiaridades del MAR, ¿qué podemos decir de su lectura crítica? Al igual que disponemos de una lista de verificación para la revisión sistemática con metanálisis habitual, la declaración PRISMA, existe una declaración específica para el MAR, la PRISMA-NMA. Esta lista incluye, como ítems específicos, aspectos como la descripción de la geometría de la red de tratamientos, la consideración de los

supuestos de transitividad y consistencia y la descripción de los métodos utilizados para analizar la estructura de la red y la idoneidad de las comparaciones, por si alguna puede tener un grado de evidencia menor. Todo esto se verá facilitado si los autores proporcionan el gráfico con la red de estudios y nos explican brevemente sus características.

De todas formas, ya sabéis que a mí me gusta más acudir a las parrillas de lectura crítica de la red CASPe. Aunque no hay una específica, os aconsejo que uséis la de la revisión sistemática con metanálisis habitual y, después, hagáis unas consideraciones sobre los aspectos específicos del MAR.

Para no alargar mucho esta entrada, vamos a saltarnos toda la parte que comparten los MAR con cualquier otra revisión sistemática e ir directamente a los aspectos específicos. Podéis consultar la entrada correspondiente donde repasábamos la lectura crítica de la revisión sistemática. Como siempre, vamos a seguir nuestros tres pilares de la sabiduría: validez, importancia y aplicabilidad.

En cuanto a la **VALIDEZ**, nos haremos tres preguntas específicas.

**1. ¿Responde la revisión a una pregunta clínica bien definida y que justifique la realización de un MAR?** Esta pregunta tiene los clásicos componentes de la pregunta PICO, aunque la intervención y la comparación englobarán las múltiples comparaciones de la red.

**2. ¿Se realizó una búsqueda exhaustiva de los estudios relevantes?** Este aspecto es importante para evitar el sesgo de publicación y la inclusión de toda la información importante disponible. Su ausencia

puede afectar la consistencia de las comparaciones.

**3. Debe haber una especificación clara de la población diana, los tratamientos evaluados y las medidas de resultado empleadas.** Todos estos aspectos pueden condicionar la validez de las comparaciones indirectas. Si queremos inferir la relación entre los efectos de A y B comparando sus efectos individuales respecto a C, es esencial que A y B se traten de forma similar en su comparación con C, que las comparaciones A-C y B-C se hagan con pacientes similares, que se usen las mismas medidas de resultado y que el riesgo de sesgo de los estudios sea bajo. Esto último puede valorarse con las herramientas habituales, como la de la Cochrane.

Para acabar este apartado, comprobaremos que los resultados son analizados y presentados de forma adecuada, qué método estadístico se ha empleado (frecuentista o bayesiano), si se proporcionan los intervalos de confianza o de credibilidad, el análisis de la red, etc.

Aunque no vamos a entrar en ello, diremos que existen múltiples tipos de red (estrella, bucle, línea...). Para que las comparaciones tengan más validez, las indirectas deben estar apoyadas por comparaciones directas. Esto podemos verlo en el esquema de la red por la presencia de triángulos similares al del gráfico que mostraba al comienzo de la entrada (u otras formas geométricas cerradas). A igualdad de los otros factores que pueden influir y que ya hemos comentado, cuantos más triángulos veamos, más validez tendrán las comparaciones.

Como último aspecto, evaluaremos si los autores han empleado los métodos adecuados para valorar la heterogeneidad y la posible existencia

de inconsistencia: análisis de sensibilidad, metarregresión, etc.

Pasamos al apartado de **IMPORTANCIA**, en el que valoramos cuáles son los resultados del metanálisis. Aquí tendremos en cuenta cinco aspectos específicos:

**1. ¿Cuál es el resultado?** Como en cualquier otro metanálisis, valoraremos el resultado y su importancia desde el punto de vista clínico.

Habrá que valorar cómo se ha podido influir el resultado por el riesgo de sesgo de los estudios incluidos: a mayor riesgo de sesgo, más podrá apartarse nuestra estimación de la verdad.

**2. ¿Son precisos los resultados?** En este sentido, debemos valorar la amplitud de los intervalos de confianza o de credibilidad, teniendo en cuenta como se afectarían las conclusiones del estudio en cada extremo del intervalo.

**3. ¿Hay consistencia de los resultados entre los diferentes estudios?** Puede haber variabilidad por puro azar o por heterogeneidad entre los estudios. Podremos valorarla observando la forma de los *forest plots* y ayudándonos de los métodos estadísticos habituales, como la  $I^2$ .

**4. ¿Son fiables las comparaciones indirectas?** Volvemos nuevamente al concepto de transitividad, que habrá que tener en cuenta junto con los otros factores que hemos comentado previamente y que pueden aumentar el riesgo de sesgo: poblaciones homogéneas, variables de resultado y comparadores comunes, etc.

**5. ¿Hay consistencia entre comparaciones directas e indirectas?** Habrá que comprobar que existen formas geométricas cerradas dentro de la red (nuestros triángulos),

además de descartar causas de inconsistencia, que son las mismas que ya hemos comentado como causantes de heterogeneidad e intransitividad.

Finalmente, acabaremos nuestra lectura crítica haciendo algunas consideraciones especiales respecto a la **APLICABILIDAD** de los resultados.

Además de tener en cuenta, como es habitual, si se han considerado todos los efectos y variables importantes para el paciente y si los pacientes son similares a los de nuestro entorno, nos haremos alguna pregunta relacionada específicamente con el empleo del MAR, como si la red ha considerado todas las posibilidades de tratamiento o si los distintos subgrupos de comparación que se hayan podido establecer tienen credibilidad desde el punto de vista clínico.

Y aquí lo vamos a dejar por hoy. Una fiera difícil de domar, este MAR. Y eso que no hemos hablado nada de su metodología estadística, bastante compleja pero que los paquetes informáticos desarrollan sin inmutarse. Además, podríamos haber hablado largo sobre los tipos de redes y las comparaciones que se pueden desprender de cada una de ellas. Pero esa es otra historia...

## Bibliografía

1. Catalá-López F, Tobías A, Roqué M. Conceptos básicos del metaanálisis en red. Aten Primaria.2014;46:573-81. ([HTML](#))
2. Mills EJ, Ioannidis JPA, Thorlund K, Schünemann HJ, Puhan MA, Guyatt GH. How to use an article reporting a multiple treatment comparison meta-analysis. JAMA.2012;308:1246-53. ([PubMed](#)) ([HTML](#))
3. Al Khalifah R, Florez ID, Guyatt G, Thabane L. Network meta-analysis: users' guide for pediatricians. BMC Pediatr.2018;18:180. ([PubMed](#)) ([HTML](#)) ([PMC](#)) ([PDF](#))

4. Hutton B, Catalá-López F, Moher D. La extensión de la declaración PRISMA para revisiones sistemáticas que incorporan metaanálisis en red: PRISMA-NMA. Med lin (Barc):2016;147:262-6. ([PubMed](#)) ([HTML](#))

---

#### Correspondencia al autor

*Manuel Molina Arias*  
[mma1961@gmail.com](mailto:mma1961@gmail.com)  
*Servicio de Gastroenterología.*  
*Hospital Infantil Universitario La Paz.*  
*Madrid. España.*

---

Aceptado para el blog en mayo de 2019.

