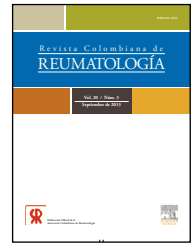


# Revista Colombiana de REUMATOLOGÍA

www.elsevier.es/rcreuma



## Artículo de revisión

# Comparaciones indirectas: una herramienta para la toma de decisiones en salud

Laura Marcela Rojas Díaz<sup>a,c</sup>, Héctor Mauricio Pérez Quevedo<sup>a,c</sup>,  
Paul Alejandro Mendez Patarroyo<sup>b,c</sup>, Paola Coral Alvarado<sup>b,c</sup>,  
Rafael Alfonso-Cristancho<sup>d</sup> y Gerardo Quintana<sup>a,b,c,\*</sup>

<sup>a</sup>Departamento de Medicina Interna, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia

<sup>b</sup>Departamento de Medicina Interna, Facultad de Medicina, Universidad de los Andes, Bogotá, Colombia

<sup>c</sup>Grupo Reumavance, Sección de Reumatología; Fundación Santa Fe de Bogotá, Bogotá, Colombia

<sup>d</sup>Departamento de Salud Pública y Farmacia, Facultad de Medicina, University of Washington, Washington, EE. UU.

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 7 de noviembre de 2012

Aceptado el 16 de agosto de 2013

### Palabras clave:

Evaluación de eficacia-efectividad  
de intervenciones

Comparaciones indirectas

Red de meta-análisis

Comparaciones de tratamientos  
mixtos

Meta-regresión

Métodos bayesianos

## R E S U M E N

Para el desarrollo de políticas sanitarias y económicas en el sector salud, es necesario que la toma de decisiones esté basada en la evidencia. El desarrollo de nuevas tecnologías para el cuidado de la salud se ha incrementado en las últimas décadas permitiendo la aparición, cada vez más rápida, de múltiples alternativas farmacológicas para el manejo de una misma patología; dada la cantidad de intervenciones se hace indispensable comparar las relevantes, en especial, mediante la realización de ensayos clínicos controlados, con el fin de evaluar y determinar la mejor alternativa en términos de beneficios, menores efectos adversos y costos. Sin embargo, debido a la ausencia de este tipo de estudios dado su alto costo, las comparaciones indirectas de tratamientos como las redes de meta-análisis y las comparaciones mixtas de tratamientos, son una herramienta útil para la selección de la mejor opción. Esta revisión, proporciona una orientación sobre la interpretación de las comparaciones indirectas, los métodos, supuestos, validez, métodos de análisis, beneficios y limitaciones, con el fin de proporcionar herramientas suficientes para la adecuada toma de decisiones en salud.

© 2012 Asociación Colombiana de Reumatología. Publicado por Elsevier España, S.L.

Todos los derechos reservados.

## Indirect comparisons: Tool for decision-making in health

### A B S T R A C T

For the development of health and economic policies in the health system, it is necessary for decision making is based on evidence. The development of new technologies for health care has increased in recent decades, allowing the appearance, fastest growing, multi-drug alternatives for handling the same pathology, given the number of interventions is

### Keywords:

Evaluation of the efficacy-  
effectiveness of interventions  
Indirect comparisons

\* Autor para correspondencia.

Correo electrónico: ge\_quintana@yahoo.com (G. Quintana).

Network meta-analysis  
Mixed treatment comparisons  
Meta-regression  
Bayesian methods

essential to compare all relevant specially through controlled clinical trials, to evaluate and determine the best alternative in terms of benefits, fewer side effects and costs. However, due to the absence of this type of study due to its high cost, indirect comparisons of treatments such as networks of meta-analysis and mixed treatment comparisons provide a useful tool for selecting the best treatment option. This review provides guidance on the interpretation of indirect comparisons, methods, assumptions, validity, analysis methods, benefits and limitations in order to provide sufficient tools for proper decision-making in health.

© 2012 Asociación Colombiana de Reumatología. Published by Elsevier España, S.L.

All rights reserved.

## Introducción

Los avances en las tecnologías farmacéuticas llevan a que las alternativas terapéuticas aumenten con el tiempo y los profesionales en salud deban decidir entre múltiples intervenciones que compiten entre sí. Sin embargo, tener un mayor número de alternativas terapéuticas crea una mayor necesidad de información relacionada con cada intervención<sup>1-3</sup>. En la práctica clínica, a pesar de los potenciales benéficos de tener un mayor número de alternativas, la selección de la mejor intervención terapéutica se puede dificultar debido a la falta de evidencia clínica disponible<sup>4-6</sup>. Por ejemplo, existen diversos medicamentos biológicos disponibles para el tratamiento de los pacientes con artritis reumatoide (AR) que no responden a medicamentos antirreumáticos modificadores de la enfermedad (DMARs, por sus siglas en inglés), principalmente el metotrexate, y aun así no existe suficiente evidencia clínica respecto a cuál sería el biológico más costo-efectivo.

La toma de decisiones en salud, idealmente debe considerar las condiciones clínicas de los pacientes, las necesidades, temores, preferencias y expectativas que estos tienen, con el fin de ofrecer el tratamiento más adecuado en cada caso, evaluando los riesgos y beneficios de cada alternativa, basándose en la evidencia clínica disponible y considerando también el costo-efectividad de cada intervención<sup>7-9</sup>. Por lo cual, la medicina basada en la evidencia (MBE) se ha consolidado como una herramienta fundamental para la toma de decisiones en salud, enfocada a comparar la efectividad entre intervenciones médicas<sup>10-13</sup>.

La información más válida para tomar una decisión con respecto a una intervención médica proviene de los ensayos clínicos controlados aleatorizados (ECC), también denominados estudios «cabeza a cabeza» y de las revisiones sistemáticas o meta-análisis de ECC de alta calidad<sup>14-16</sup>. Estas comparaciones directas (CD) se han establecido como el «Gold Standard» para evaluar la eficacia y seguridad de las intervenciones médicas y para hacer las evaluaciones económicas<sup>17-19</sup>. Para lograr ese objetivo, los ECC deben ser sólidos metodológicamente y ser capaces de detectar las diferencias mínimas clínicamente importantes y significativas entre las intervenciones evaluadas. Si un ECC se realiza con pobre calidad y múltiples sesgos, sus resultados pueden

ser confusos y originar falsas interpretaciones, así, una aleatorización inadecuada puede incrementar la razón de disparidad, también conocida como «Odds Ratios» (OR) entre las intervenciones hasta en 41%, y un cegamiento inadecuado hasta en 17%. Los OR se consideran como una medida de asociación, definida como la razón entre la probabilidad de que ocurra un evento y de que no ocurra en términos de eficacia, siendo una medida similar al riesgo relativo<sup>20</sup>. En conclusión, un bajo poder del ensayo se traduce en una disminución en la probabilidad de detectar las diferencias entre las intervenciones<sup>17</sup>.

Bajo el contexto de la toma de decisiones y el proceso de evaluación de las tecnologías sanitarias, los estudios de efectividad o costo-efectividad, y especialmente los meta-análisis de los ECC de este tipo, se caracterizan por la combinación cuantitativa de resultados de estudios similares con diseño y población comparables<sup>16</sup>. En múltiples ocasiones los nuevos medicamentos son comparados contra placebo o con medicamentos de referencia, pero generalmente no se reportan CD con más de dos intervenciones, o con dos o más medicamentos recientemente aprobados<sup>1,4</sup>. Los meta-análisis tradicionales, que usan ECC de CD no pueden discriminar el efecto independiente de las diferentes intervenciones, ni comparar las diferencias en el efecto entre ellas, por lo cual, usualmente, se limitan a evaluar un solo tipo de intervenciones o las agrupan dentro de la misma clase terapéutica para calcular el efecto promedio de esas intervenciones pero no por separado<sup>21,22</sup>.

El desarrollo de cualquier ECC requiere de recursos importantes en tiempo, número de pacientes y costo, lo cual hace que no se realicen de forma rutinaria para todas las comparaciones posibles o relevantes en la práctica clínica, por lo anterior las CD entre los medicamentos no están siempre disponibles<sup>2,5,7,12,17</sup>. Cuando estas CD son insuficientes, o no existen, puede ser posible usar los diferentes resultados de los ECC para estimar los efectos relativos de los diferentes tratamientos. En este sentido, las comparaciones indirectas (CI) surgen como una muy buena alternativa para hacer objetiva la toma de decisiones en salud cuando los ECC de un tratamiento no estén disponibles<sup>4-6,23</sup>. Las CI y las redes de meta-análisis (RMA), pueden proporcionar información útil para comparar los efectos del tratamiento entre las intervenciones de interés y de esta manera apoyar la toma de decisiones para escoger la mejor opción terapéutica<sup>1,24-26</sup>.

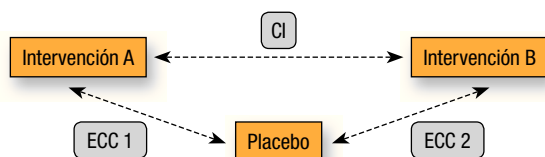
## Materiales y métodos

El presente manuscrito corresponde a una revisión narrativa no sistemática de literatura desarrollada en los últimos 10 años en idiomas inglés y español. Se realizó una búsqueda y revisión de literatura en las bases de datos PubMed, Embase y Scielo utilizando los términos y frases «indirect treatment comparisons», «bayes theorem» [Mesh], «Meta-analysis as topic» [Mesh], «decision making» [Mesh], «models statistical» [Mesh], «comparative effectiveness researchs» [Mesh], «network meta-analysis» «meta regression analysis» y «mixed treatment comparison». Se utilizaron los descriptores booleanos AND/OR para cada uno de los términos y tener una búsqueda más sensible, dado el escaso material publicado hasta la fecha. Adicionalmente, se realizó selección manual de los artículos considerados de interés sobre el tema y se amplió la información mediante la revisión de artículos referenciados en los documentos encontrados inicialmente y literatura gris. Tanto la búsqueda como la selección de los artículos se realizaron de forma independiente por los diferentes autores y cada uno realizó lectura, análisis, asociación de nuevas referencias y construcción literaria para posteriormente ser analizada, validada por el grupo de autores y corregida bajo el mejor estilo literario. Se realizó lectura de cada artículo y se seleccionaron los relacionados al tema en cuestión. Se incluyeron en total 79 artículos.

## Resultados

### Comparaciones indirectas

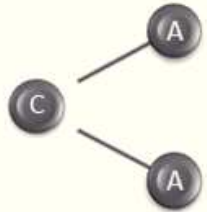
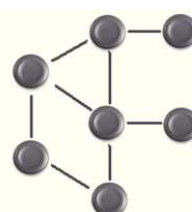
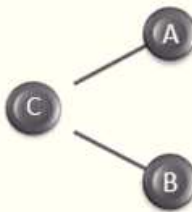
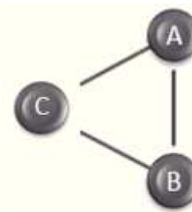
Las comparaciones entre intervenciones en salud se definen como indirectas si los efectos de las intervenciones que se comparan entre sí por su desempeño y efectividad, están mediadas por un comparador común (fig. 1), el cual puede ser placebo o una intervención activa, es decir, un medicamento de uso estándar<sup>15,16</sup>. Las CI no solamente se utilizan para realizar comparaciones entre medicamentos de la misma clase o dosis de la misma intervención, cuando por ejemplo, se realiza un sub-grupo de análisis, las comparaciones entre diferentes intervenciones farmacológicas versus cirugía o entre diferentes clases de medicamentos también se pueden apoyar en las CI<sup>3,24</sup>. Aunque la validez de las CI ha sido altamente cuestionada durante los últimos años, las evaluaciones de efectividad comparativa tendrán que depender de ellas en



**Figura 1 – Esquema básico de 2 ensayos clínicos aleatorizados con intervenciones diferentes A y B con un comparador común (placebo) que puede usarse para realizar una comparación indirecta. CI: comparación indirecta; ECC: ensayo clínico aleatorizado.**

ausencia de CD y su interpretación dependerá de la metodología empleada<sup>10,16,17,27</sup> (fig. 1).

Existen diferentes formas de realizar una CI y las definiciones para cada una de ellas varían en la literatura. En general, los métodos para las comparaciones indirectas pueden clasificarse en dos categorías: comparaciones no ajustadas (CINA) o «Naive», en las cuales entrarían las comparaciones informales (CII); y comparaciones ajustadas formales (CIA) que utilizan el método de Bucher<sup>23</sup>. Dentro de las CIA existen métodos que las pueden subclasificar como análisis de meta-regresión (MR), la red de meta-análisis (RMA) que utiliza el método de Lumley, y los métodos bayesianos<sup>1,14,18,28</sup>. Por otro lado, hay un híbrido compuesto por las comparaciones mixtas que incluyen CD y CI, que igual pueden usar los métodos descritos anteriormente. Entre ellas las CINA son las que tienen mayor riesgo al sesgo, mientras que CIA y los análisis MR ofrecen un alto grado de validez<sup>14,15</sup> (fig. 2).

|  |   |   |
|--|---|---|
| Meta-análisis                                | Combina los resultados de los estudios comparables de la misma intervención para obtener una estimación global del efecto               |   |
| Comparaciones indirectas no ajustadas (CINA) | Compara los resultados de los resultados individuales de diferentes estudios, como si se tratara de los ECC mismos                      | Narrativa<br>Toma resultados de comparaciones de tratamientos, hace una discusión y un juicio subjetivo, sin un cálculo estadístico formal  |
| Comparaciones indirectas ajustadas (CIA)     | Permiten la comparación entre dos intervenciones A y B, si ambas han sido comparadas con una tercera alternativa C; creando así una red | Redes de meta-análisis (RMA)<br>Más de dos estudios en red<br>   |
|  |   | Mixtas (CTM)<br>Reúne la información de varias comparaciones directas e indirectas «bucles cerrados»<br> |
|  | Métodos de análisis   | Meta-regresión  |
|  |   | Enfoque bayesiano   |
|  |   | Enfoque frecuentista  |

**Figura 2 – Tipos de comparaciones indirectas. ECC: ensayo clínico aleatorizado.**

### Comparación indirecta no ajustada o «Naive»

Las CINA se basan en la comparación de diferentes intervenciones usando datos de distintos estudios con un comparador común. La comparación de interés es entre las intervenciones A y B, que tienen a la intervención C como el comparador común, bien sea con placebo u otro fármaco, se comparan A y B de forma indirecta ignorando a C. Se realiza una agrupación meta-analítica de todos los grupos tratados con A para obtener una estimación global  $\theta_A$ , y de la misma manera con todos los grupos tratados con B para obtener  $\theta_B$ . Se rompe la asignación aleatoria dentro de los ECC y se tienen en cuenta los diferentes brazos de intervención individuales, de todos los estudios analizados, como si ellos fueran del mismo ECC<sup>15,16</sup>. Se denomina «CI no ajustada» porque no se ajusta para los eventos de la intervención C. Existen cuatro maneras de comparar el efecto global estimado  $\theta_A$  y  $\theta_B$ : cálculo de la estimación global de efectos  $\theta_A$  versus  $\theta_B$ , evaluar la diferencia entre  $\theta_A$  y  $\theta_B$  mediante significancia estadística, comprobar los intervalos de confianza entre ambos mediante sobre-posición, y comparación narrativa de la eficacia de A y B<sup>16</sup>.

La CINA ha sido ampliamente criticada por incrementar los sesgos y las sobre estimaciones, por ello los métodos de ésta no son recomendados y se consideran equivalentes a datos observacionales sujetos a sesgos similares<sup>1,14-16,22,29-32</sup>. Otro tipo de CINA, es la también llamada CINA narrativa, con menor poder de evidencia, la cual consiste en la discusión y juicio subjetivo, sin cálculo estadístico formal del efecto relativo ni de su significación estadística, a partir de los resultados de varias comparaciones de tratamientos en ECC. Comúnmente son empleadas en revisiones sistemáticas o narrativas temáticas, en las que por la gran diferencia entre las características de los estudios no es posible llevar a cabo un meta-análisis<sup>8,14,18</sup>.

### Comparación indirecta ajustada

Las CIA son el primer método válido disponible que hace posible llevar a cabo CI<sup>15</sup>. Cuando no existen CD cabeza a cabeza que permitan obtener información de efectividad de intervenciones se puede recurrir a las CIA<sup>32</sup>. Las CIA han sido usadas durante los últimos años de manera común y más frecuentemente en revisiones sistemáticas que evalúan los efectos de las intervenciones farmacológicas disponibles. La síntesis de los datos no es derivada de ensayos cabeza a cabeza necesariamente, pero utilizan datos de múltiples ECC de manera indirecta<sup>7,32-35</sup>. Siendo la comparación de interés A vs. B, y C el comparador común para ambos tratamientos, que puede ser un medicamento o placebo, mediante métodos derivados del meta-análisis se calculan los efectos globales estimados  $\theta_{AC}$  y  $\theta_{BC}$ <sup>8,16</sup>. Para la comparación de ambos efectos globales estimados se utilizan los mismos métodos empleados para las comparaciones indirectas no ajustadas<sup>16</sup>.

El propósito de mantener la aleatorización es asegurar que los factores de confusión conocidos y desconocidos se distribuyan uniformemente a través de los grupos de tratamiento, en las CI no es posible conocer si estos factores de confusión A y B fueron distribuidos uniformemente entre los ensayos AC

y BC, por lo tanto, la clave de este tipo de comparación, a diferencia de las comparaciones no ajustadas, es que no rompen la aleatorización inicial de cada ECC individual<sup>1,4,24,30,34</sup>. Las CI se pueden considerar similares a un análisis retrospectivo observacional<sup>29</sup>.

Los métodos estadísticos para las CI ajustadas deben cumplir con algunos supuestos, en este sentido, unos estudios consideran que el único requerimiento es que las poblaciones de los estudios individuales contengan alguna subpoblación en común, de otro lado, 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 en términos clínicos y metodológicos<sup>8,14</sup>. Para mejorar el poder estadístico de una CI ajustada, ésta puede combinarse con evidencia derivada de intervenciones cabeza a cabeza, para ello se cuenta, actualmente, con herramientas estadísticas que utilizan supuestos similares a los adoptados por los meta-análisis y las comparaciones pares, aunque requieren tener varios ECC<sup>2,14</sup>. En general, la validez de una CIA se garantiza por la calidad de los ECC, la homogeneidad de los resultados y la similitud entre los brazos de comparación<sup>8</sup>.

Las CI ajustadas utilizan el método sugerido en 1997 por Bucher et al.<sup>10</sup>, gracias a su menor complejidad, en éste las CI de las intervenciones nuevas y las clásicas son calculadas mediante los resultados de las CD entre ellas con una intervención de control, utilizando los OR como medida de estimación de los efectos<sup>30</sup>. Para ello, se pueden realizar las estimaciones mediante el uso de modelos de efectos fijos o de efectos aleatorios. El modelo de efectos aleatorios es el recomendado para prevenir la subestimación de los errores estándar en las CIA<sup>33</sup>. La CIA asume que la eficacia relativa de un tratamiento es la misma en todos los ECC incluidos, y asume la independencia entre las comparaciones por pares, que no se encuentran en los ECC de tres brazos<sup>26</sup>. Entre las limitaciones de este modelo se encuentra que únicamente puede ser aplicado a datos generados a partir de ECC de dos brazos, sólo se deben considerar efectos medidos mediante OR y comparaciones entre tres intervenciones A, B, C. Es importante recordar que cualquier sesgo en los estudios utilizados para las comparaciones podrá, también, afectar la validez de la CI<sup>27,33</sup>.

Entre los resultados de los efectos de los medicamentos comparados mediante CD y CI, no se han encontrado diferencias significativas, aunque los efectos en ECC cabeza a cabeza tienden a ser mayores en comparación con las CI ajustadas. Entre los factores que influyen en las diferencias se encuentran el azar, los sesgos de los ECC, los sesgos de la CIA, la heterogeneidad clínicamente significativa, las medidas del desenlace, las diferencias metodológicas y la variación de los efectos entre los subgrupos poblacionales<sup>35,36</sup>. En este sentido, las CIA surgen como una herramienta útil para rectificar e interrogar los resultados derivados del azar de los ECC cabeza a cabeza, el uso de las CIA debe ser extendido hacia los casos en que los ECC de CD son abundantes, como en la introducción de nuevas terapias. Las discrepancias entre las CD y las CIA deben ser cuidadosamente investigadas, con el fin de encontrar las explicaciones que deben incluir, principalmente, las deficiencias metodológicas y la heterogeneidad clínica significativa entre los ECC<sup>34</sup>.

### Meta-análisis convencional

Los MA son una forma de revisión sistemática, cuantitativa, en la que los estudios individuales son combinados y resumidos matemáticamente, todos los estudios incluidos comparan la misma intervención con el mismo comparador<sup>1,36-39</sup>. El principal beneficio del MA consiste en resumir los estimados de los efectos de la intervención, el cual resulta ser, la mayoría de las veces, más preciso que los obtenidos de los estudios individuales, sin embargo, si la heterogeneidad de los estudios es muy grande y la calidad de los mismos es baja, el efecto estimado podría ser, incluso, menor. El estudio de la heterogeneidad y la cuantificación del riesgo también son beneficios derivados de este tipo de estudio<sup>40</sup>. Esta herramienta se ha convertido en indispensable para la toma de decisiones en salud, considerándose como el más importante tipo de estudio en la jerarquía de la evidencia. El primer intento de resumir los resultados de estudios individuales fue realizado en 1904 por Karl Pearson, cuando examinaba el efecto preventivo de las inoculaciones de suero en el manejo de la fiebre entérica<sup>1,22,31,36,38,39,41-43</sup>.

La base estadística de los meta-análisis puede resumirse en dos modelos: 1) *el modelo de efectos fijos*, se basa en la suposición de que la única fuente de variación en los resultados es la que ocurre intra-estudio, es decir, que el efecto esperado de cada estudio es el mismo, así las diferencias encontradas en los resultados se deben a errores aleatorios (azar). Asume que los estudios son homogéneos, que no hay diferencias en población, criterios de selección de los sujetos y aplicación de tratamientos; y 2) *el modelo de efectos aleatorios*, asume que existe una distribución en los efectos, lo que da lugar a heterogeneidad entre los resultados del estudio, supone que los estudios incluidos en el meta-análisis son una muestra aleatoria de un universo de todos los estudios posibles, que respondan a una determinada pregunta de investigación. El modelo de efectos aleatorios incluye dos fuentes de variabilidad, intraestudio (error aleatorio) y entre estudios (heterogeneidad). Los efectos se ponderan según estas fuentes, así, los resultados varían en torno al efecto promedio global<sup>36,37</sup>.

Entre las limitaciones para garantizar los supuestos de los MA se encuentran la baja calidad de los estudios incluidos, los sesgos y la heterogeneidad; en este sentido es posible manejar esta última mediante el modelo de efectos aleatorios o de modelos mixtos que incluye efectos aleatorios más covariables. Cuando se realiza un MA debe ponerse de manifiesto que existe más heterogeneidad entre los resultados de los estudios incluidos de lo que se podría esperar por azar<sup>38,41</sup>.

La evaluación de la consistencia de los efectos a través de los estudios mediante la interpretación de la heterogeneidad entre ECC es un punto clave, por lo tanto, la heterogeneidad entre los resultados de los estudios debe ser analizada detenidamente para hallar su causa; ésta puede derivarse de las diferencias en las características de los pacientes incluidos en cada ECC, las cuales están relacionadas con la eficacia del tratamiento y también deben tenerse en cuenta las diferencias entre los criterios de inclusión y exclusión de los estudios (edad, severidad de la enfermedad, estratificación del riesgo). La identificación de las características de los pacientes como causa de heterogeneidad puede

ayudar a identificar los tipos de pacientes que se pueden beneficiar o no de las intervenciones estudiadas<sup>34,40,43</sup>. De la misma manera, las características del diseño del estudio, tales como diferencias en los métodos de administración de la intervención, tiempo de seguimiento y calidad del diseño, pueden conducir a la heterogeneidad observada en el MA. Las variaciones entre poblaciones de áreas geográficas diferentes también aportan a la heterogeneidad<sup>40</sup>.

### Red de meta-análisis

La RMA es una técnica de CI desarrollada por Lumley<sup>18</sup>, es un tipo especial de meta-análisis, considerada como una extensión de MA de comparaciones pares indirectas<sup>1,44</sup>. En los MA todos los estudios incluidos comparan la misma intervención con el mismo comparador, en un «network meta-analysis» o RMA este concepto se amplía mediante la inclusión de múltiples intervenciones pares (dos intervenciones son comparadas), a través de una serie de comparaciones y proporciona estimaciones del efecto con fines de efectividad comparada mediante los tratamientos múltiples<sup>1,22,44</sup>.

La RMA se caracteriza por combinar pruebas directas e indirectas para comparaciones pares, con lo que se obtiene una mayor síntesis de la evidencia disponible que el tradicional meta-análisis, permite la combinación de ECC que comparen diversos grupos de tratamiento, proporciona estimaciones del efecto relativo del tratamiento entre las intervenciones de la competencia y permite una estimación de los efectos heterogéneos de cualquier medicamento<sup>1,26,45</sup>. Por ejemplo, si dos tratamientos A y B nunca han sido comparados entre sí, pero se han comparado frente a un placebo C: los medicamentos A y B han sido superiores a C, mediante una red de meta-análisis se puede hacer una comparación indirecta entre A y B<sup>1,18,46</sup>. Las RMA se basan en la suposición de que en una escala adecuada, se pueden sumar y restar estimaciones de los efectos relativos del tratamiento, la diferencia ( $\Delta$ ) en el efecto entre los tratamientos A y B ( $\Delta_{AB}$ ) es igual a la diferencia de efectos entre los tratamientos AC y BC y C<sup>47</sup>.

$$\Delta_{AB} = \Delta_{AC} - \Delta_{BC}$$

Una RMA se inicia con una red de prueba, las intervenciones de interés y los ECC que han comparado los tratamientos directamente. Su estructura, generalmente, es evidente a partir de un diagrama en el que cada nodo representa una intervención o una clase de tratamientos, y cada línea representa el enlace que conecta los tratamientos que han sido comparados directamente en uno o más ECC<sup>26</sup>. El principal supuesto de la RMA es que la comparación entre dos intervenciones puede ocurrir en un circuito cerrado (fig. 1), este diseño es necesario para calcular la inconsistencia entre la evidencia<sup>27</sup>. Otros supuestos son que los estudios combinados únicamente deben ser considerados si son clínica y metodológicamente similares y cuando se combinan evidencia directa e indirecta se asume la consistencia entre ambas<sup>26</sup>.

Un MA de una CD, usualmente, requiere que los estudios individuales estimen tanto el efecto del tratamiento común o los efectos de los tratamientos específicos del estudio, dis-

tribuidos alrededor de un valor típico. La selección entre un efecto común y la distribución del efecto da lugar a los enfoques de efectos fijos o aleatorios, respectivamente<sup>26,48</sup>. La heterogeneidad entre los estudios de una CD es aceptable, sin embargo, en la combinación de evidencia directa e indirecta la heterogeneidad conduce a una pérdida de similitud. Por lo tanto, en la RMA es necesario evaluar en todos los casos la heterogeneidad, consistencia y similitud<sup>26</sup>.

Las ventajas de la RMA son la capacidad de detectar inconsistencias entre los ECC de los diferentes tratamientos, estimar las diferencias de tratamiento y evaluar la incertidumbre en estas estimaciones<sup>18</sup>. La RMA se ha usado para comparar la eficacia de diferentes terapias antihipertensivas, evaluación de biológicos para el manejo de la artritis reumatoide, VIH y en análisis de costo-efectividad de diferentes terapéuticas<sup>10,22,25,49,50</sup>.

Las redes de MA son menos fiables que las CD aleatorizadas, pero son muy útiles en la planificación de CD y en los casos en que las CD no sean éticamente posibles o no se encuentren disponibles. Sin embargo, los resultados de las pruebas indirectas en conjunto con las directas pueden reforzar la evaluación entre las intervenciones consideradas, obteniendo una estimación más precisa y ampliar la inferencia a la población de la muestra<sup>1,2,14</sup>. Aun así, la inconsistencia posible entre los resultados de la comparación podría, en determinado momento, alterar la confiabilidad general de la red e impide la extrapolación de dichos resultados<sup>22</sup>. Otra limitación es que puede combinar la evidencia de ensayos muy heterogéneos de manera desapercibida, por lo que estudios de baja calidad metodológica o con diferencias clínicas pueden tener una influencia indebida en los resultados<sup>45</sup>.

### Comparaciones de tratamientos mixtos

Nuevos métodos para evaluar la efectividad relativa de intervenciones competitivas con mayor coherencia y eficiencia, ofrecen oportunidades únicas para investigación en efectividad comparativa y soportar la toma de decisiones en salud<sup>13,51,52</sup>. El factor clave para decidir entre el uso de evidencia directa sola o una combinación con indirecta, es la cantidad de evidencia directa que existe, es decir, si se encuentra disponible una cantidad suficiente de evidencia directa con resultados conclusivos probablemente no sea necesario considerar las CI<sup>8,15</sup>. La combinación de evidencia directa e indirecta proveniente de los ECC, se conoce como comparaciones de tratamientos mixtos<sup>53,54</sup> (CTM). Otro término para las CTM es meta-análisis de tratamiento mixto (MTM), y dada su similitud con el modelo propuesto por Lumley también se hace referencia a ellas como red de meta-análisis o parte de ellas<sup>27</sup>.

Las CTM son una generalización de MA de comparaciones pares, dirigidas a sintetizar la información sobre más de dos comparaciones, consideradas extensiones de los métodos de meta-análisis con gran potencial para estimar la efectividad comparativa de múltiples intervenciones, con base en ECC que individualmente no comparan todas las opciones de tratamiento<sup>13,15,30,53,55</sup>, mediante la realización de juicios coherentes sobre cuál de los tratamientos es el más efectivo y generación de estimaciones de los efectos relativos de cada

intervención comparada con cada una de las otras analizadas en una red<sup>56</sup>.

El objetivo de una CTM es combinar toda la información disponible en un conjunto de estimados coherentes en la cual se respeta la aleatorización original de toda la evidencia incluida<sup>55</sup>. Las CTM permiten la estimación de la efectividad de cada intervención en relación con cada una de las otras, se hayan o no comparado directamente en ECC<sup>15,54,56</sup>. También permite calcular la probabilidad de cada tratamiento para ser el más efectivo entre el conjunto estudiado<sup>57</sup>. Este enfoque únicamente puede ser aplicado a redes conectadas de ECC que relacionen cada intervención de interés con otro tratamiento<sup>30,54,56,58</sup>.

La combinación de la evidencia directa e indirecta para tres o más tratamientos es posible mediante la manipulación de los resultados de comparaciones pares, por ejemplo, por medio del uso de resúmenes meta-analíticos de estas comparaciones para cada intervención. Extensiones de los modelos de efectos estándar (fijos y aleatorios) son comúnmente usados para este propósito, sin embargo, también se propone el uso de un modelo jerárquico bayesiano para k-comparaciones, donde k es el número de intervenciones por evaluar, el cual es una extensión natural de un modelo meta-analítico diseñado para sintetizar comparaciones entre tratamientos pares<sup>58,59</sup>. Usando métodos bayesianos que involucran cadenas de Markov de Monte Carlo, mediante herramientas de software como el «BayesianWinBUGS», es posible clasificar los tres o más tratamientos de interés y determinar una probabilidad que indique cuál de ellos es el «mejor». Ésta es una muestra potencial de la habilidad de los métodos estadísticos bayesianos para realizar declaraciones de probabilidad<sup>27,30,58</sup>.

Teniendo k cantidad de tratamientos por comparar con  $k \geq 2$ , el tratamiento A representa el tratamiento de referencia en un modelo de meta-análisis de k comparaciones de tratamiento. Cuando es de interés comparar dos tratamientos pero la evidencia de ECC es limitada, las CTM pueden realizarse si existe suficiente evidencia para comparar cada tratamiento A y B frente a C. El tratamiento C actúa como el tratamiento de referencia y cada uno de los anteriores (A y B) debe ser comparado con C. Las  $\Delta AC$  y  $\Delta BC$  representan la eficacia relativa de los tratamientos A y B con relación al tratamiento C, respectivamente. Posteriormente, la eficacia relativa del tratamiento B en relación con el A puede obtenerse así<sup>27</sup>:

$$\Delta AB = \Delta BC - \Delta AC$$

Los beneficios potenciales de extender los MA a enfoques de CI y CMT son multifacéticos<sup>30</sup>. Las principales ventajas de las CTM sobre otras CI como las propuestas por Bucher, et al.<sup>10</sup>, y Song et al.<sup>33</sup>, son su aplicación para evaluar la eficacia relativa entre dos o más intervenciones, mediante la combinación de toda la evidencia directa e indirecta en un análisis único y simultáneo de una gran cantidad de comparaciones, facilitar la deducción simultánea en relación con todos los tratamientos con el fin de compararlos, clasificarlos y seleccionar el mejor, optimizar el poder estadístico de las CI proporcionando estimaciones más precisas y reduciendo la incertidumbre con intervalos de confianza más estrechos en comparación con los resultados basados en CD o CI sola-

mente, y permitir la comprobación formal de la consistencia entre la evidencia directa e indirecta mediante el análisis de ésta<sup>27,30,54,57,58</sup>. Por otro lado, mediante el uso de métodos bayesianos, como se comentó previamente, existe la posibilidad de determinar la probabilidad que tiene cada una de las intervenciones incluidas para ser la mejor entre el grupo de análisis. Sin embargo, entre las limitaciones de las CTM se encuentran la validez de los juicios hechos a priori y la complejidad del método.

#### *Supuestos de una comparación de tratamientos mixtos*

Los supuestos requeridos para una CTM son exactamente los mismos relacionados con un meta-análisis de comparación par o una CI individual, excepto que se deben mantener para todo el conjunto de estudios y se basan en la consistencia entre las comparaciones. Factores como la variación del efecto del tratamiento dependiente de la severidad de la enfermedad pueden invalidar una CTM. No debe considerarse una limitación de la metodología, sino una limitación de los datos, ya que para una población definida el modelo será compatible por definición<sup>30</sup>.

Un supuesto fundamental en el modelo de efectos fijos de los meta-análisis es que el efecto real es una constante entre todos los estudios, asumiendo que todos los ensayos AB estiman el mismo efecto relativo de tratamiento de AB (B con respecto a A), y que todos los estudios AC, BC e inclusive CD, podrían estimar este efecto si ambos brazos A y B han sido incluidos en ellos. De otro lado, para un modelo de efectos aleatorios, se asume que los efectos relativos de tratamiento ensayo-específico provienen de una distribución común independientemente de si las intervenciones A y B fueron incluidas en el mismo estudio. Más fácilmente, si uno puede imaginar que cada ensayo incluye cada tratamiento, entonces, el supuesto clave es que los brazos faltantes son aleatorios, sin importar la eficacia del tratamiento. La mayoría de las dudas respecto a las CMT se refieren a si este supuesto de consistencia entre la evidencia directa e indirecta es seguro<sup>52,56,58</sup>.

En el caso en que se desee realizar CTM para combinaciones de tratamientos, los supuestos se basan en cómo los tratamientos interactúan entre sí cuando son combinados, de esta forma se pueden construir diferentes modelos de CTM: en los cuales todos los efectos son asumidos para ser completamente aditivos «modelo de efectos aditivos», cada tratamiento puede interactuar «modelo de interacción de dos vías» o cada conjunto de tratamiento puede interactuar «modelo de interacción total». En términos de estimación de efectos comparativos entre todos los tratamientos el «modelo de interacción total» es idéntico al modelo de CTM convencional, el cual puede ignorar la dependencia entre la efectividad de las combinaciones de medicamentos, incluyendo la efectividad de la intervención individual<sup>57</sup>.

#### *Evaluación de la consistencia en comparaciones de tratamientos mixtos*

La consistencia se refiere a que cada intervención para el efecto verdadero del tratamiento se asume como similar para cada estudio en el análisis, independientemente de si el estudio asignó los pacientes a esos dos tratamientos. El supuesto

de consistencia se puede considerar como una extensión del supuesto de homogeneidad y no debe ser tomado cuando uno o más de los efectos verdaderos de tratamiento son modificados por características particulares y los ensayos incluidos difieren con respecto a las características, por consiguiente, si las interacciones son detectadas, el supuesto podría ser poco realista y los resultados del ECC original serían cuestionables<sup>54</sup>. Los «modelos de inconsistencia» han sido propuestos para valorar la consistencia entre evidencia directa e indirecta en redes complejas. En tales modelos los factores de inconsistencia son adicionados a las ecuaciones básicas de consistencia y esta es valorada mediante la comparación de los modelos con y sin dichos factores. Sin embargo, es necesario revisar con más detalle las comparaciones específicas, para evidenciar más claramente en cada caso si la evidencia indirecta se combina o se suma a la evidencia directa para formar las estimaciones de las CTM<sup>56</sup>.

Existen, en principio, dos métodos para verificar la consistencia entre la evidencia directa e indirecta, ambos permiten la comparación de estimados basados en evidencia directa contra los basados en evidencia indirecta. El primer método es el «back-calculation» o cálculo retrógrado, que deduce la contribución de evidencia indirecta de la evidencia directa y el desenlace de un análisis de CTM, siendo útil cuando sólo los datos disponibles consisten en resúmenes combinados de comparaciones pares. El segundo «node-splitting» o separación de nodos es más general pero computacionalmente intensivo, separa la evidencia de una comparación particular (nodo) en directa o indirecta, y puede ser aplicado a redes en donde se encuentran disponibles datos a nivel de ensayo<sup>56</sup>.

Ambos métodos pueden ser usados para graficar las consistencias posteriores de los efectos de tratamiento obtenidos, usando evidencias directas e indirectas únicamente, lo cual ofrece una herramienta visual para valorar la concordancia entre ambos tipos de evidencia. Estas gráficas posteriores de consistencia no muestran únicamente cómo se comporta la consistencia entre las evidencias directa e indirecta, sino también el grado de precisión con el que se comparan entre sí y cómo afectan el estimado final de la CTM. El grado de concordancia entre las evidencias directa e indirecta (solapamiento entre ambas curvas) es medido formalmente por los p-valores bayesianos<sup>56</sup>.

---

## **Métodos de análisis**

Para sintetizar los resultados de los diferentes ECC en las CI que comparan más de dos intervenciones simultáneamente y obtener sus efectos relativos estimados, es necesario utilizar métodos estadísticos complejos que preserven la aleatorización y brinden beneficios adicionales según el resultado que se desea obtener. Para la mayoría de los modelos de análisis se podría seguir ya sea un enfoque frecuencial o una aproximación bayesiana. En la práctica actual de la mayoría de los meta-análisis para las comparaciones directas se utilizan métodos frecuentistas. Para los modelos más complicados, en particular las redes que implican comparaciones de tratamientos mixtos, los métodos bayesianos han sufrido sustancialmente mayor desarrollo, facilitado por los avances en



la informática que apoyan su intensidad computacional y la necesidad de controlar la convergencia a la distribución posterior<sup>1,15,27</sup>.

### Enfoque frecuentista

La estadística frecuentista ha sido el enfoque dominante en la investigación clínica durante las últimas cinco décadas<sup>26,60</sup>. La etiqueta de «frecuentista» se aplica a la mayoría de los métodos estadísticos tradicionales aplicados al hacer comparaciones, incluyendo las medias ponderadas con intervalos de confianza (sobre la base de una distribución normal asumida). En este modelo, la probabilidad es el límite de la frecuencia relativa, y asume que el experimento se puede repetir bajo condiciones idénticas. Sin embargo, solo aplica a eventos que aún no han ocurrido y a fenómenos repetitivos como lanzar una moneda<sup>26</sup>.

La principal virtud del enfoque tradicional es el rigor extremo y la estrechez del enfoque para el experimento de interés, sin embargo, la inflexibilidad relacionada con este enfoque limita la innovación en el diseño y el análisis de los ECC. Por esta razón los ECC tienden a ser extensos y costosos. La inferencia estadística frecuentista es un enfoque inverso al bayesiano. Se centra en que la probabilidad de los resultados de un ensayo supone que una hipótesis particular es verdadera. Así, un *p*-valor frecuentista es la probabilidad de observar resultados más extremos que los resultados observados, suponiendo que la hipótesis nula es verdadera por lo que existe una fuerte dependencia de las conclusiones sobre el diseño, ambas características actualmente son cuestionables científicamente<sup>13,60-62</sup>.

### Técnicas de meta-regresión para CI

Los métodos de meta-regresión (MR) han sido usados cada vez más frecuentemente en CI, su popularidad consiste en su potencial para explicar la heterogeneidad de los estudios y ajustar los factores de confusión, con el fin de apoyar las decisiones clínicas en cuanto a cómo, dónde y para quién un tratamiento es benéfico<sup>36,40,43</sup>.

Las MR involucran el estudio de múltiples variables mediante modelos de regresión, que explican la variación en el desenlace y si la heterogeneidad es dada por una o más características del estudio, de la población o del tratamiento. Su objetivo es relacionar la magnitud del efecto de una o más características de los diferentes estudios incluidos<sup>36,40,43,53,63-65</sup>. En ella los efectos globales estimados  $\theta_{AC}$  y  $\theta_{BC}$ , siendo C el comparador, son calculados de manera separada en dos modelos de regresión. Además de ajustar los efectos en los grupos de los comparadores, los modelos de regresión pueden ajustarse para los efectos de las diferentes variables consideradas para controlar la heterogeneidad, por ejemplo: la edad de la población o la severidad de la enfermedad<sup>16</sup>. En este sentido, las CIA pueden ser vistas como una de las formas más simples de MR<sup>24</sup>.

Los factores predictores de la regresión de cada estudio incluyen el protocolo, las características de la población o variables que describen la logística del mismo. Las preguntas que se pueden resolver mediante una MR incluyen la esti-

mación de los efectos de las intervenciones, controlando las diferencias entre los estudios, y la determinación de múltiples variables que pueden comprometer la heterogeneidad<sup>64</sup>. Las asociaciones derivadas de una MR son observacionales y, por lo tanto, tienen una interpretación más débil que las relaciones causales derivadas de comparaciones aleatorizadas<sup>52</sup>.

### Enfoque Bayesiano

Los métodos estadísticos bayesianos han venido siendo usados en investigación clínica y farmacológica cada vez más, el uso del enfoque Bayesiano representa un cambio fundamental en la conducta de la investigación médica, como consecuencia del progreso rápido en el desarrollo de medicamentos y el menor costo de los mismos<sup>53,60,61,66,67</sup>. En la teoría frecuentista los parámetros distribucionales (media, varianza, proporción, etc.) son fijos y desconocidos. En el enfoque bayesiano estos parámetros se asumen aleatorios. La inferencia bayesiana exige el conocimiento de una probabilidad *a priori*, es decir, conocida de antemano (estudios previos, experiencia, opinión de expertos, un ensayo previo o actual) y de los datos obtenidos, con lo que se procede al cálculo de la **verosimilitud** (likelihood) y, finalmente, se obtiene una probabilidad *a posteriori*<sup>26,27</sup>. La verosimilitud es una función de los parámetros de un modelo estadístico que permite realizar inferencias acerca de su valor, a partir de un conjunto de observaciones. No debe confundirse con el término probabilidad el que a partir de una serie de parámetros conocidos permite realizar predicciones acerca de los valores que tomaría una variable aleatoria<sup>62</sup>.

Si la distribución posterior se encuentra en la misma familia que la distribución de probabilidad anterior son, entonces, llamadas distribuciones conjugadas o un «*conjugate prior*» de la verosimilitud. Por ejemplo, la familia gaussiana se conjuga a sí misma con respecto a una función de probabilidad gaussiana: si la función de probabilidad es gaussiana, la elección de una gaussiana antes sobre la media asegurará que la distribución posterior también sea gaussiana. Consideremos el problema general de inferir una distribución para un parámetro  $\theta$  dado algún dato  $x$ , por el teorema de Bayes, la distribución posterior  $p(\theta|x)$  es igual al producto de la función de verosimilitud  $\theta \rightarrow p(\delta|x)$  y la distribución anterior,  $p(\theta)$  dividido por la probabilidad de los datos  $p(x)$ <sup>66,68</sup>.

$$f(\theta|x) = \frac{p_{\frac{x}{\theta}} \times p(\theta)}{p(x)}$$

Una RMA puede realizarse dentro de un marco bayesiano o frecuentista. Con un enfoque frecuentista, el resultado del análisis es una estimación puntual con un intervalo de confianza del 95% (IC). Un IC del 95%, bajo muestreo repetido, contendría el parámetro real de la población del 95% de las veces. Debe tenerse en cuenta que la CI obtenida con un enfoque frecuentista no puede ser interpretada en términos de probabilidades, el IC del 95% no significa que hay 95% de probabilidad de que la población valor está entre los límites del intervalo<sup>1</sup>. Los métodos bayesianos implican una combinación formal de una distribución de probabilidad previa, lo que refleja la creencia previa de los valores posibles del parámetro del modelo de interés, con una distribución de



probabilidad de estos parámetros sobre la base de los datos observados, para obtener una distribución de probabilidad posterior correspondiente<sup>1,27</sup>.

El modelo jerárquico bayesiano fue propuesto en 1995 por Smith, Spiegelhalter, Caldwell y Thomas<sup>59</sup>; y, a su vez, el más completo de los diferentes modelos propuestos para la realización de la red meta-análisis Bayesiano parece ser el propuesto por Lu y Ades<sup>27</sup>. Es la continuación de los modelos anteriores propuestos por Caldwell y Smith, basado en la implementación de las técnicas de integración numérica de cadenas de Markov de Monte Carlo (CMMC), como metodología alternativa a los modelos de efectos fijos y aleatorios para meta-análisis de comparaciones pares, y se ha extendido al modelo general para CTM de k-tratamientos<sup>15,27,58,59</sup>.

La necesidad de una distribución a priori para el elemento que no se conoce, pero que se desea conocer a través de la observación, es el principal problema con este enfoque. Este es un punto importante, ya que la distribución posterior estimada puede cambiar considerablemente en función de la anterior elegida. Afortunadamente, es posible encontrar antecedentes poco informativos, que no afectan el resultado. Tales distribuciones previas poco informativas parecen decir que antes del análisis (es decir, antes de que se han observado los primeros datos), cualquier valor es posible que el verdadero valor de un determinado parámetro y que, por lo tanto, no se justifica para favorecer a uno más que a otro<sup>27</sup>.

Una ventaja de los enfoques bayesianos es que el método conduce naturalmente a un marco de decisión que apoya la toma de decisiones, además, permite a los investigadores actualizar el conocimiento mediante la acumulación de información en ensayos. Tal actualización puede ocurrir en un momento dado o de manera continua en el tiempo con la posibilidad de modificar el diseño del estudio, lo que incluye desequilibrar la aleatorización para favorecer terapias con mejor desempeño, quitar o adicionar brazos de tratamiento y cambiar la población de estudio para concentrarse en los subgrupos de pacientes que responden mejor a las terapias experimentales<sup>66</sup>. Para una RMA, una ventaja específica es que la distribución de probabilidad posterior permite el cálculo de la probabilidad de que una de las intervenciones que compiten es mejor. Este aspecto de un análisis bayesiano permite proporcionar información que es directamente relevante para los tomadores de decisiones de atención de salud. En este sentido los enfoques bayesianos son únicos<sup>27,53,61,62</sup>.

Los métodos bayesianos tienen grandes ventajas en el campo del desarrollo de medicamentos. Entre ellas se incluye la capacidad de sintetizar información, el proceso de actualizar el conocimiento de manera progresiva en cualquier momento y sin restricciones, como resultado de la acumulación de información en cada ECC, la utilidad vinculada a la toma de decisiones en salud, la capacidad de calcular y utilizar probabilidades de predicción de distribución de resultados futuros con base en los resultados actuales, la construcción de modelos jerárquicos para apropiarse de la información entre terapias o subtipos de enfermedad, donde la fuerza de endeudamiento depende de la homogeneidad de los datos y estimación de la evidencia directa para efectos de interés, usando la probabilidad posterior en lugar de calcular indirectamente

la probabilidad de los datos observados, suponiendo efectos particulares mediante el uso de p valores<sup>13,61,65,66,69,70</sup>.

La investigación en efectividad comparativa, es el análisis de toda la evidencia disponible de los diferentes estudios que abordan un problema determinado en términos de riesgo vs. beneficio. Se considera el enfoque estadístico bayesiano como el método ideal para este tipo de investigaciones, dado que es inherentemente sintético y además no tiene limitaciones filosóficas en cuanto a la capacidad de analizar en simultáneo toda la evidencia disponible<sup>53,60,69</sup>. Sin embargo, no es la panacea, aunque tiene múltiples ventajas sobre el uso del enfoque frecuentista, por ejemplo, no es inmune a los sesgos ya que no puede convertir la evidencia de pobre calidad en buena evidencia, aunque puede incluir la posibilidad de hallar sesgos, es usualmente difícil obtener información sobre ellos y, generalmente, se incluyen desapercibidos en la distribución a priori y lucha de la misma manera que el enfoque frecuentista con los datos perdidos; y aunque intenta el rescate de datos perdidos mediante imputación múltiple, no logra la credibilidad suficiente y, por lo tanto, el vacío persiste de alguna manera<sup>70,71</sup>.

Las consecuencias de tomar un enfoque bayesiano para el diseño de un ECC son la eficiencia, la comparación de intervenciones en ausencia de comparaciones cabeza a cabeza, la selección del mejor tratamiento para los pacientes en el ECC y la gran precisión en relación con los desenlaces primarios, por ejemplo, la relación entre desenlaces tempranos y tardíos<sup>13,67</sup>.

---

## Validez, calidad y problemas metodológicos

Las CI tienen muchas limitaciones y no logran sustituir los ECC, por lo tanto, es importante ser consciente de ellas para definir cuándo se pueden utilizar y en qué medida sus resultados son válidos y confiables<sup>28</sup>. Es imprescindible, entonces, analizar su validez interna y externa, de la misma manera que para todos los estudios y métodos. Son tres los pilares fundamentales para garantizar la **validez interna** de las CI: 1) la identificación adecuada de los estudios, 2) la alta calidad de los estudios incluidos y 3) el grado de sesgo de confusión debido a violaciones de similitud y consistencia<sup>1,6,27</sup>.

La calidad de una CI dependerá de diversos factores, incluida la metodología escogida y la validez de los supuestos. En varias áreas clínicas la evidencia de alta calidad no existe o es inconclusiva, por lo que el desarrollo de CI es una alternativa útil<sup>3</sup>. La validez está condicionada a la validez interna de los ECC individuales incluidos, y es importante tener en cuenta que no es posible obtener resultados de buena calidad en una combinación, si los estudios incluidos son de baja calidad, aunque hayan sido bien seleccionados<sup>28</sup>.

Los supuestos clave de las CI son la similitud clínica y metodológica, la consistencia y la homogeneidad, éstos pueden considerarse como una extensión del supuesto original de homogeneidad en los MA. La valoración de dichos supuestos es vital para asegurar la calidad, validez e interpretación correcta de los resultados de las CI, por lo que cualquier violación a estos principios puede invalidar la CI<sup>3,28</sup>.

El supuesto principal de las metodologías de CI es la similitud entre los efectos de tratamiento, esto es, el efecto

verdadero del tratamiento en la comparación de dos o más intervenciones sería similar entre todos los estudios independientemente si se incluyen una o ambas de dichas intervenciones. Si el supuesto de similitud se viola, la validez de la CI será cuestionable y es necesario valorar en qué forma influye en el resultado. La similitud, sin embargo, es difícil de valorar dado que no existen métodos estándar que determinen el supuesto de similitud aunque se proponen los análisis de sensibilidad, análisis de subgrupos y MR<sup>1,3,6,14,24,26,27</sup>.

La consistencia se refiere a que los resultados entre la evidencia directa e indirecta para la misma comparación de tratamiento sean similares, en caso contrario se deben investigar las causas de la discrepancia<sup>3,28</sup>. Las razones potenciales para que no haya consistencia entre las CD y CI incluyen: el azar, la heterogeneidad clínica, la heterogeneidad metodológica, la heterogeneidad estadística, la medida del efecto, los sesgos de publicación y, por último, el informe selectivo de desenlaces o de los análisis<sup>1,28,33</sup>.

Por último, si en la CI se van a utilizar resultados obtenidos de un MA, es necesario que los estudios que se combinen sean similares y sus resultados sean homogéneos. Esta variabilidad se define como heterogeneidad y es la variabilidad en el resultado medido en cada estudio con respecto al resultado global promedio<sup>3,6,72</sup>. Según Thompson, la heterogeneidad se clasifica en clínica y de diseño<sup>1,6,64</sup>. La heterogeneidad clínica se refiere a que los estudios no se relacionan con el objeto de interés, es decir, comprende las diferencias entre las poblaciones estudiadas para el mismo tratamiento o las diferencias en los desenlaces medidos, por lo que no depende del investigador<sup>64</sup>. Por su parte, la heterogeneidad de diseño es susceptible de control por parte del investigador ya que puede decidir cuándo el estudio debe ser retrospectivo o prospectivo, cuánto es el tiempo de seguimiento, qué desenlace será medido y los métodos estadísticos que serán empleados, entre otros<sup>1,6,28,64</sup>.

A diferencia de la similitud, si existen métodos estándar para valorar la homogeneidad, ésta puede cuantificarse mediante la prueba de Chi-cuadrado (Estadístico Q), aunque tiene un bajo poder estadístico, y el Test I<sup>2</sup> que es la proporción de la variación de cada estudio respecto a la variación total. Valores del 25%, 50% y 75% corresponderían a una heterogeneidad baja, moderada y alta, respectivamente. De la misma manera, las características del diseño del estudio tales como diferencias en los métodos de administración de la intervención, tiempo de seguimiento, factores de calidad del diseño, pueden conducir a heterogeneidad que es observada en el MA. Las variaciones entre poblaciones de áreas geográficas diferentes también aportan a la heterogeneidad<sup>3,6,40,73</sup>.

Es importante recordar que la calidad de las revisiones y, por lo tanto, de las CI depende de preservar la aleatorización original de los ECC incluidos, la aplicación de medidas del efecto del tratamiento y de la precisión. También se considera alta calidad si la homogeneidad se mide por algún método como el test de Chi-cuadrado, I-cuadrado estadístico o la estimación entre la varianza de los estudios en un modelo de efectos aleatorios. La valoración de la consistencia y el supuesto de inconsistencia cuando se combina evidencia directa con indirecta, también mejoran la calidad de la comparación<sup>3,26</sup>.

En general, se considera que el supuesto básico de los meta-análisis de efectos fijos es la homogeneidad, para las CIA se requiere, además, suponer la similitud, y para la combinación de evidencia directa e indirecta ajustada, es necesario añadir el supuesto de consistencia entre la evidencia mediante los modelos de efectos fijos o aleatorios, dado que aunque la CIA sea válida puede no ser consistente con la evidencia directa de ECC, debido a la heterogeneidad clínica significativa<sup>14</sup>.

La **validez externa** analiza en qué medida se pueden extrapolar los resultados a una población particular o de interés y está limitada por la validez externa de cada ECC incluido en la red de ensayos. Desde el punto de vista de la toma de decisiones, cierta heterogeneidad entre los ECC en la red de ensayo puede aumentar la validez externa, siempre que la heterogeneidad en las CD sea superior a la variación de los modificadores del efecto entre los estudios de la CI para evitar violaciones de similitud y los sesgos tanto como sea posible<sup>1,28</sup>.

---

### Utilidad de las comparaciones indirectas en artritis reumatoide

La artritis reumatoide (AR), es una enfermedad crónica, autoinmune, de etiología desconocida, que se caracteriza por la inflamación progresiva y el daño articular irreversible; llevando finalmente a una disminución de la calidad de vida y mortalidad prematura<sup>74-76</sup>. Los medicamentos modificadores de la enfermedad (DMARs) solos o combinados han demostrado su capacidad para retrasar la progresión de la enfermedad<sup>77</sup>. Aquellos pacientes con AR activa temprana y establecida, de intensidad moderada-severa, que han fracasado a tratamiento con DMARs, incluido metotrexate (MTX), son candidatos a recibir terapias biológicas<sup>76</sup>. Actualmente, se dispone de ocho agentes biológicos para el tratamiento de la AR: cinco medicamentos que actúan bloqueando el factor de necrosis tumoral alfa (anti-TNF): infliximab, etanercept, adalimumab, certolizumab y golimumab; un inhibidor de la IL-1: anakinra; una proteína de fusión moduladora de la activación de células T: abatacept y un anticuerpo monoclonal contra el receptor de la IL-6: tocilizumab<sup>19,76,77</sup>. Todos estos agentes han sido probados frente a MTX para la eficacia en la prevención de progresión de la enfermedad en varios ECC<sup>77</sup>.

Aunque los biológicos típicamente han sido comparados con placebo, no existen ECC comparando medicamentos biológicos cabeza a cabeza. Se hace necesario conocer la efectividad y seguridad para decidir sobre el mejor tratamiento para cada paciente<sup>19,77</sup>. Por lo tanto, dado el elevado costo de estos medicamentos se hace necesario usar otros métodos como las CI para ayudar a la toma de decisiones en el área de la reumatología. Hasta el momento se han realizado diversas CI para evaluar la eficacia de los diversos tratamientos para la AR. En una de las CI encontradas en la literatura, que comparó la efectividad de DMARs y biológicos, encontró que todos los DMARs y biológicos fueron significativamente más eficaces que el placebo en el siguiente orden: certolizumab, tocilizumab, rituximab, infliximab, etanercept, adalimumab, golimumab, abatacept, anakinra, y MTX. En las comparaciones por parejas, cada clase fue más eficaz que el MTX, pero

ninguno fue más eficaz que otro<sup>19</sup>. Mientras una CI sobre biológicos, evidenció que en comparación con el placebo, los biológicos se asocian con una mayor mejoría clínica ACR50 (American College of Rheumatology). Sin embargo, los biológicos se asociaron con más eventos adversos y anakinra fue el menos eficaz de todos los otros biológicos. Dentro de los resultados más representativos de este estudio, anakinra fue menos efectivo que el adalimumab y etanercept, y etanercept fue más seguro que adalimumab, anakinra e infliximab<sup>78</sup>. En otra CI de biológicos, en pacientes con AR previamente tratados con MTX y biológicos en ausencia de comorbilidades, no existen diferencias entre infliximab, etanercept, adalimumab, certolizumab, golimumab, abatacept, anakinra y tocilizumab en la eficacia en términos de respuesta ACR20, ACR50, ACR70, ni en la interrupción del tratamiento debida a efectos adversos<sup>76</sup>. Por otro lado, los resultados de una CI publicados en Cochrane respecto a la eficacia de los biológicos en el tratamiento de AR, mostraron que los anti-TNF son similares, se encontró una única diferencia respecto al infliximab que fue superior a anakinra en la mejoría del ACR20<sup>79</sup>.

Finalmente, en ausencia de ECC de comparaciones cabeza a cabeza que evalúen la eficacia y seguridad de medicamentos biológicos para el tratamiento de enfermedades autoinmunes, y especialmente los que comparen la eficacia de biológicos en el tratamiento de la AR, las CI, como son las CTM y las RMA surgen como una herramienta complementaria para ayudar a profesionales de la salud y dirigentes públicos en la toma de decisiones en salud<sup>21</sup>.

## Conclusiones

Para la toma de decisiones en salud es necesario contar con evaluaciones económicas, de efectividad y seguridad de las diferentes alternativas de intervención; para ello los estudios de comparaciones directas tipo ECC «cabeza a cabeza» serían la mejor elección. Sin embargo, este tipo de estudios casi nunca están disponibles debido, principalmente, a su elevado costo. En estas situaciones, se hace necesario contar con otras herramientas de análisis como lo son las CI. Las CI de tratamientos permiten utilizar modelos estadísticos que evalúan el efecto de cada terapia y compararlas entre sí, convirtiéndose en una herramienta complementaria para apoyar la toma de decisiones en salud. Por ejemplo, si una comparación entre las intervenciones A y B es necesaria, pero sólo existe una comparación entre A y C, y una entre B y C, es posible emplear las CI con el fin de determinar estadísticamente cuál sería la mejor intervención. En cuanto a qué tipo de CI usar, los diferentes métodos tienen sus ventajas e inconvenientes. Las RMA y especialmente las CTM parecen ser los mejores métodos cuando se tiene tanto evidencia directa como indirecta. Por último, para las buenas prácticas de investigación, así como para una adecuada toma de decisiones en salud, se debe tener en cuenta que los resultados obtenidos de las CI pueden llevar a erradas interpretaciones de sus resultados dada su naturaleza y ser susceptibles a sesgos y errores. Por lo tanto, se hace necesario conocer las ventajas y limitaciones de las CI, así como comprender adecuadamente los supuestos básicos de ellas y aplicar la más adecuada para cada situación.

## Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## BIBLIOGRAFÍA

1. Jansel JP, Fleurence R, Devine B, Itzler R, Barrerr A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: Report of the ISPOR task force on indirect treatment comparison good research practices: Part 1. *Value Health*. 2011;14:417-28.
2. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. 2005; 351:897-900.
3. Donegan S, Williamson P, Gamble C, Tudor-Smith C. Indirect Comparisons: A review of reporting and methodological quality. *PLoS ONE*. 2010;5:1-11.
4. Schneeweiss S. Developments in post-marketing comparative effectiveness research. *Clin Pharmacol Ther*. 2007;82:143-56.
5. Ioannidis JPA. Indirect comparisons: the mesh and mess of clinical trials. *Lancet*. 2006;368:1470-2.
6. Song F. What is an indirect comparison? What is ...? series. 2009 [en línea; consultado 30 Ago 2012]. Disponible en: [http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What\\_is\\_ind\\_comp.pdf](http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_ind_comp.pdf).
7. Lazo-Langner A, Rodger MA, Barrowman NJ, Ramsay T, Wells PS, Coyle DA. Comparing multiple competing interventions in the absence of randomized trials using clinical risk-benefit analysis. *BMC Med Res Methodol*. 2012;12:1-13.
8. Zuleta JJ. Ronda clínica y epidemiológica: comparaciones indirectas. *IATREIA*. 2010;23:79-83.
9. Sánchez-Pedraza R, Gamboa O, Díaz JA. Modelos empleados para la toma de decisiones en el cuidado de la salud. *Rev Salud Pública*. 2008;19:178-88.
10. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50:683-91.
11. Normand SLT, McNeil BJ. What is evidence? *Stat Med*. 2010;29:1985-7.
12. Castillo-Laborde C. Evaluaciones económicas en el proceso de toma de decisiones en salud: El caso de Inglaterra. *Rev Med Chile*. 2010;138:103-7.
13. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: An introduction to mixed treatment comparisons. *Value Health*. 2008;11:956-64.
14. Song F, Loke Y, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ*. 2009;338:B1147.
15. Hauté Autorité de Santé. Departament of Medecines Assessment. Indirect comparisons. Methods and validity. Francia: CEDEX; 2009.
16. Shöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirect comparisons of therapeutic interventions. *GSM Health Technol Assess*. 2009;5:1-13.
17. Edwards SJ, Arke MJ, Wordsworth S, Borrill J. Indirect comparisons of treatments based on systematic reviews of randomized controlled trials. *Int J Clin Pract*. 2009;63:823-3.

18. Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics Med.* 2002;21:2313-24.
19. Devine EB, Alfonso-Cristancho R, Sullivan SD. Effectiveness of biologic therapies for rheumatoid arthritis: An indirect comparison approach. *Pharmacotherapy.* 2011;31:39-51.
20. Aedo S, Pavlov S, Clavero F. Riesgo relativo y odds ratio. ¿Qué son y cómo se interpretan? *Rev Obstet Ginecol.* 2010;5:51-4.
21. Alfonso-Cristancho R. Métodos de comparaciones mixtas e indirectas: Herramientas para evaluar tecnologías de salud en ausencia de evidencia directa [en línea; consultado 16 Ago 2012]. En: ISPOR Tercera Conferencia Latinoamericana. México 2011 Sep 10. Disponible en: <http://www.ispor.org/conferences/mexico0911/presentations/W8-ALLSLIDES.pdf>
22. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol.* 2011;11:61.
23. Edwards SJ, Borrell J. Network meta-analysis: importance in appropriate trial selection. *Value Health.* 2010;13:681-2.
24. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess.* 2005;9:1-134.
25. Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics.* 2010;28:935-45.
26. Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Capperelli JC, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the ISPOR task force on indirect treatment comparisons good research practices- Part 2. *Value Health.* 2011;14:429-37.
27. Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect evidence: Indirect treatment comparisons in meta-analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009 [en línea; consultado 26 Jul 2012]. Disponible en: [http://www.cadth.ca/media/pdf/H0462\\_itc\\_tr\\_e.pdf](http://www.cadth.ca/media/pdf/H0462_itc_tr_e.pdf).
28. Ortega A, Fuentes F, Alegre EJ, Ventayol P. Comparaciones indirectas. *FarmHosp.* 2012. doi:10.1016/j.farma.2011.06.007.
29. Report of the Indirect Comparisons Working Group to the Pharmaceutical Benefits Advisory Committee: Assessing Indirect Comparisons. ICWG Report [en línea; consultado 22 Jul 2012]. Disponible en: [http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\\$File/ICWG%20Report%20FINAL2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/$File/ICWG%20Report%20FINAL2.pdf).
30. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics.* 2008;26:753-67.
31. Green S, Higgins JPT, Alderson P, Clarke M, Mulrow CD, Oxman AD. *Cochrane handbook for systematic reviews of interventions* 5.1.0. 2011.
32. Bending M, Hutton J, McGrath C. Indirect treatment comparisons and network meta-analysis: A review of manufacturers' submissions to the NICE Single Technology Appraisal Process (STA). 28<sup>th</sup> June 2011. HTAi 8<sup>th</sup> Annual Meeting, Rio de Janeiro, Brazil [en línea; consultado 30 Jul 2012]. Disponible en: [http://php.york.ac.uk/inst/yhec/web/posters/Bending\\_IndirectTreatmentComparison\\_HTAi2011\\_429.pdf](http://php.york.ac.uk/inst/yhec/web/posters/Bending_IndirectTreatmentComparison_HTAi2011_429.pdf).
33. Song F, Harvey I, Lilford R. Adjusted indirect comparisons may be less biased than direct comparisons for evaluating new pharmaceutical interventions. *J Clin Epidemiol.* 2008;61:455-63.
34. O'Regan C, Ghement I, Eyawo O, Guyatt GH, Mills EJ. Incorporating multiple interventions in meta-analysis: an evaluation of the mixed treatment comparison with the adjusted indirect comparison. *Trials.* 2009;10:1-12.
35. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ.* 2011;243. doi:10.1136/bmj.d4909.
36. Khoshdel A, Attia J, Carney SL. Basic concepts in meta-analysis: a primer for clinicians. *Int J Clin Pract.* 2006;60:1287-94.
37. Haidich AB. Meta-analysis in medical research. *Hippokratia.* 2010;14:29-37.
38. Nordmann AJ, Kasenda B, Briel M. Meta-analyses: what they can and cannot do. *Swiss Med Wkly.* 2012;142:w13518.
39. Peura P, Asseburg C, Turunen J, Purmonen T, Martikainen J. Indirect comparison and network meta-analyses: new tools for the assessment of evidence on the relative efficacy of drugs. *Duodecim.* 2011;127:900-10.
40. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol.* 2002;55:86-94.
41. Sutton AJ, Higgins JPT. Recent developments in meta-analysis. *Statist Med.* 2008;27:625-50.
42. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ.* 2009;181:488-93.
43. Schmid CH, Stark PC, Berlin JA, Landais P, Lau J. Meta-regression detected associations between heterogeneous treatment effects and study-level, but no patient-level, factors. *J Clin Epidemiol.* 2004;65:683-97.
44. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. *Statist Med.* 2012. doi:10.1002/sim.5442.
45. Baumgartner RW. Network meta-analysis of antiplatelet treatments for secondary stroke prevention. *Eur Heart J.* 2008;29:1082-3.
46. Li T, Puhon MA, Vedula SS, Singh S, Dickersin K; Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med.* 2011;27:79.
47. Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol.* 2010;10:54.
48. Del Giovane C, Vacchi L, Mavridis D, Filippini G, Salanti G. Network meta-analysis models to account for variability in treatment definitions: application to dose effects. *Statist Med.* 2012. doi:10.1002/sim.5512.
49. Ades AE, Mavranzouli I, Dias S, Welton NJ, Whittington C, Kendall T. Network meta-analysis with competing risk outcomes. *Value Health.* 2010;13:976-83.
50. Singh JA, Wells GA, Christensen R, Tanjong Ghogomu E, Maxwell L, McDonald JK, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. *Cochrane Database Syst Rev.* 2011;16:CD008794.
51. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Statist Med.* 2002;21:1559-73.
52. Berry SM, Ishak KJ, Luce BR, Berry DA. Bayesian meta-analyses for comparative effectiveness and informing coverage decisions. *Med Care.* 2010;48:S137-S144.
53. The Cochrane Collaboration open learning material. Investigating sources of heterogeneity. En: *Diversity and*

- heterogeneity [en línea; consultado 30 Jul 2012]. Disponible en: <http://www.cochrane-net.org/openlearning/html/mod13-5.htm>.
54. Donegan S, Williamson P, D'Alessandro U, Smith CT. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. *Statist Med*. 2012. doi:10.1002/sim.51.
  55. Turkstra E, Ng SK, Scuffham PA. Review. A mixed treatment comparison of the short-term efficacy of biologic disease modifying anti-rheumatic drugs in established rheumatoid arthritis. *Curr Med Res Opin*. 2011;27:1885-97.
  56. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Statist Med*. 2010;29:932-44.
  57. Thorlund K, Mills E. Stability of additive treatment effects in multiple treatment comparison meta-analysis: a simulation study. *Clinical Epidemiology*. 2012;4:75-85.
  58. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statist Med*. 2004;23:3105-24.
  59. Smith T, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Statist Med*. 2005;14:2685-99.
  60. Berry DA. Bayesian clinical trials. *Nat Rev Drug Discov*. 2006;5:27-36.
  61. Berry DA. Is the Bayesian approach ready for prime time? Yes! *Stroke*. 2005;36:1621-2.
  62. Donnan GA, Davis SM, Ludbrook J. The Bayesian principle. Can we adapt? *Stroke*. 2005;36:1623-4.
  63. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med*. 1999;18:2693-708.
  64. Morton SC, Adams JL, Suttrop MJ, Shekelle PG. Meta-regression approaches: What, why, when, and how? Technical review 8. Rockville: Agency for Healthcare Research and Quality; 2004.
  65. Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Statist Med*. 2007;26:1237-54.
  66. Berry DA. Introduction to Bayesian methods III. Use and interpretation of Bayesian tools in design and analysis. *Clin Trials*. 2005;2:295-300.
  67. Gönen M. Bayesian clinical trials: no more excuses. *Clin Trials*. 2009;6:203-4.
  68. Raiffa H, Schlaifer R. *Applied statistical decision theory*. Boston: Division of Research, Graduate School of Business Administration, Harvard University; 1961.
  69. Berry DA. Bayesian approaches for comparative effectiveness research. *Clin Trials*. 2012; 9:37-47.
  70. Donnan GA, Davis SM, Ludbrook J. The Bayesian principle. Can we adapt? *Stroke*. 2005;36:1623-4.
  71. Biswas S, Liu DD, Lee JJ, Berry DA. Bayesian clinical trials at the University of Texas M.D. Anderson Cancer Center. *Clinical Trials*. 2009;6:205-16.
  72. Hawkins N, Scott DA, Wood BS. How far do you go? Efficient searching for indirect evidence. *Med Decis Making*. 2009;29:273-81.
  73. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statist Med* 2002;21:1539-58.
  74. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez Olivo MA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. *Cochrane Database Syst Rev*. 2009;(4):CD007848:1-57.
  75. Singh JA, Beg S, Lopez-Olivo MA. Tocilizumab for rheumatoid arthritis. *Cochrane Database Syst Rev*. 2010;(7):CD008331:1-68.
  76. Castillo MA, Ubago R, Flores S, Beltrán C. Terapias biológicas en el tratamiento de la artritis reumatoide. Eficacia y seguridad comparada entre diferentes agentes biológicos. Sevilla: Agencia de Evaluación de Tecnologías Sanitarias; 2011.
  77. Favalli EG, Pregnotato F, Biggioggero M, Meroni PL. The role of biologic agents in damage progression in rheumatoid arthritis: indirect comparison of data coming from randomized clinical trials. *Ther Adv Musculoskelet Dis*. 2012;4:213-23.
  78. Singh JA, Christensen R, Wells GA, Suárez- Almazor ME, Buchbinder R, López-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *CMAJ*. 2009. doi:10.1503/cmaj.091391.
  79. Krüger K, Hellmich B, Rubbert-Roth A, Müller-Ladner U, Kielhorn A. Is it possible to compare different studies on efficacy of biologicals in patients with rheumatoid arthritis? Mixed treatment comparison as a new tool for indirect comparison of clinical studies. *Z Rheumatol*. 2011;70:517-24.