

Ronda clínica y epidemiológica

Comparaciones indirectas

John Jairo Zuleta Tobón¹

INTRODUCCIÓN

Un oncólogo recibe en su consulta a una mujer con cáncer de mama avanzado. Él es consciente de que existen múltiples alternativas de tratamiento farmacológico y también de que, después de considerar las condiciones clínicas de la mujer y de conocer las necesidades, temores, preferencias y expectativas que ella tiene, su deber es ofrecer el tratamiento que, según el respaldo de la información científica más válida, ofrezca la mayor efectividad. Una vez su juicio clínico y su conocimiento le han permitido tener un panorama claro e integral de la situación, adecuándose a las circunstancias y el entorno en el cual está actuando, procede a hacer una búsqueda de la información disponible del tema.

La información más válida para tomar una decisión con respecto a una intervención médica proviene de los ensayos clínicos e, idealmente, de una revisión sistemática con metanálisis de varios ensayos clínicos de alta calidad. En una búsqueda inicial en la base de datos PubMed, se encuentran más de 230 ensayos clínicos que evaluaron el tratamiento de primera línea del cáncer de mama avanzado. Más abrumador aún, se encuentran más de 100 comparaciones que incluyen más de 20 esquemas terapéuticos diferentes. Se encuentran también metanálisis del tema. Sin embargo, similar a lo que sucede con los ensayos clínicos, estos metanálisis incluyen comparaciones de solo dos tratamientos ¿Cómo manejar tal cantidad de información? ¿Cómo decidir entre dos tratamientos que nunca se han comparado entre sí? Esta no es una situación exclusiva de las enfermedades complejas: para el tratamiento de la hipertensión se encuentran múltiples alternativas farmacológicas, con un fármaco o combinación de dos o más medicamentos, con gran cantidad de ensayos clínicos, donde igualmente se presenta la situación que algunas de ellas nunca se han evaluado frente a frente.

Los avances en la ciencia y en la tecnología llevan a que aparezcan nuevas alternativas terapéuticas y el clínico debe decidir entre múltiples intervenciones que compiten entre sí. Los entes regulatorios promueven la comparación de una nueva terapia farmacológica frente al placebo y no en todos los casos se cumple la recomendación ética de comparar contra la alternativa más efectiva existente al momento. La práctica de la medicina basada en la evidencia implica que las decisiones se fundamenten en información objetiva, y la comparación directa entre las diferentes intervenciones para una misma enfermedad es la excepción. Esta necesidad ha llevado a que se desarrollen estrategias metodológicas de comparación para solucionar ese vacío.

¹ Profesor de Ginecología y Obstetricia, Magíster en Epidemiología Clínica. Universidad de Antioquia, Medellín, Colombia.
Correspondencia: John Jairo Zuleta Tobón jjzuleta@une.net.co

Comparaciones indirectas

La situación más sencilla se da cuando cada uno de dos tratamientos (A y B) para una misma enfermedad se han comparado contra placebo de manera independiente en diferentes estudios. A partir de estos resultados, donde nunca se comparó A contra B en un mismo estudio, se debe decidir cuál alternativa es mejor. La existencia de más tratamientos vuelve más compleja la situación. Se han planteado diferentes alternativas para obtener la respuesta a esa pregunta.

Comparación narrativa

Método muy empleado previamente, pero sin buena validez. A partir de los resultados de varias comparaciones de tratamientos en ensayos clínicos, o de diferentes metanálisis, se hace una discusión y un juicio subjetivo, sin un cálculo estadístico formal del efecto relativo ni de su significación estadística.¹ Son comparaciones típicas de las revisiones narrativas de tema o de las revisiones sistemáticas, que por la gran diferencia entre los estudios no puede culminar en un metanálisis.

Comparación cruda o no ajustada

Se comparan directamente los resultados de uno de los brazos de tratamiento de cada uno de los estudios evaluados, sin tener en cuenta el grupo control de ninguno de los estudios. Esta es una comparación con riesgo de sesgos porque los estudios se pudieron haber realizado en poblaciones que no son comparables.¹ Sanchez-Ramos presentó un artículo en el cual evaluó la efectividad y la seguridad de dosis de 25 comparado con 50 µg de misoprostol para la inducción del trabajo de parto. Por un lado realizó un metanálisis tradicional con cinco ensayos clínicos que compararon frente a frente las dos concentraciones, y por otro tuvo en cuenta los resultados obtenidos en los grupos de misoprostol de dos estudios adicionales que compararon misoprostol y dinoprostone y dejó por fuera la información que aporta el grupo de dinoprostone.²

Revisiones "sombrija"

Es un término poco empleado en la literatura que corresponde a revisiones sistemáticas que consideran varias comparaciones de varios tratamientos para el manejo de la misma enfermedad o condición. La fuente

primaria para las revisiones sombrilla son revisiones sistemáticas previas de esos tratamientos y no los estudios primarios. En estos casos cada comparación se hace entre dos alternativas. Idealmente se debe planear de manera prospectiva, antes que reunir múltiples revisiones previamente implementadas, debido a las diferencias temporales, de criterios de inclusión-exclusión, de evaluación y de objetivos que pueden tener estas revisiones cuando no se planearon previamente para ser reunidas en una revisión de revisiones. En ellos no se hace una agrupación estadística de los resultados³.

Comparación indirecta ajustada

En ausencia de una comparación directa entre dos terapias, se puede emplear una comparación indirecta ajustada. El método permite la comparación entre dos medicamentos A y B, si ambos han sido comparados con una tercera alternativa C, que puede ser un medicamento activo o un placebo.^{4,5} Coomarasamy y colaboradores querían evaluar la efectividad de nifedipino (el uteroinhibidor actualmente aceptado como de primera elección) comparado con atosibán (el uteroinhibidor de más reciente introducción en el mercado). Uno de los objetivos fue evaluar la efectividad para prevenir el parto antes de 48 horas en mujeres con amenaza de parto prematuro. Como nifedipina y atosibán no se han comparado entre sí, pero ambos se han comparado en varios estudios contra betamiméticos, los autores realizaron el análisis indirecto. Por medio de dos metanálisis independientes obtuvieron los OR agrupados de las comparaciones entre nifedipina y betamiméticos (nueve ensayos clínicos, OR: 1,47 IC95% 0,9-2,42) y entre atosibán y betamimético (cuatro ensayos clínicos, OR: 1,27 IC95% 0,94-1,71) y, a partir de esos dos OR, realizaron una comparación ajustada de nifedipino vs atosibán (OR: 1,2 IC95% 0,73-1,95), con lo cual se concluyó que ambas alternativas son igualmente efectivas para prevenir el parto antes de 48 horas. Un hallazgo bien importante en este metanálisis indirecto fue que la nifedipina disminuyó la probabilidad de síndrome de dificultad respiratoria.⁶

Esta estrategia ofrece como ventaja que mantiene los beneficios reconocidos de la asignación por azar. Tres aspectos fundamentales para garantizar la validez en las comparaciones indirectas ajustadas son: que los ensayos clínicos que surten la comparación sean de alta

calidad, que sus resultados sean suficientemente (no es necesario que lo sean completamente) homogéneos y que los estudios de ambos brazos de comparación sean similares entre sí. El supuesto de similitud asume que los factores modificadores del efecto del tratamiento son similares en los dos estudios o grupos de estudios que se comparan. Esta similitud se debe entender en términos clínicos y metodológicos. La similitud clínica incluye las características clínicas y demográficas, los tratamientos, los escenarios, los desenlaces medidos, la duración del seguimiento, entre otros, y la similitud metodológica se refiere a todos los aspectos conocidos que pueden inducir sesgos tanto en el estudio primario como en la revisión sistemática, como son el método de asignación, el ocultamiento de las secuencias de asignación, el enmascaramiento, las pérdidas de seguimientos, los abandonos de tratamiento, el tipo de búsqueda de estudios, la forma de evaluar la calidad de los mismos y la publicación, entre otros. Algunos de los métodos empleados frecuentemente para evaluar o para mejorar la similitud entre los grupos son los análisis de subgrupos y la metarregresión.¹

La evaluación de este método ha encontrado que en 93% de los casos los resultados de las comparaciones indirectas son confirmados por ensayos clínicos de comparaciones frente a frente,⁴ e incluso un estudio encontró que en algunos casos los resultados obtenidos en las comparaciones indirectas son menos sesgados,⁷ aunque igualmente otros autores han cuestionado su validez.

Redes de tratamientos y metanálisis de múltiples tratamientos

Las comparaciones indirectas ajustadas tienen la desventaja que la comparación simultánea de más de dos intervenciones es bastante laboriosa y es inconveniente. Adicionalmente, no incluye dentro del resultado la información derivada de las comparaciones frente a frente que puedan existir de las intervenciones en evaluación. Una red de tratamientos reúne la información de varias comparaciones directas e indirectas y un metanálisis de esa red respeta el concepto de la asignación al azar e incrementa la precisión de los resultados al mezclar esos dos tipos de comparaciones.^{3,8} La figura n.º 1 representa una red relativamente sencilla donde tres tratamientos diferentes se han comparado

contra placebo, y los tratamientos B y C se han comparado frente a frente. A partir de esta red se puede evaluar la efectividad del tratamiento A comparativamente contra los tratamientos B o C y se puede definir cuál de ellos es el más efectivo, o se puede comparar B contra C incluyendo la información que pueda aportar la comparación entre A y el placebo. Cuantos más tratamientos y comparaciones entre ellos existan, más compleja se va volviendo esa red, como es el caso representado en la figura n.º 2. En las figuras n.º 1 y n.º 2 los círculos representan los tratamientos y las líneas las comparaciones entre dos de ellos, esta línea es más gruesa cuantos más estudios hayan evaluado tal comparación.

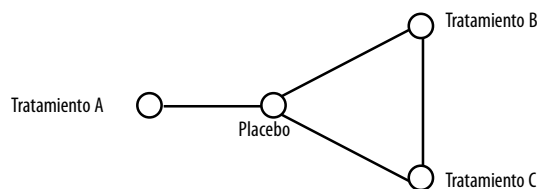


Figura n.º 1. Red de tratamientos

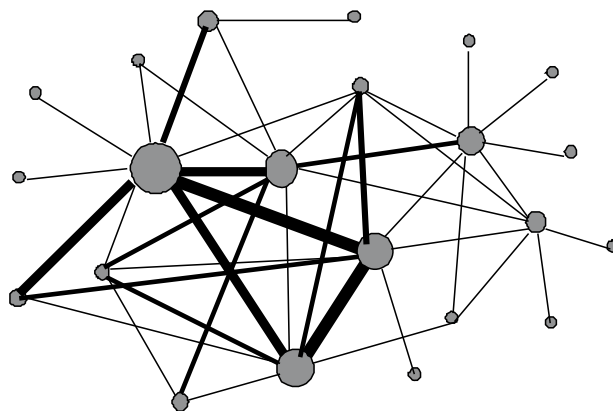


Figura n.º 2. Red de tratamientos de primera línea para pacientes con cáncer de mama avanzado

Una red de tratamiento se evalúa inicialmente en términos de su geometría: cuando todos los tratamientos se han comparado contra un control común pero no entre ellos, por ejemplo, todos contra placebo, se obtiene una red en forma de estrella. Si todas las intervenciones se han comparado entre sí, se obtiene un polígono. Entre estos dos extremos existen muchas alternativas. Formalmente esto se evalúa con la diversidad y la co-ocurrencia.⁹ La diversidad mide la probabilidad de que todos los tratamientos se hayan comparado entre sí y sirve para identificar si ha habido tratamientos más evaluados que otros. Esta diversidad se cuantifica con la Probabilidad de Encuentro Interespecífica, cuyo valor tiende a cero cuando la diversidad es limitada, y tiende a uno cuando es grande: valores de 0,75 o menores reflejan baja diversidad. La co-ocurrencia refleja si ha habido preferencia por, o se ha evitado, la comparación entre dos tratamientos, e igualmente se evalúa por intermedio de una probabilidad, Índice C de co-ocurrencia, para el cual un valor menor de 0,05 ó 0,1 indica que ha habido co-ocurrencia importante. En algunos casos esta co-ocurrencia es justificada, como cuando los tratamientos han estado disponibles por periodos de tiempo muy diferentes, pero en otras ocasiones es consecuencia de una política injustificada de no comparar un nuevo tratamiento contra la mejor alternativa existente. Al evaluar la geometría de la red se deben contestar dos preguntas: ¿Se han realizado las comparaciones correctas? ¿Existe una comparación para la cual la evidencia es poca o inexistente pero es esencial tenerla? y a partir de las respuestas se debe evaluar la presencia de potenciales sesgos en las comparaciones.

Los resultados de los ensayos incluidos en una red se pueden sintetizar con un metanálisis de múltiples tratamientos que incluyen la información de las comparaciones directas e indirectas. Igual que en las comparaciones indirectas ajustadas, los supuestos que se deben cumplir para realizar un metanálisis de red son la homogeneidad de resultados y la similitud, pero adicionalmente se debe respetar el de consistencia o coherencia.^{1,8} El supuesto de consistencia se refiere a que los resultados de las comparaciones directas e indirectas sean compatibles entre sí, por lo tanto dice si el efecto estimado con las comparaciones indirectas es diferente del estimado con las comparaciones directas. La situación más conflictiva se da cuando las dos comparaciones arrojan resultados contradictorios,

en cuyo caso es necesario evaluar cuidadosamente la calidad de los estudios que dan soporte a la comparación para definir si se deben combinar y sobre cuál de los dos resultados se debe aceptar como verdadero. Las razones potenciales para que no haya consistencia entre las comparaciones directas e indirectas incluyen el azar, diferencias en los pacientes incluidos, en las intervenciones, en la definición de las variables, de los antecedentes y del manejo, sesgos de publicación, reporte selectivo de desenlaces o de los análisis, terminaciones tempranas de los estudios o defectos metodológicos en los estudios primarios.^{1,3}

Las comparaciones se pueden hacer mediante métodos estadísticos frecuentistas tradicionales, o con métodos Bayesianos. Ambos métodos generan el efecto del tratamiento en comparaciones por parejas entre las múltiples alternativas, así como la incertidumbre o credibilidad de ese efecto (intervalos de confianza en los métodos frecuentistas e intervalos de credibilidad en los métodos Bayesianos). Una particularidad de los métodos Bayesianos es que adicionalmente permiten calcular la probabilidad de que cada uno de los tratamientos pueda ser el mejor.

El metanálisis de los múltiples tratamientos para cáncer de mama avanzado, representados esquemáticamente en la figura n.º 2, tuvo en cuenta 370 estudios elegibles con información de 54.189 mujeres asignadas al azar en un periodo de 35 años, 148 comparaciones diferentes con 22 tipos diferentes de tratamiento. De 231 comparaciones directas teóricamente posibles, sólo hubo 45.¹⁰ En general hubo una buena diversidad en las comparaciones (Probabilidad de Encuentro Interespecífica de 0,88, visualizado en la gráfica por la geometría de la red poligonal). Hubo claras preferencias por algunas comparaciones específicas (Índice C de co-ocurrencia 132 con $p < 0,001$, representado en la red por la diferencia de espesor en las líneas). La conclusión del estudio es que existen muchas opciones de tratamiento que se pueden usar como primera línea, porque tienen un perfil de efectividad similar.

En conclusión, la mejor evidencia acerca de una intervención se obtiene de la comparación frente a frente de dos tratamientos, e idealmente de la reunión de varios de esos estudios en un metanálisis. Sin embargo, el rápido avance tecnológico hace imposible tener una comparación directa entre todas las

alternativas posibles, lo que da cabida a las comparaciones indirectas. Estos métodos todavía están en desarrollo, y aunque falta evidencia empírica objetiva de su validez y del efecto de la falta de cumplimiento de los supuestos, es necesario familiarizarse con ellos, conocer sus ventajas, desventajas y utilidad, porque actualmente son la única alternativa para resolver esta situación.

REFERENCIAS BIBLIOGRÁFICAS

1. Song F, Loke Y, Walsh T, Glenny A, Esatwood A, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338:b1147. Disponible en http://www.bmj.com/cgi/reprint/338/apr03_1/b1147 Consultado el 25 de enero de 2010.
2. Sanchez-Ramos L, Kaunitz AM, Delke I. Labor induction with 25 microg versus 50 microg intravaginal misoprostol: a systematic review. *Obstet Gynecol* 2002;99(1):145-151.
3. Ioannidis J. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ*. 2009; 181(8): 488-493.
4. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003; 326:472-476. Disponible en <http://www.bmj.com/cgi/content/full/326/7387/472> Consultado el 25 de enero de 2010.
5. Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9: 1-134.
6. Coomarasamy A, Knox EM, Gee H, Song J, Khan K. Effectiveness of nifedipine versus atosiban for tocolysis in preterm labour: a meta-analysis with an indirect comparison of randomised trials. *BJOG*. 2003; 110: 1045-1049.
7. Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol*. 2008; 61:455-463.
8. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897-900.
9. Salanti G, Kavvoura FK, Ioannidis J. Exploring the Geometry of Treatment Networks *Ann Intern Med* April 1, 2008 148:544-553.
10. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis J. Multiple-Treatments Meta-analysis of Chemotherapy and Targeted Therapies in Advanced Breast Cancer. *J Natl Cancer Inst* 2008; 100: 1780-1791.

