



EVALUACIÓN DE MEDICAMENTOS CON EL NUEVO PROGRAMA MADRE 4.0. NOVEDADES EN ANÁLISIS DE SUPERVIVENCIA, COMPARACIONES INDIRECTAS Y EVALUACIÓN ECONÓMICA

GENESIS



Sociedad Española de
Farmacia Hospitalaria





COMPARACIONES INDIRECTAS

Emilio Jesús Alegre del Rey
M^a Dolores Fraga Fuentes

GENESIS



Sociedad Española de
Farmacia Hospitalaria



10:00 – 10:10 Presentación de la jornada

Análisis de supervivencia

10:10 – 10:50 Análisis de supervivencia en oncología: variables relacionadas (SG, SLP, SLE, HR) e interpretación de curvas de supervivencia (media vs mediana). A. Ortega

10:50 – 11:30 Comparación de curvas de supervivencia y cálculo de NNT. E.Lopez Briz

11:30-12:00 Café



Comparaciones Indirectas

12:00 – 12:45 Lista de comprobación para el análisis crítico de CI. MD. Fraga

12:45 – 13:00 Comparación indirecta sencilla con la calculadora de Wells. E. Alegre

13:00 – 13:30 Evaluación y posicionamiento de medicamentos como ATE. E. Alegre.

13:30- 15:00 Comida



Evaluación económica

15:00-15:30 Adaptación de evaluaciones económicas a nuestro entorno. A. Ortega

15:30-16:00 Aplicabilidad de los datos de ensayo clínico en las evaluaciones económicas. R. Marin

16:00-16:30 Análisis de sensibilidad. A. Ortega

16:30-17:00 Cálculo de impacto presupuestario. R. Marin

17:00-18:00 Discusión y conclusiones





Nueva lista de comprobación para el análisis crítico de comparaciones indirectas

Ma Dolores Fraga Fuentes

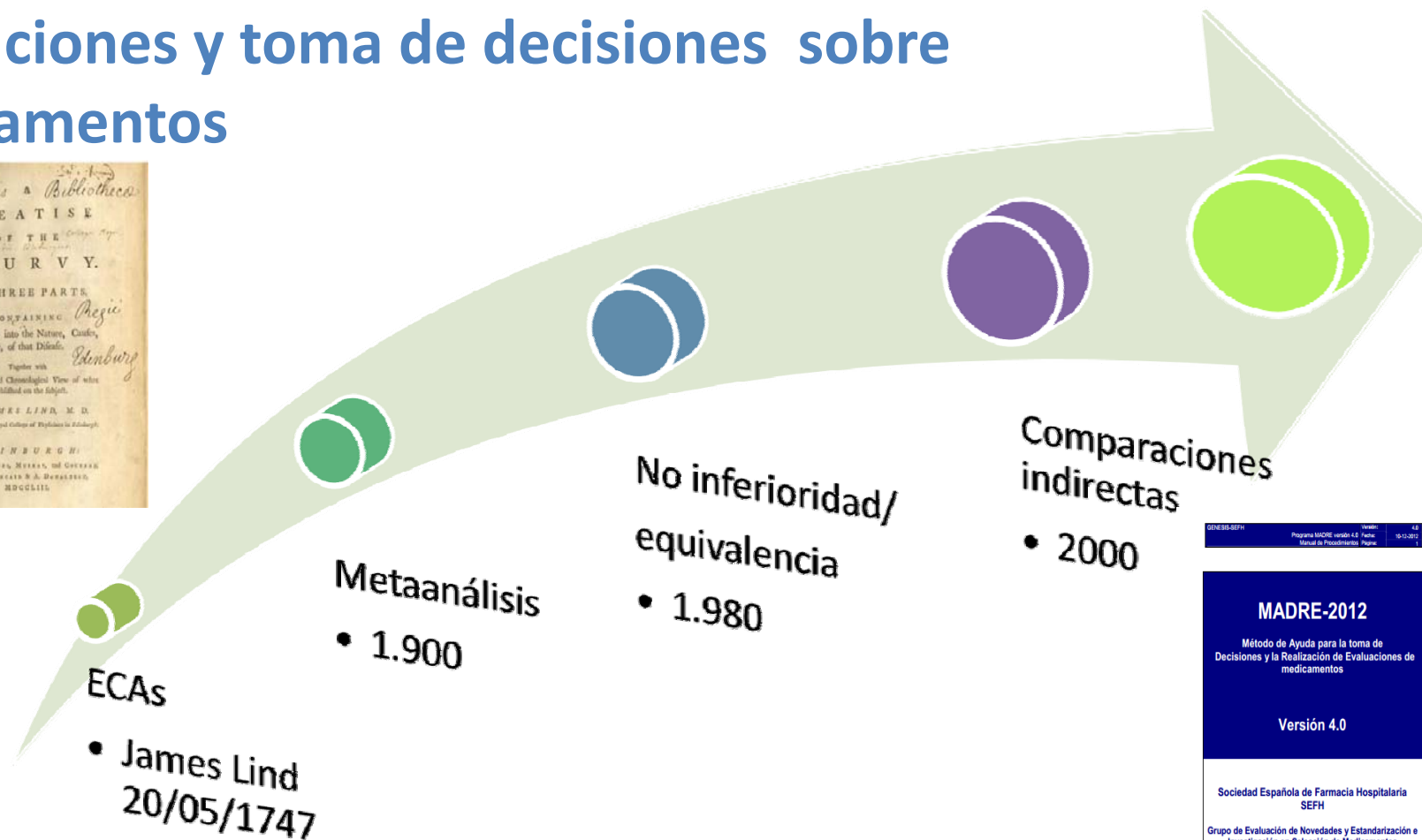
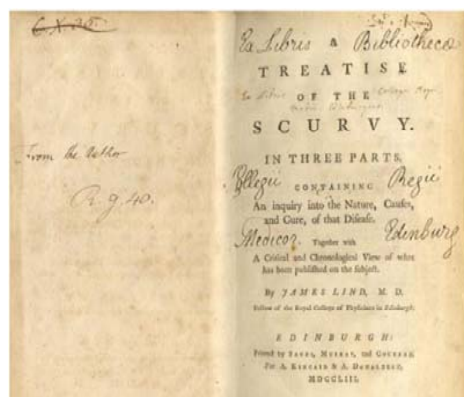
GENESIS



Sociedad Española de
Farmacia Hospitalaria



Dónde estamos: Crónica de un proceso.....en evaluaciones y toma de decisiones sobre medicamentos



GENESIS-00711
Programa MADRE versión 4.0
Manual de Procedimientos Regla

MADRE-2012

Método de Ayuda para la toma de Decisiones y la Realización de Evaluaciones de medicamentos

Versión 4.0

Sociedad Española de Farmacia Hospitalaria
SEFH

Grupo de Evaluación de Novedades y Estandarización e Investigación en Selección de Medicamentos
GENESIS

GENESIS



....En la evaluación de medicamentos

EFICACIA Y SEGURIDAD

Incremento en opciones de tratamiento que requieren efectividad comparativa

Escasez de ensayos comparativos directos

Metaanálisis herramienta estadística que proporciona un resumen estimado de los efectos de los tratamientos, pero la **limitación** de los metaanálisis clásicos es que **sólo comparan dos intervenciones al mismo tiempo**

Necesidad de toma de decisiones ante varias alternativas terapéuticas



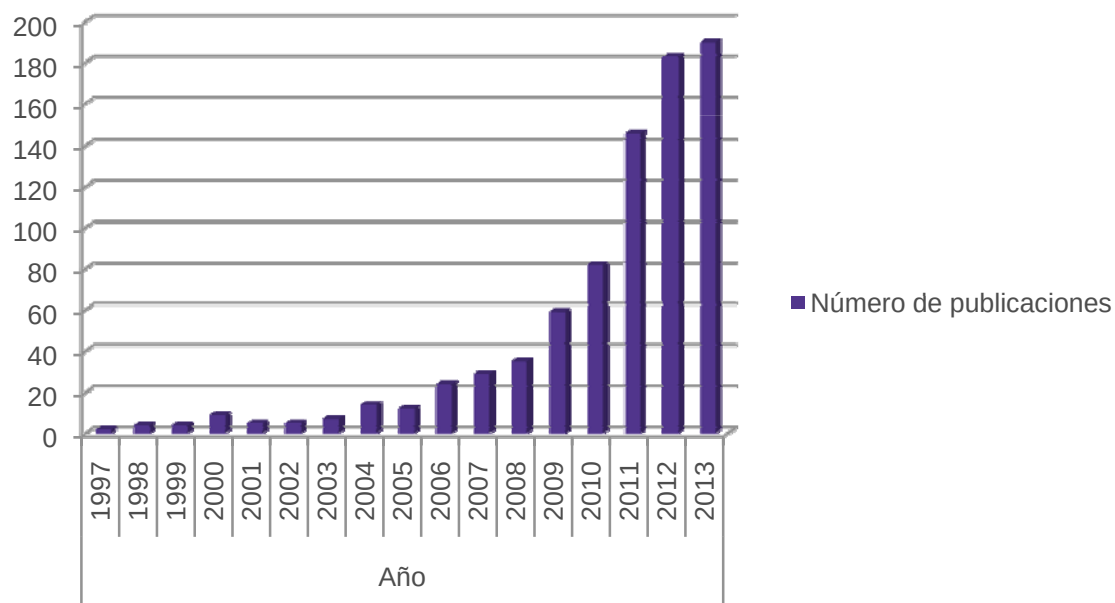
Comparaciones indirectas





...En la evaluación de medicamentos

Artículos publicados sobre comparaciones indirectas en PubMed



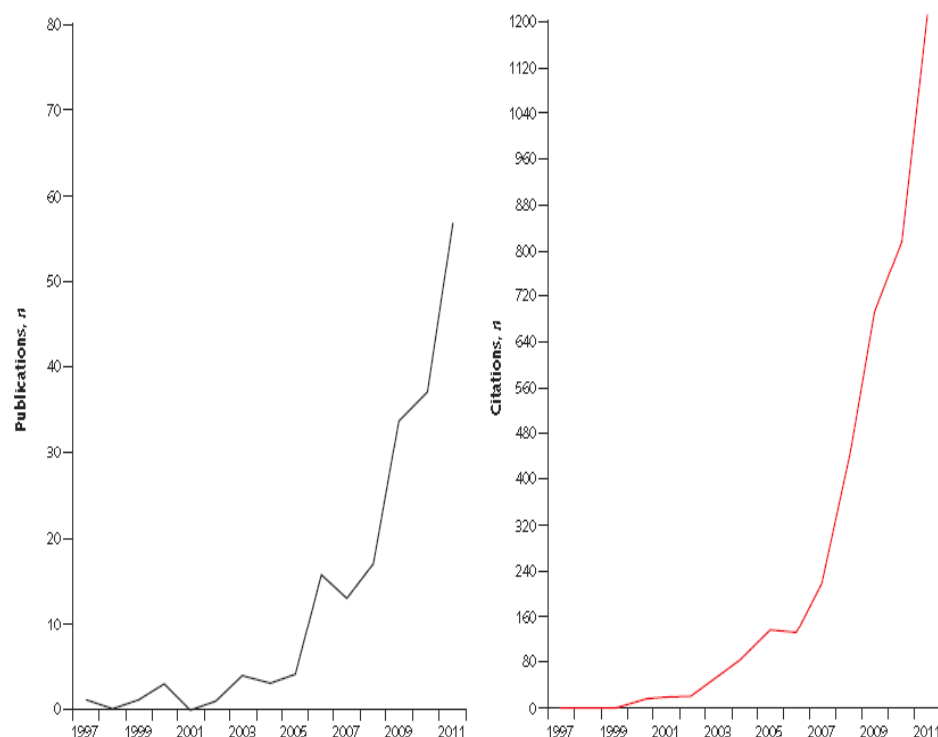
Pueden ser
más por la
falta de
consenso en la
terminología





...En la evaluación de medicamentos

Number of network meta-analyses published in the scientific literature and their citations since 1997.



Journal of Clinical Epidemiology 67 (2014) 138–143

Journal of
Clinical
Epidemiology

REVIEW ARTICLES

Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009

Andrew W. Lee*

Department of Continuing Education, University of Oxford, Rewley House, 1, Wellington Square, Oxford OX1 2JA

Accepted 19 July 2013; Published online 3 October 2013

Abstract

Objectives: To identify and summarize published systematic reviews that report results of meta-analyses that combined direct and indirect comparisons.

Study Design and Setting: Narrative review of mixed treatment comparisons (MTCs) reported in systematic reviews of health interventions. MEDLINE, MEDLINE In-Process, Embase, CINAHL, DARE, the Cochrane Database of Systematic Reviews, and SIGLE were searched for reviews published up to June 2012 in which a meta-analysis had been conducted that combined direct and indirect comparisons among more than two interventions.

Results: Reviews reporting MTCs are difficult to identify when searching major databases. These databases offer no way to identify MTCs, and authors use various names when reporting them. Of the 201 eligible reviews identified, more than three-quarters had been published in full. MTC methods have been used to study a wide range of clinical topics. The reported use of these methods has increased rapidly since 2009, and results from MTCs are commonly used in health policy decisions, through the evidence considered in health technology assessments.

Conclusions: In view of the increasing use of MTCs, indexing of this study type in databases and consensus on terminology and standards for conduct and reporting would be timely.

Keywords: Research design/ends, Mixed treatment comparison, Evidence-based medicine, Review literature, Indirect comparison, Meta-analysis



Opinión de las agencias

CI solo si
no existen
ensayos
H2H

Menos peso
que los ECAS
H2H

No
deberían
utilizarse
datos
observacio
nales

CI sólo
válidos si la
contribución
de los ECAS
es similar



UK National Health Service (NHS) Health Technology Assessment (HTA) Programme
US Agency for Healthcare Research and Quality (AHRQ)
Canadian Agency for Drugs and Technologies in Health (CADTH)
Australian Pharmaceutical Benefits Advisory Committee (PBAC) and PBAC Working Group
German Institute of Medical Documentation and Information



CRITICA

Constructiva



La CI ajustada respeta la asignación aleatoria pero no es una evidencia aleatorizada:

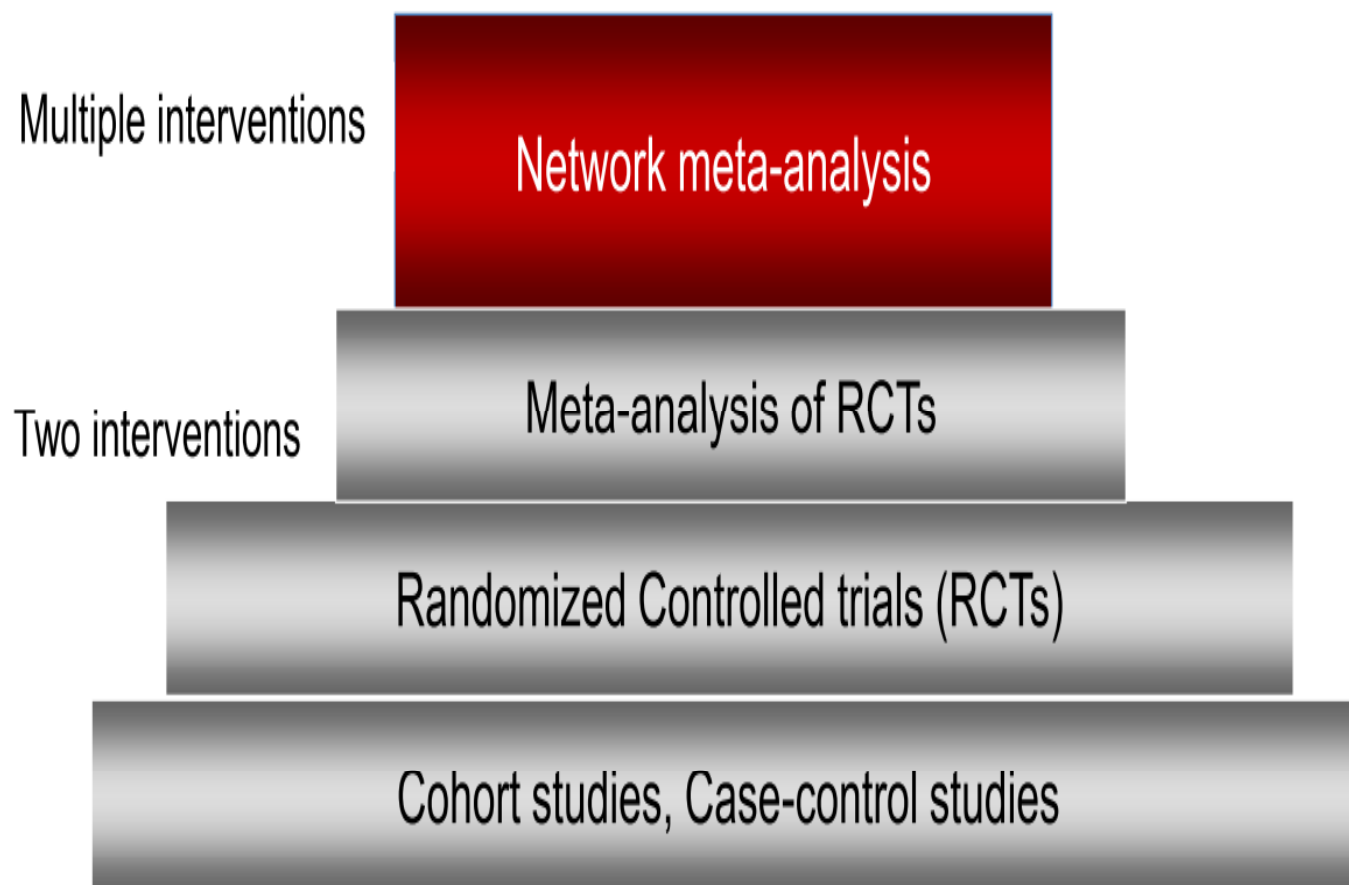
- Las comparaciones de tratamientos no están asignados aleatoriamente entre estudios
- La CI es un tipo especial de regresión
- ¿Es una evidencia observacional?

¿La evidencia directa es preferible a la indirecta?

¿Debe utilizarse solo en caso de ausencia de evidencia directa?



¿Cambiarán los niveles de evidencia?



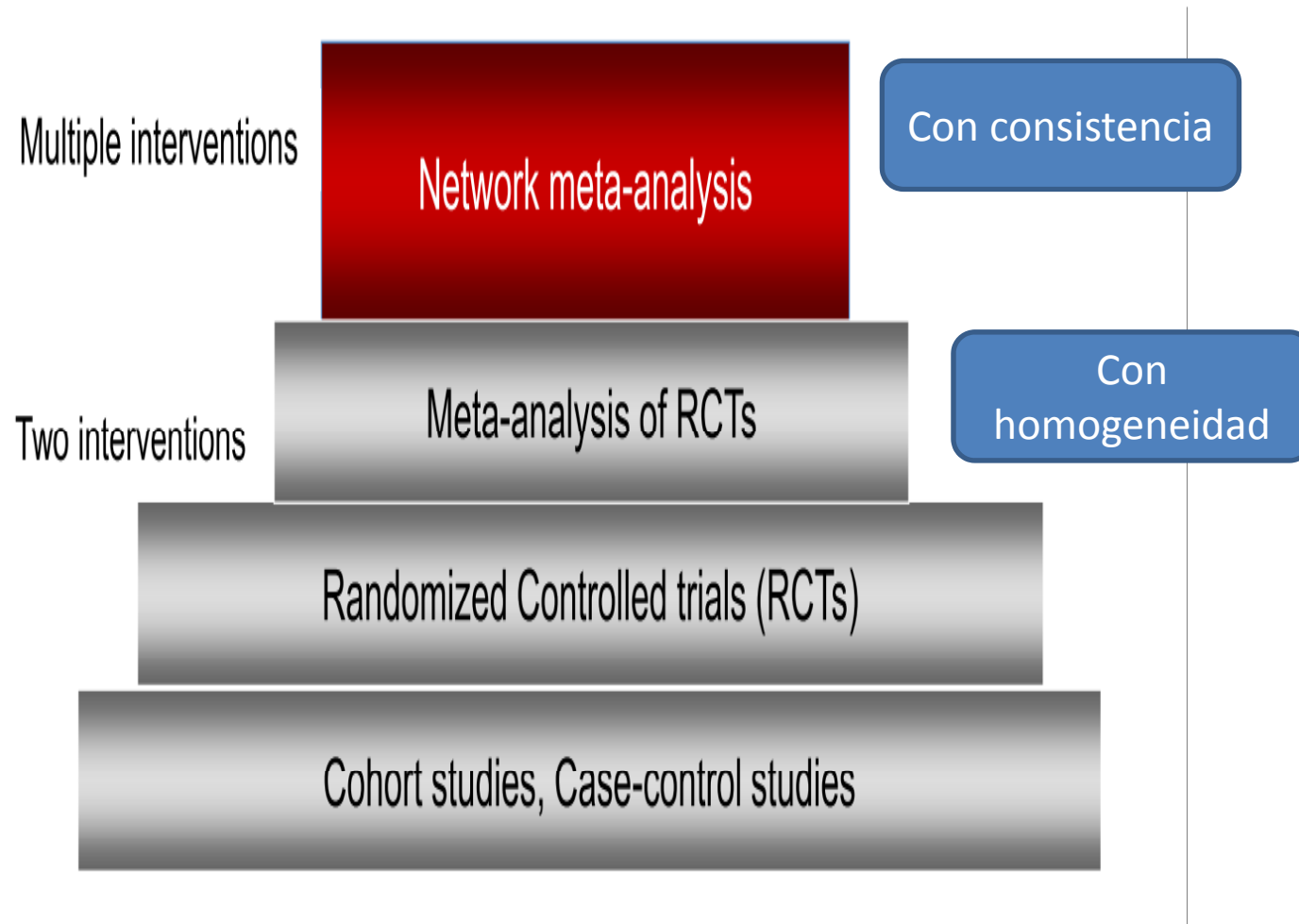
GENESIS



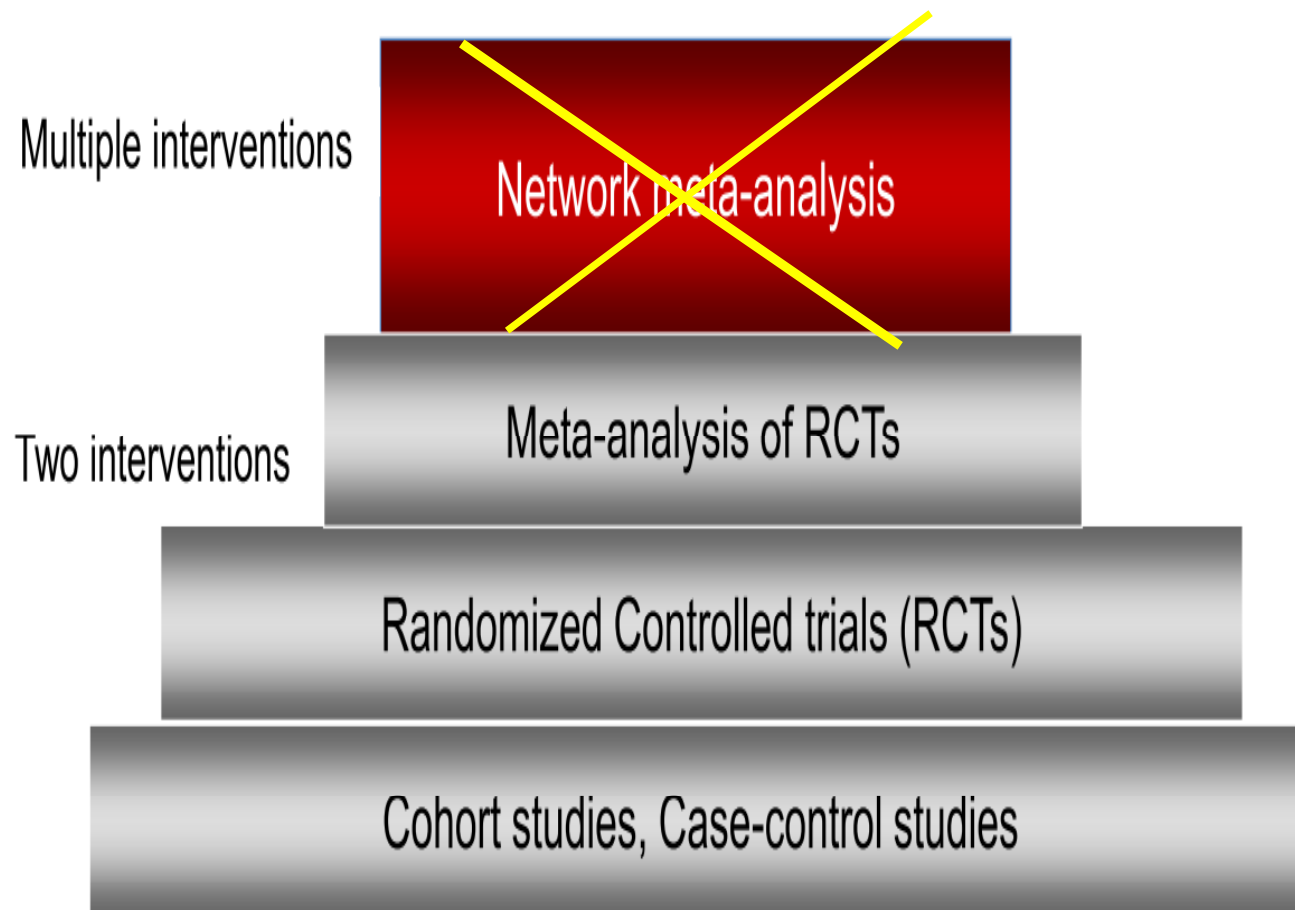
Sociedad Española de
Farmacia Hospitalaria



¿Cambiarán los niveles de evidencia?



¿Cambiarán los niveles de evidencia?



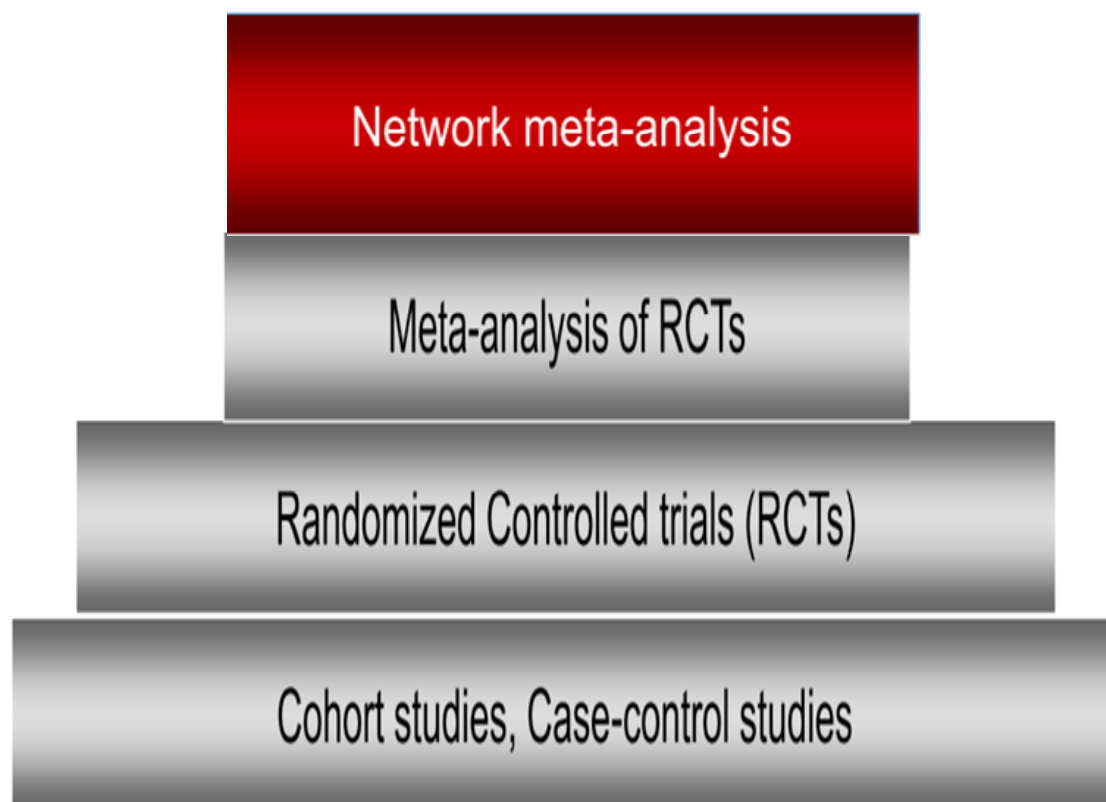
GENESIS



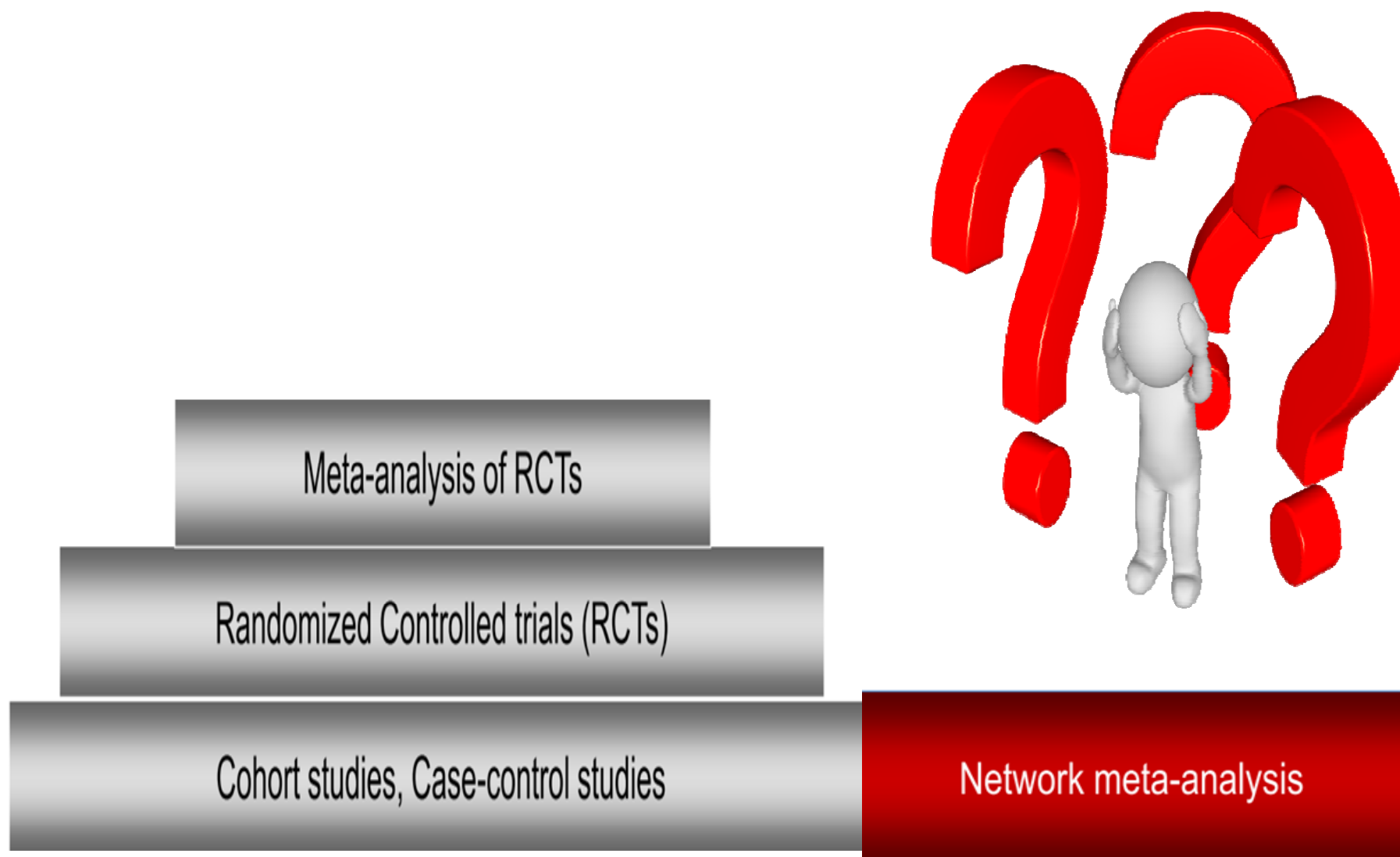
Sociedad Española de
Farmacia Hospitalaria



¿Cambiarán los niveles de evidencia?



¿Cambiarán los niveles de evidencia?



¿Cambiarán los niveles de evidencia?

RESEARCH METHODS & REPORTING

A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis

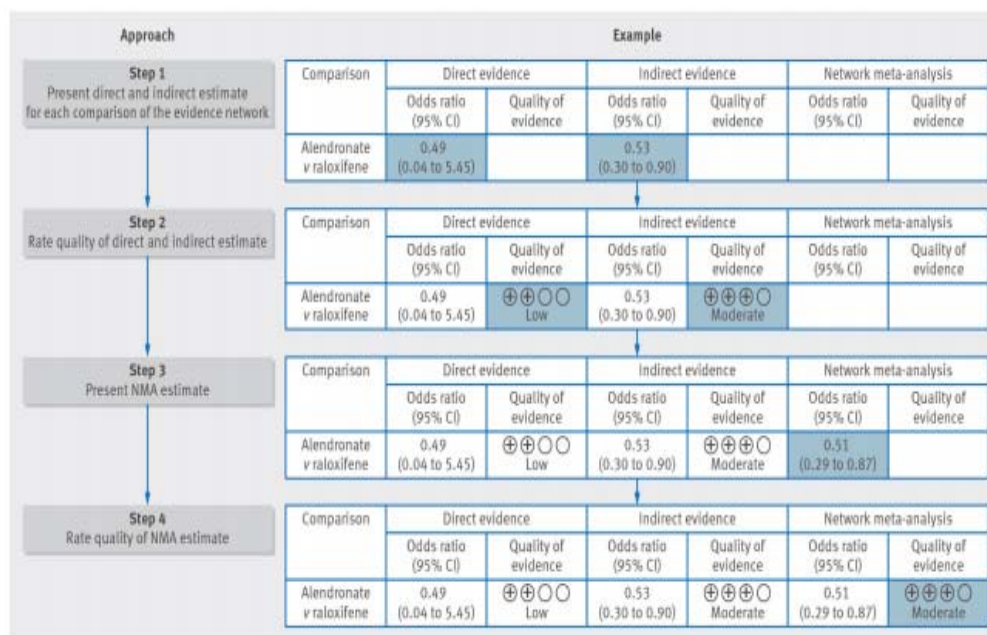
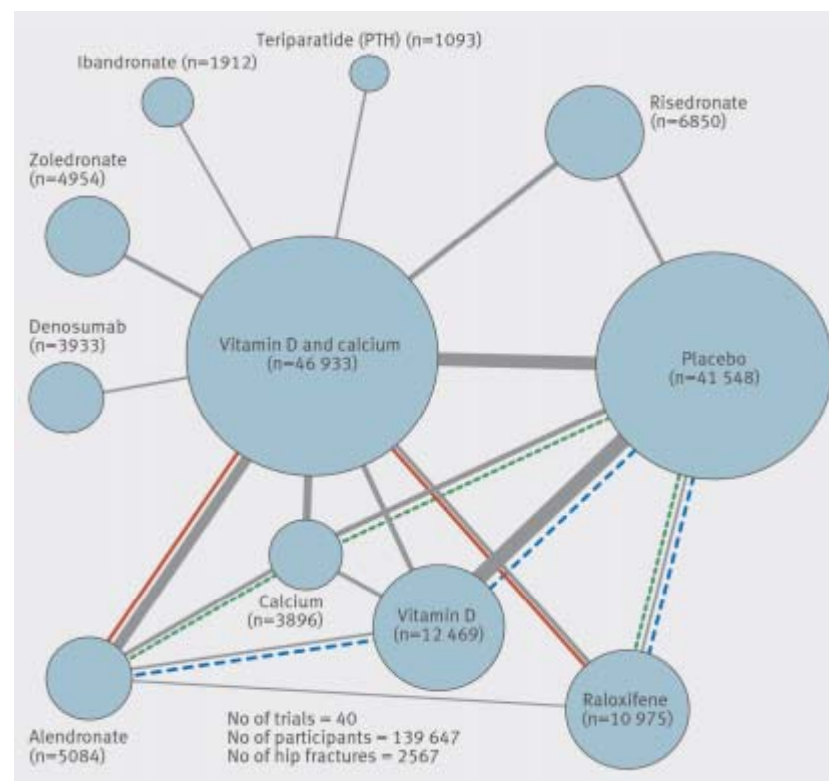


Fig 1 Approach for rating the quality of network meta-analysis (NMA) estimates



5.3 Revisiones sistemáticas publicadas, comparaciones indirectas y sus conclusiones,

5.3.a Revisiones sistemáticas publicadas,

5.3.b Comparaciones indirectas,

5.3.b.1 Comparaciones indirectas publicadas,

5.3.b.2 Comparaciones indirectas de elaboración propia,

5.4 Evaluación de fuentes secundarias,

5.4.1 Guías de Práctica clínica

5.4.2 Evaluaciones previas por organismos independientes

5.4.3 Opiniones de expertos

5.4.4 Otras fuentes.

6. EVALUACIÓN DE LA SEGURIDAD, 74

6.1.a Descripción de la búsqueda bibliográfica,

6.1.b Descripción de los efectos adversos más significativos (por su frecuencia o gravedad),

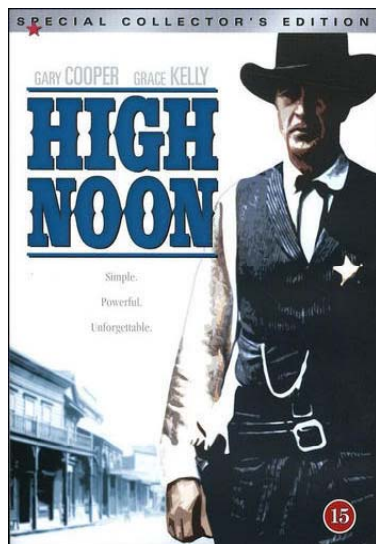
6.2. Seguridad. Ensayos Clínicos comparativos.

6.3. Fuentes secundarias sobre seguridad,

6.4. Precauciones de empleo en casos especiales,



Comparaciones indirectas: muchos nombres



Comparaciones
indirectas ajustadas

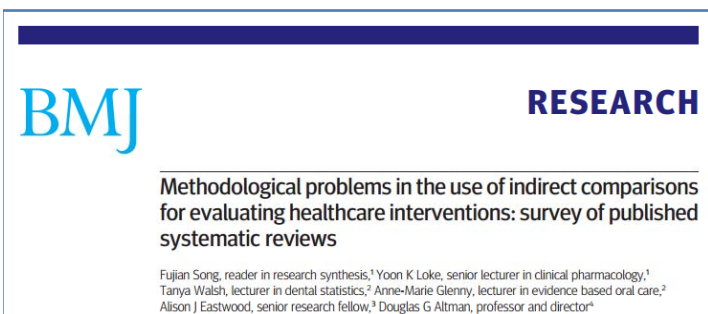
- **Comparaciones mixtas**

~~Comparaciones
indirectas naïve~~

- **Metanálisis en red (Net-work meta-analysis)**

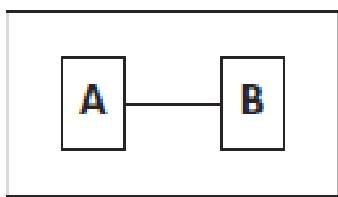
Metanálisis con comparaciones múltiples (multiple-treatments meta-analysis)

Metanálisis con comparaciones mixtas (mixed-treatments comparisons meta-analysis)





Tipos-Definiciones



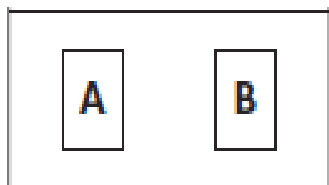
Comparación directa entre dos tratamientos

Comparados mediante:

Ensayo clínico

Varios ensayos clínicos

Un metaanálisis



Comparación indirecta naive o no ajustada

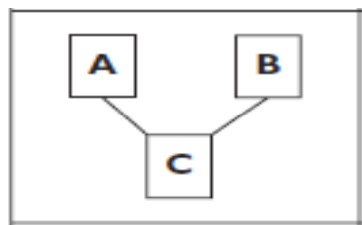
Se comparan directamente los resultados de los brazos individuales de varios estudios como si se tratara del mismo estudio.

NO ADECUADO





Tipos-Definiciones



Comparación indirecta ajustada

Se comparan los efectos relativos de los tratamientos

A y B no están comparados directamente

Ambos están comparados con un comparador común C

A partir de ECAs que valoren por separado la eficacia de dos tratamientos de interés (A y B) frente a un comparador o control común C (A vs C y B vs C), se puede comparar de manera indirecta la eficacia de A vs B, en este caso se **preserva la asignación aleatoria original de los pacientes incluidos en los diferentes ECAs**. La validez de los resultados de este método implica asumir que el efecto de ambos tratamientos es coherente entre los ECAs de ambos grupos, así los resultados de A y B vs C debería ser intercambiables.

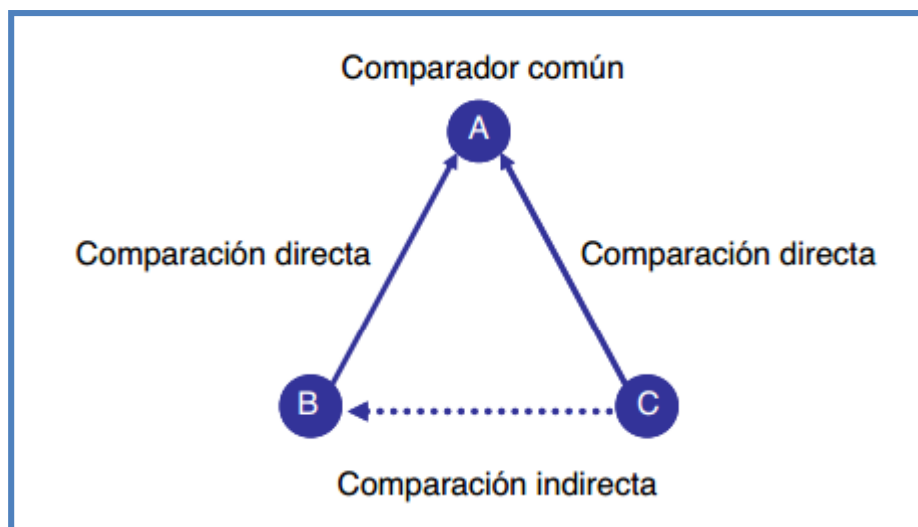




Si sabemos cuánto más alto es B que A y cuánto más alto es C que A, sabemos cuánto más alto es B en comparación con C

Para cualquier pareja B y C

Ventaja de C sobre B = Ventaja de C sobre A – Ventaja de B sobre A



$$\theta_{BC,indirecta} = \theta_{AC,directa} - \theta_{AB,directa}$$

$$V[\theta_{BC,indirecta}] = V[\theta_{AC,directa}] + V[\theta_{AB,directa}]$$

$$EE[\theta_{BC,indirecta}] = \sqrt{V[\theta_{BC,indirecta}]}$$

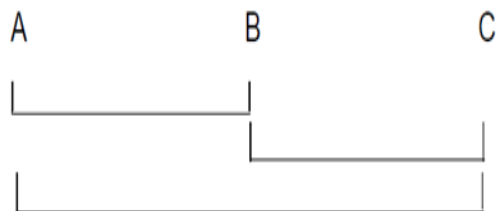
$$IC\ 95\%: \theta_{BC,indirecta} \pm 1,96 \times EE[\theta_{BC,indirecta}]$$

Donde:

$$V[\theta_{AC,directa}] = ((IC_{superior} - IC_{inferior})/3,92))^2$$

$$V[\theta_{AB,directa}] = ((IC_{superior} - IC_{inferior})/3,92))^2$$





Ejemplos

$$\ln(OR')_{AB} = \ln(OR)_{AC} - \ln(OR)_{BC}$$

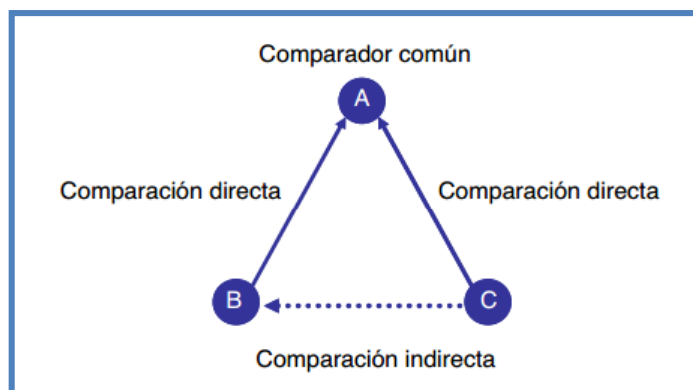
(Bucher et al J Clin Epidemiol 1997)

Comparación	n/N	n/N	RR	IC
A vs C	314/400	391/403	0,95	0,87-0,93
B vs C	107/216	250/260	0,39	0,33-0,46

A imipramina, B alarma, C no tto

$$\begin{aligned} \ln(RR)_{AC} &= -0,051 \\ \ln(RR)_{BC} &= -0,941 \\ \ln RR_{AB} &= \ln RR_{BC} - \ln RR_{AC} = -0,941 - (-0,051) = -0,89 \end{aligned}$$

$$\begin{aligned} RR \text{ indirecto } AB &= \exp(\ln RR_{AB}) = 0,41 \\ \text{Var}(\ln RR_{AB}) &= \text{Var}(\ln RR_{BC}) + \text{Var}(\ln RR_{AC}) = 0,008, EE = 0,09 \\ \ln RR_{AB} \pm 1,96 EE(\ln RR_{AB}) & \quad IC_{95\%} 0,34-0,49 \end{aligned}$$



Para OR, RR, HR:

$$OR_{BC} = OR_{AC} / OR_{AB}$$

$$\ln OR_{BC} = \ln OR_{AC} - \ln OR_{AB}$$

$$\text{Var}(\ln OR_{BC}) = \text{Var}(\ln OR_{AC}) + \text{Var}(\ln OR_{AB})$$

$$IC_{95\%} \ln OR_{BC} \pm 1,96 EE(\ln OR_{BC})$$



Indirect Treatment Comparisons

Effect measure: ☒ Relative Risk (RR) ☐ Odds Ratio (OR) ☐ Risk Difference (RD) ☐ Mean Difference (MD) ☐ Hazard Ratio (HR)

Number of Treatments:

	Estimate	95% LCL	95% UCL	Reverse
(1,2)				☐
(2,3)				☐
(3,4)				☐
(4,5)				☐
(5,6)				☐
(6,7)				☐
(7,8)				☐
(8,9)				☐
(9,10)				☐

Calculate

Indirect Estimate: Treatments (1,10)

Effect measure:

Estimate:

95% confidence interval: LCL UCL

Test of association:

Clear Save Open Exit

<http://gruposdetrabajo.sefh.es/genesis/genesis/Enlaces/Calculadoras.htm>

	A vs C	B vs C
Estimador (RR) directo:	0,39	0,95
Varianza ln estimador directo:	-0,0600	-0,9400
Estimador (RR) ajustado indirecto A vs B:	0,41	#NUM!
IC 95%	#NUM!	#NUM!
Z	#NUM!	#NUM!
P	#NUM!	#NUM!

- 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.

	A vs C	B vs C
Estimador (RR) directo:	0,39	0,95
Limite inferior IC 95% estimador directo:	0,33	0,87
Limite superior IC 95% estimador directo:	0,46	0,92
Estimador (RR) ajustado indirecto A vs B:	0,41	0,35 a 0,49
IC 95%		
Z		-10,36
P		0,00000

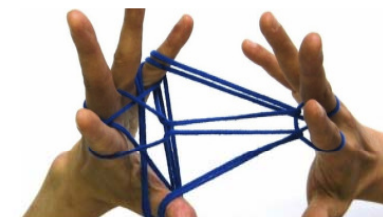
- Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003;326:219.

	Directo A vs B	Indirecto A vs B
Estimador (RR):		
Limite inferior IC 95% estimador:		
Limite superior IC 95% estimador:		

Comparación indirecta ajustada RR DerSimonian-Laird

Comparaciones Indirectas y Metaanálisis en Red

Home
Curso
Galería
Profesores
Publicaciones
Software
Contacto

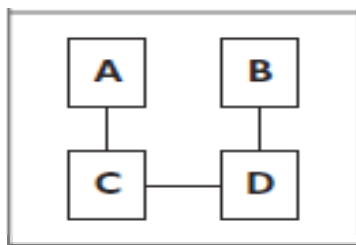


Software > Comparaciones indirectas y mixtas con Excel





Tipos-Definiciones



Comparación indirecta múltiple ajustada

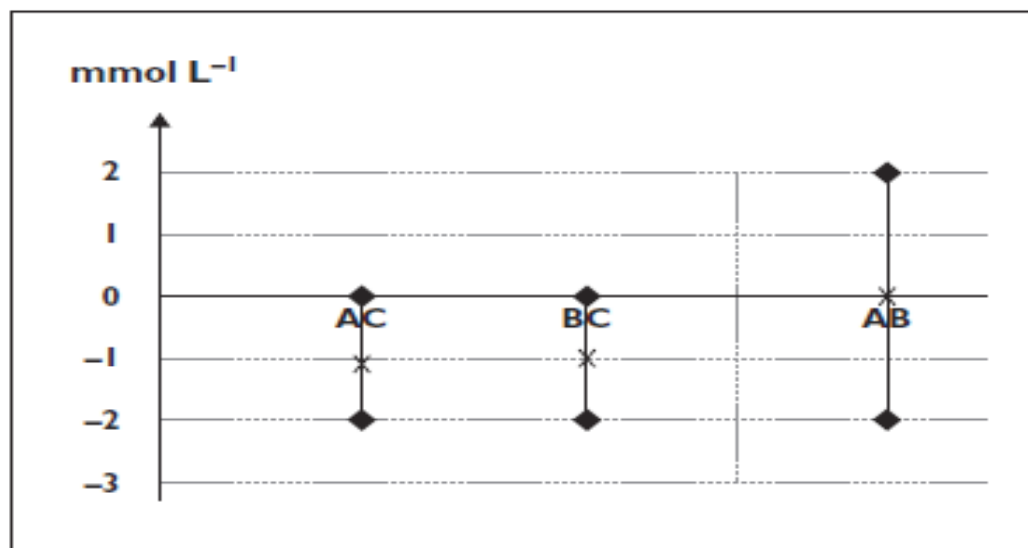
Cuando no se puede identificar un comparador común entre dos tratamientos, se puede construir una serie donde los dos medicamentos están enlazados indirectamente vía dos o más comparadores

	Sitagliptin vs. placebo	Exenatide vs. placebo	Insulin vs. exenatide
Results of the head-to-head clinical trials: (absolute change in HbA1c)	0.57 (95% CI: 0.45, 0.72)	0.52 (95% CI: 0.32, 0.92)	0.40 (95% CI: 0.27, 0.60)
Adjusted indirect comparison: sitagliptin vs exenatide using placebo as common comparator	0.30 (95% CI: 0.17, 0.53)		
Adjusted indirect comparison: sitagliptin vs insulin using exenatide as common comparator	0.74 (95% CI: 0.40, 1.37)		





Tipos-Definiciones

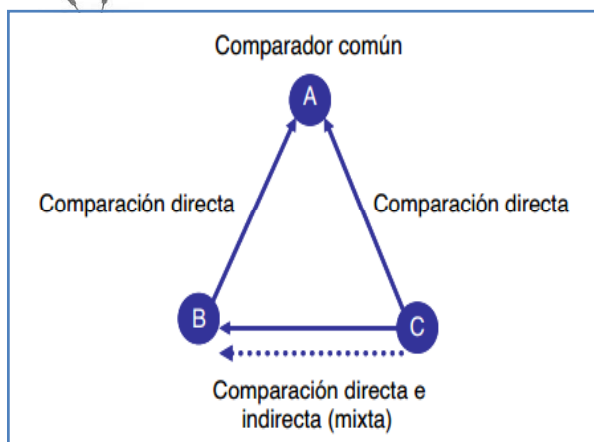


Principal desventaja, la incertidumbre, debido a que las incertidumbres estadísticas de los estudios de los componentes de la comparación se suman





Tipos-Definiciones



Comparación mixta (combinan evidencia directa e indirecta)

Metaanálisis en red (Network meta-analysis) también conocidos como metaanálisis con comparaciones múltiples (MTM-multiple-treatments meta-analysis) o metaanálisis con comparaciones mixtas (MTC-mixed-treatments comparisons meta-analysis)

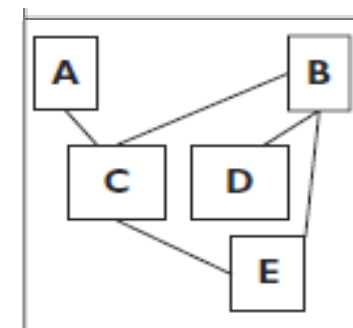
Cuando existen más de tres alternativas y más estudios

Requieren programas complejos. Generalmente utilizan métodos bayesianos

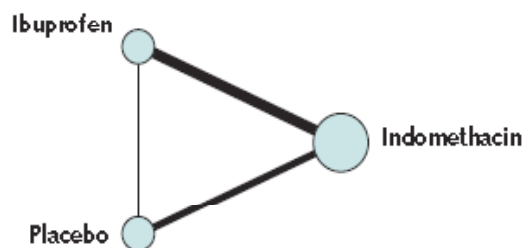
Los resultados de las comparaciones vienen con un intervalo de credibilidad del 95%

Permiten obtener estimaciones de los efectos relativos de los distintos tratamientos considerando toda la red completa de estudios disponibles

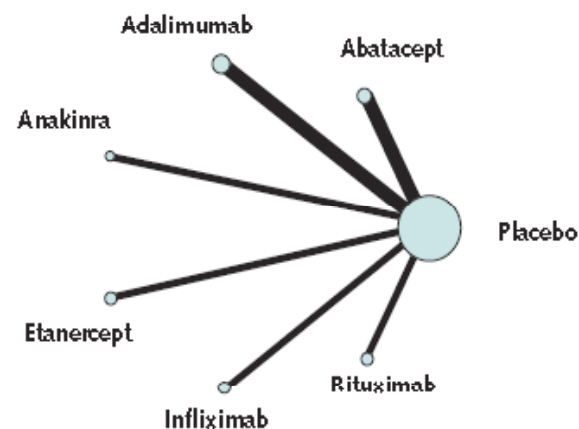
Un metaanálisis en red presenta información procedente de las CD, CI y mixtas establecidas entre un conjunto de alter



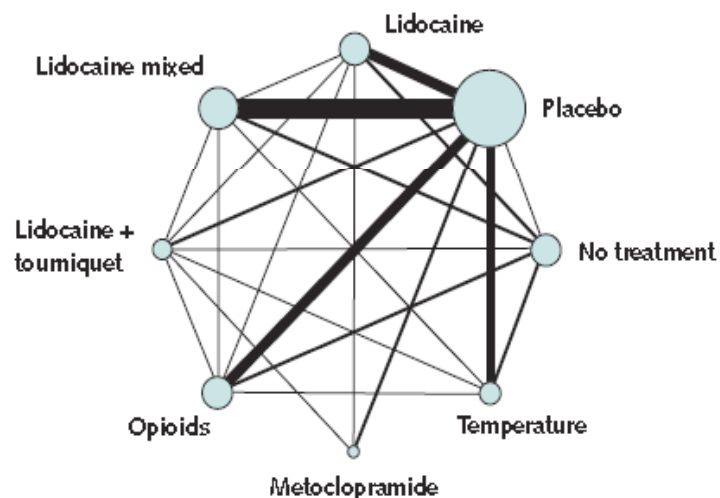
Punto clave para comprender la red de tratamientos es la evaluación de su geometría



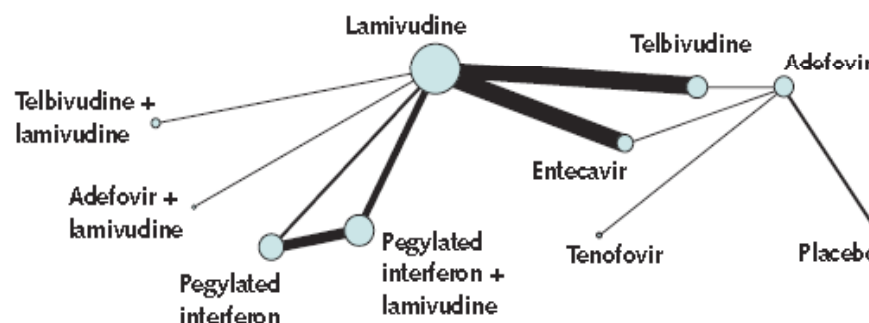
A. Single closed loop



B. Star network

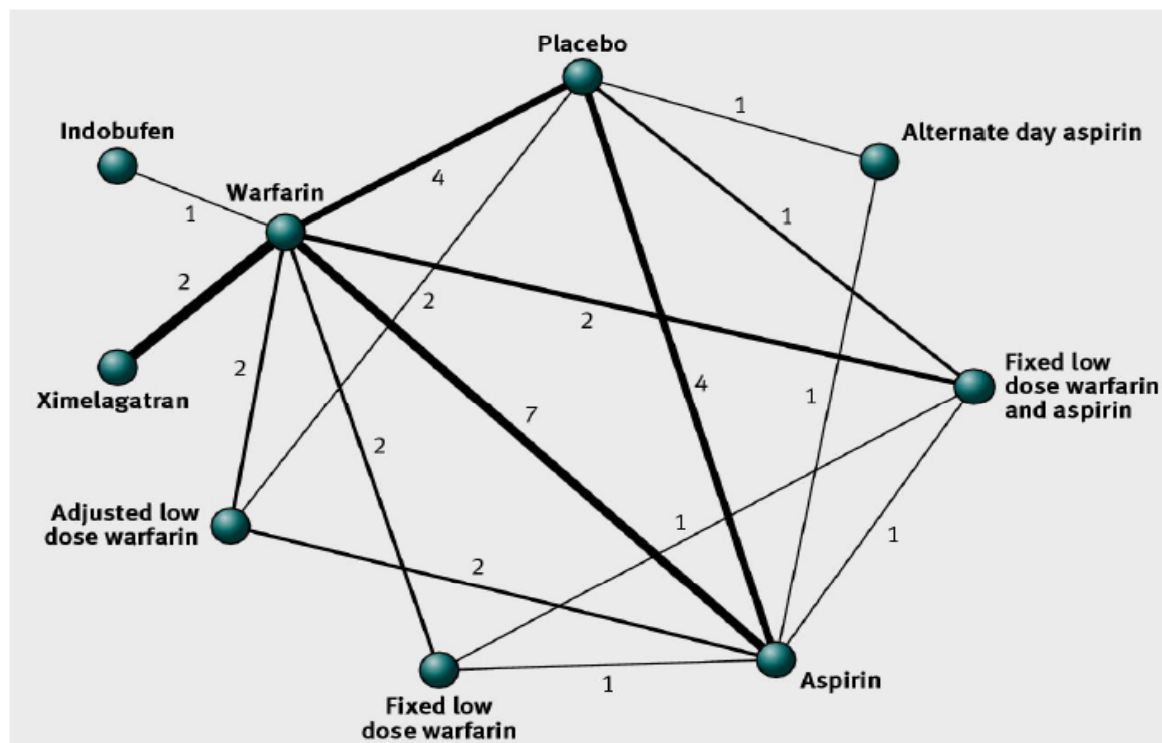


C. Connected network



D. Complex network





Cooper NJ et al. Ann Intern Med 2008;148:544-53

34 parejas de comparaciones

Warfarina (n=20), aspirina (n=16) y placebo (n=12)

Comparación más común (n=7) warfarina y aspirina

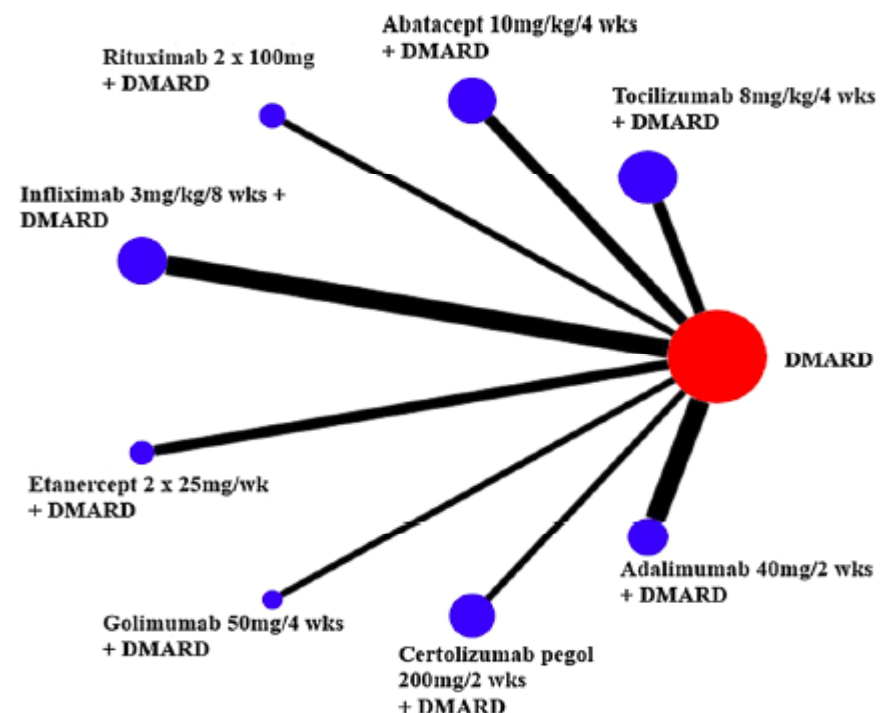
4 comparaciones warfarina con placebo y 4 de aspirina con placebo

Posibilidad de 45 parejas de comparaciones

16 comparaciones directas, pero 6 con 1 ensayo



La metodología del metaanálisis en red no se limita a la estimación numérica de las distintas comparaciones, también explora sesgos y permite proponer una jerarquía de intervenciones

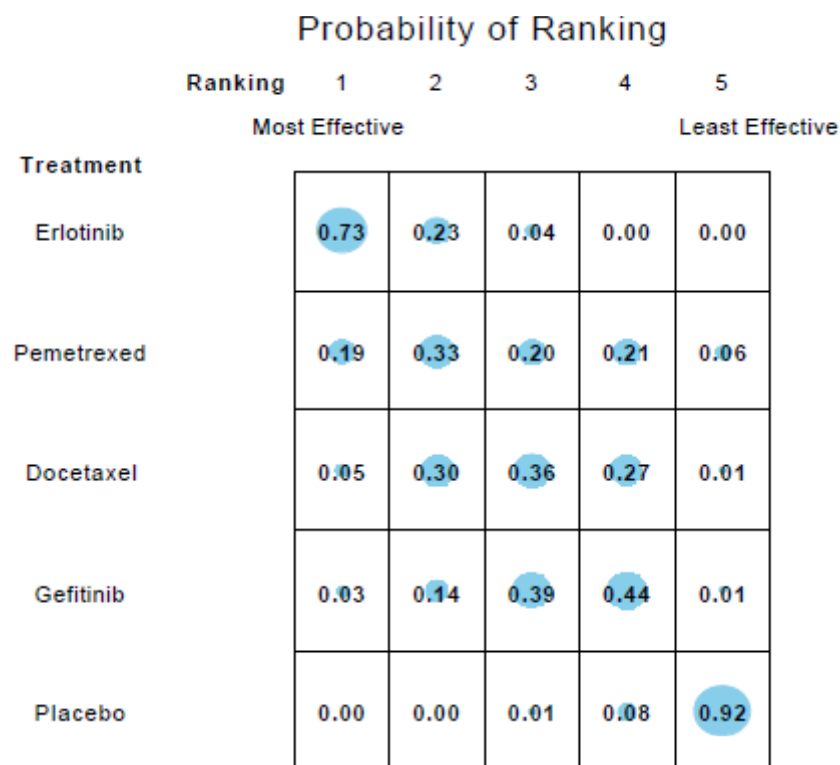


Treatment regimen (vs. DMARD + placebo)	Random effects—odds ratio (95 % credible interval)	% of patients with ACR50 response (95 % credible interval) ^a	Probability best (%)
Abatacept + DMARD	3.63 (2.09, 6.34)	33.1 (21.1, 47.8)	0.0
Adalimumab + DMARD	3.87 (2.30, 6.60)	34.5 (22.6, 49)	0.1
Certolizumab + DMARD	9.77 (4.60, 22.65)	57.1 (37.5, 76.3)	37.6
Etanercept + DMARD	11.15 (4.95, 27.95)	60.4 (39.2, 79.8)	55.6
Golimumab + DMARD	4.92 (2.05, 12.34)	40.1 (21.2, 63.6)	3.6
Infliximab + DMARD	3.60 (2.25, 5.92)	32.9 (22.3, 46.3)	0.1
Rituximab + DMARD	4.10 (1.82, 9.73)	35.9 (19.2, 57.8)	1.4
Tocilizumab + DMARD	5.40 (2.91, 9.56)	42.3 (27.4, 58)	1.6



CUIDADO cuando la red de evidencia es débil, existen pocos estudios para algunas comparaciones o son de pequeño tamaño y de baja calidad. Prestar atención a la magnitud de los efectos y su incertidumbre.

El ranking puede verse afectado y cambiar cuando falta alguna alternativa de tratamiento de interés o existen indicios de sesgo de publicación



	Placebo	Erlotinib	Gefitinib	Pemetrexed	Docetaxel
Placebo	NA	1.43 (1.18:1.72)	1.14 (1.01:1.29)	1.2 (0.92:1.52)	1.18 (1:1.39)
Erlotinib	0.7 (0.58:0.85)	NA	0.8 (0.64:1)	0.84 (0.61:1.13)	0.83 (0.64:1.06)
Gefitinib	0.88 (0.78:0.99)	1.26 (1:1.57)	NA	1.05 (0.84:1.29)	1.04 (0.93:1.16)
Pemetrexed	0.85 (0.66:1.08)	1.22 (0.88:1.64)	0.97 (0.77:1.19)	NA	1 (0.82:1.2)
Docetaxel	0.85 (0.72:1)	1.22 (0.94:1.55)	0.97 (0.86:1.08)	1.01 (0.83:1.22)	NA



Network comparator treatments	Rate ratio (95% credible interval)	Treatment	Probability of being best treatment (%)
—	—	Placebo	0
Adjusted standard dose warfarin v placebo	0.37 (0.26 to 0.53)	Adjusted standard dose warfarin	3
Adjusted low dose warfarin v placebo	0.32 (0.18 to 0.56)	Adjusted low dose warfarin	16
Fixed low dose warfarin v placebo	0.76 (0.30 to 1.76)	Fixed low dose warfarin	1
Aspirin v placebo	0.62 (0.43 to 0.86)	Aspirin	0
Fixed low dose warfarin and aspirin v placebo	0.98 (0.60 to 1.67)	Fixed low dose warfarin and aspirin	0
Ximelagatran v placebo	0.35 (0.19 to 0.65)	Ximelagatran	11
Alternate day aspirin v placebo	0.17 (0.01 to 1.15)	Alternate day aspirin	66
Indobufen v placebo	0.46 (0.19 to 1.14)	Indobufen	5





RESEARCH

The effects of excluding treatments from network meta-analyses: survey

OPEN ACCESS

Edward J Mills *associate professor*^{1,2}, Steve Kanfers *biostatistician*¹, Kristian Thorlund *associate professor*^{2,3}, Anna Chaimani *PhD student*⁴, Areti-Angeliki Veroniki *PhD student*⁴, John P A Ioannidis *professor and director*²

Table 2 | Nodes leading to greatest change in estimated effects and average fold changes in estimated treatment effects when single treatment nodes are removed

Network	Node leading to largest Brier score*	Connectivity of removed node	Removal of largest expected impact node		Removal of other active intervention nodes	
			Average fold change	Largest fold change	Average fold change	Largest fold change
Change in top three ranked treatments:						
Brown 2006 ¹⁶	Cox 2 preferentials (20%)	3 (max 4)	1.83	3.13	1.38	4.40
Cipriani 2009 ¹⁷	Fluoxetine (0%)	11 (max)	1.06	1.18	1.03	1.16
Cooper 2006 ¹⁸	Fixed low dose warfarin+aspirin (0%)	4 (max 7)	1.25	2.65	1.06	1.60
Gollinopoulos 2007 ²¹	Fluorouracil (0%)	5 (max 7)†	1.03	1.05	1.01	1.09
Imamura 2010 ²²	Pelvic floor muscle training+biofeedback (5%)	6 (max 7)	1.22	1.40	1.07	2.18
Mauri 2008 ²⁵	Standard dose anthracycline (0%)	10 (max)	1.05	1.12	1.01	1.09
Mills 2012 (6 mth) ²⁷	Low dose nicotine replacement therapy (0%)	4 (max)	1.12	1.39	1.01	1.03
Sciarretta 2011 ²⁹	Calcium channel blockers (1%)	6 (max)	1.27	1.53	1.04	1.31
Wang 2010 ³¹	Chlorhexidine and silver sulfadiazine (0%)	4 (max)	1.27	4.64	1.01	1.08
No changes in top three ranked treatments:						
Ara 2009 ³⁴	Simvastatin 40 mg (21%)	4 (max)	1.13	1.19	1.04	1.11
Bansback 2009 ³⁵	Adalimumab (0%)	2 (max 3)	1.08	1.32	1.01	1.02
Dias 2010 ³⁶	Accelerated alteplase (0%)	7 (max)	1.08	1.16	1.03	1.07
Elliot 2007 ³⁸	Angiotensin converting enzyme inhibitors (22%)	4 (max 5)	1.08	1.12	1.02	1.06
Kyrgiou 2006 ³⁹	Platinum monotherapy (0%)	6 (max)	1.08	1.25	1.02	1.13
Lam 2008 ⁴⁴	Implantable cardioverter defibrillator (9%)	3 (max 4)†	1.03	1.07	1.01	1.04
Mills 2011 ⁴⁶	Inhaled corticosteroids (0%)	3 (max 6)	1.01	1.02	1.00	1.01
Psaty 1997 ⁴⁸	Diuretics (10%)	5 (max)	1.06	1.14	1.03	1.08
Thijs 2008 ⁵⁰	Aspirin (0%)	4 (max)	1.03	1.04	1.01	1.03

*Probability of being best treatment.

†Most connected node that was removed.

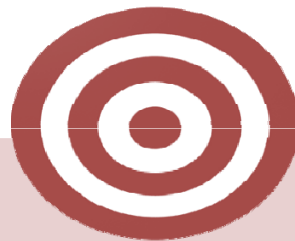


Puntos claves



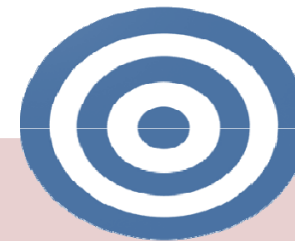
Lo más simple

Un sólo ensayo de A vs C
y uno de B vs C.
Combinados estiman A
vs B



2ª Posibilidad

Uno de los brazos o
ambos son metanálisis



Otra posibilidad

Comparaciones
múltiples entre distintos
tratamientos

Puntos clave:
Homogeneidad
Consistencia (coherencia)
Transitividad (similitud)

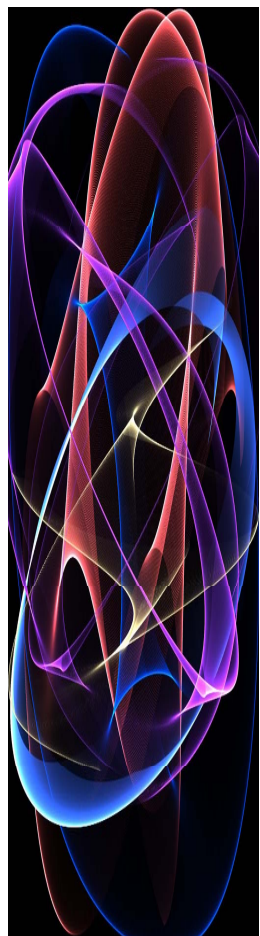


Heterogeneidad

- Estudios en diferentes lugares, tiempos, poblaciones..resultados diferentes
- Ver forest plot
- Estadística, test χ^2 : $p < 0,1$ heterogeneidad
- Test I^2 : 25% heterogeneidad baja; 50% moderada; 75% alta
- *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]:
 - 0% a 40%: puede no ser importante
 - 30% a 60%: puede representar heterogeneidad moderada
 - 50% a 90%: puede representar heterogeneidad sustancial
 - 75% a 100%: heterogeneidad considerable
- Los autores deben indicar los motivos de la heterogeneidad
- Test I^2 : Cuantifica el porcentaje de la variabilidad de las estimaciones del efecto del tratamiento que se debe a la variabilidad entre estudios



Modelos estadísticos



- El modelo de efectos fijos asume que los estudios incluidos en el metaanálisis están estimando, a un único efecto poblacional, variabilidad debida al azar, variabilidad intraestudio.
- La generalización se limita a la población de estudios de características similares a los incluidos en el meta-análisis

- El modelo de efectos aleatorios asume que cada tamaño del efecto individual estima a un tamaño del efecto poblacional.
- Los resultados pueden generalizarse a una población mayor de estudios



Puntos claves en la evaluación crítica

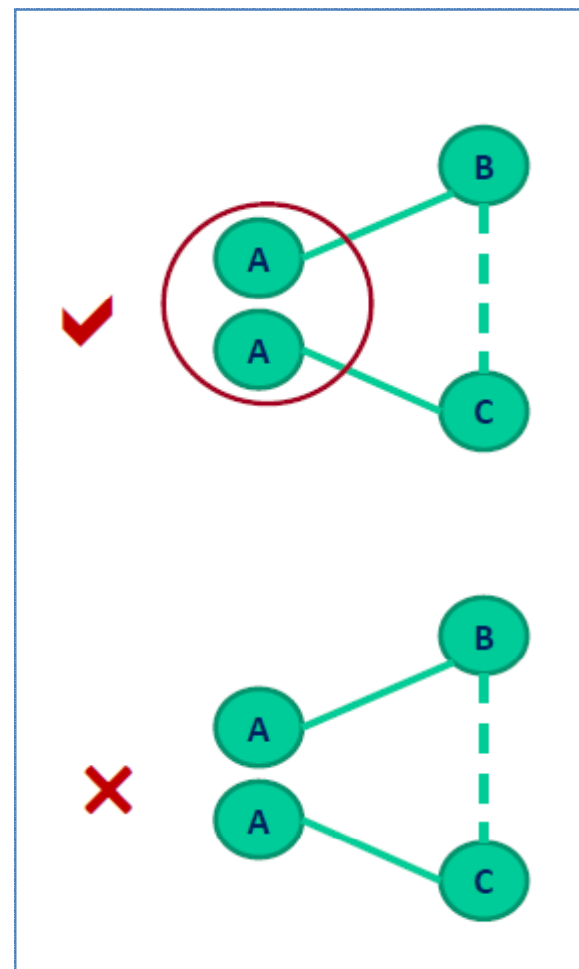
Transitividad

No se puede probar estadísticamente pero puede ser evaluada conceptual y epidemiológicamente

Grupo control similar

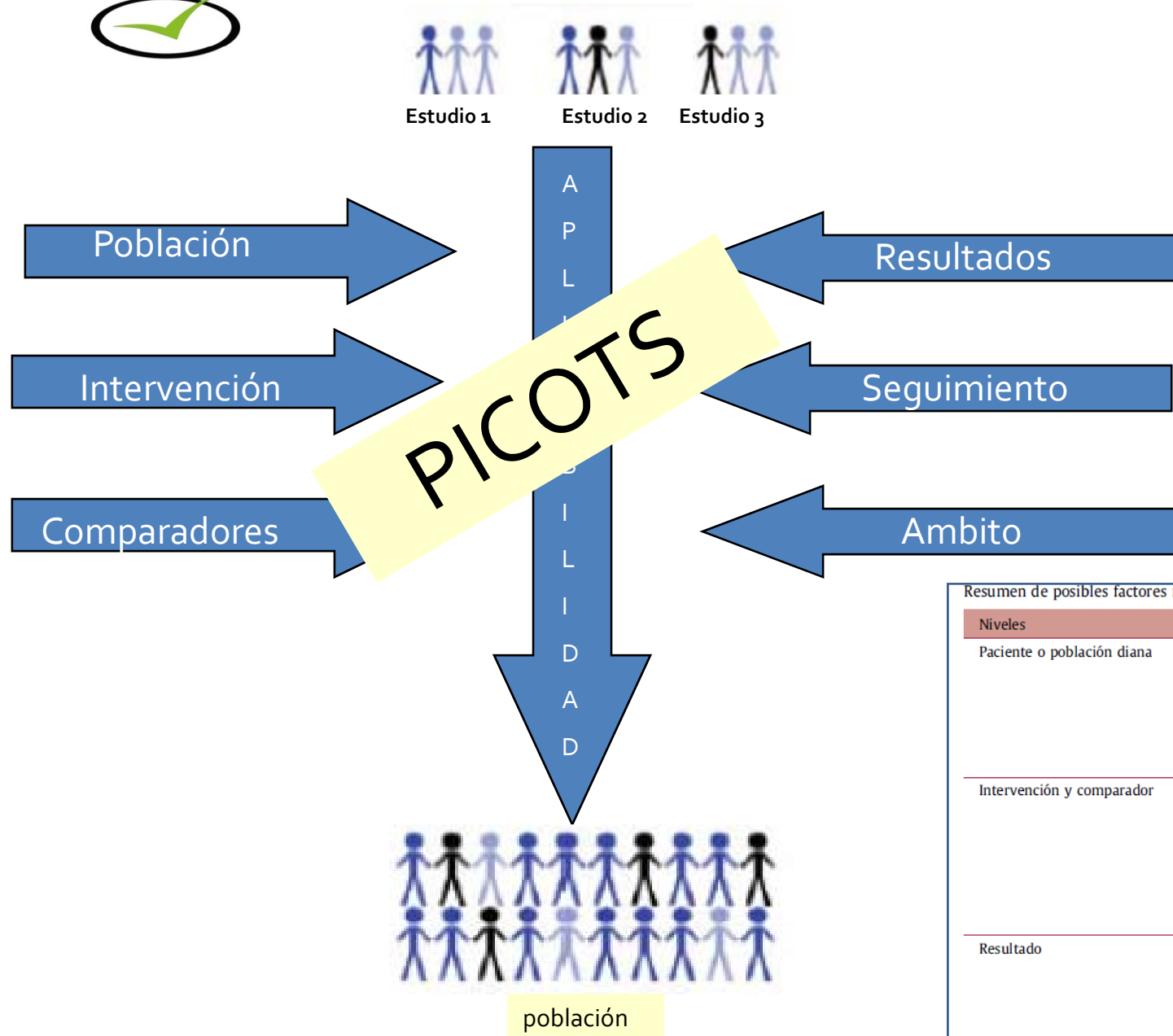
Ensayos AC y AB no difieran en cuanto a la distribución de los factores modificadores del efecto

Se asume para la media de los tamaños de los efectos y no para los estudios individuales





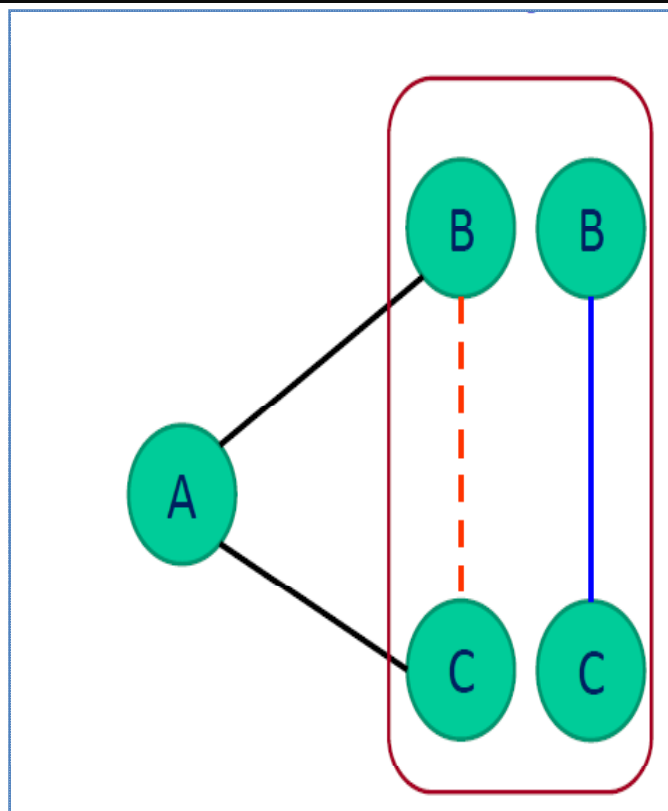
Puntos críticos



Resumen de posibles factores modificadores del efecto	
Niveles	Posibles factores modificadores del efecto
Paciente o población diana	• Edad • Sexo • Gravedad (riesgo basal) • Duración de la enfermedad • Comorbilidad(es) • No responde al tratamiento previo • Proveedor o ámbito*
Intervención y comparador	• Tipo de intervención (preventiva, terapéutica, etc.) • Dosis o intensidad • Duración • Vía de administración • Tratamientos concomitantes • Complejidad • Proveedor o ámbito*
Resultado	• Definición de la variable de resultado (<i>outcome</i>) • Herramientas para medir o monitorizar • Métodos y procedimientos • Duración del seguimiento

*Proveedor (público o privado) o ámbito (atención primaria u hospitalaria).

Consistencia



Analiza la coherencia entre las comparaciones directas e indirectas
Se pueden ver las discrepancias entre lo estimado mediante comparación directa e indirecta

$$\mu_{BC}^I \quad \mu_{BC}^D \rightarrow \mu_{BC}^M$$

$$\mu_{BC}^I = \mu_{BC}^D$$

$$\mu_{AC}^D - \mu_{AB}^D = \mu_{BC}^I = \mu_{BC}^D$$



Consistencia

Consistencia es la manifestación estadística de la transitividad. Necesario bucle cerrado

Inconsistencia se refiere al grado de desacuerdo entre los efectos del tratamiento de fuentes específicas. Se mide por las diferencias entre estimaciones directas e indirectas que no pueden ser explicadas por el azar.

Inconsistencia estadística diferencia estadísticamente significativa entre μ^D_{BC} y μ^I_{BC}

Consistencia puede ser evaluada estadísticamente comparando μ^D_{BC} y μ^I_{BC} mediante el z-test.

Alternativa: Factores de inconsistencia $IF = |\mu^D_{BC} - \mu^I_{BC}|$

Si la consistencia se mantiene podemos combinar evidencia directa e indirecta



Consistencia

CALCULATOR FOR MIXED COMPARISONS (for relative effects)

		95% CI			
A vs B (B/A)	Effect size	Lower	Upper		
Direct comparison	0,79	0,59	1,06		
Indirect comparison	0,72	0,45	1,16		

		95% CI			
A vs B (B/A)	Effect size	Lower	Upper	z	p
Mixed comparison	0,77	0,61	0,99	-2,06	0,0397

		95% CI			
	IF	Lower	Upper	z	p
Inconsistency factor	1,10	0,63	1,91	0,33	0,7440

Hay evidencia para no rechazar la hipótesis nula de consistencia entre ambos tipos de comparaciones ($p=0,7440$), por lo que estadísticamente la comparación mixta es consistente



Box 3. Potential Reasons for Incoherence Between the Results of Direct and Indirect Comparisons

Chance

Genuine diversity

Differences in enrolled participants (eg, entry criteria, clinical setting, disease spectrum, baseline risk, selection based on prior response)

Differences in the interventions (eg, dose, duration of administration, prior administration [second-line treatment])

Differences in background treatment and management (eg, evolving treatment and management in more recent years)

Differences in definition or measurement of outcomes

Bias in head-to-head (direct) comparisons

Optimism bias with unconcealed analysis

Publication bias

Selective reporting of outcomes and analyses

Inflated effect size in early stopped trials and in early evidence

Limitations in allocation concealment, blinding, loss to follow-up, analysis as randomized

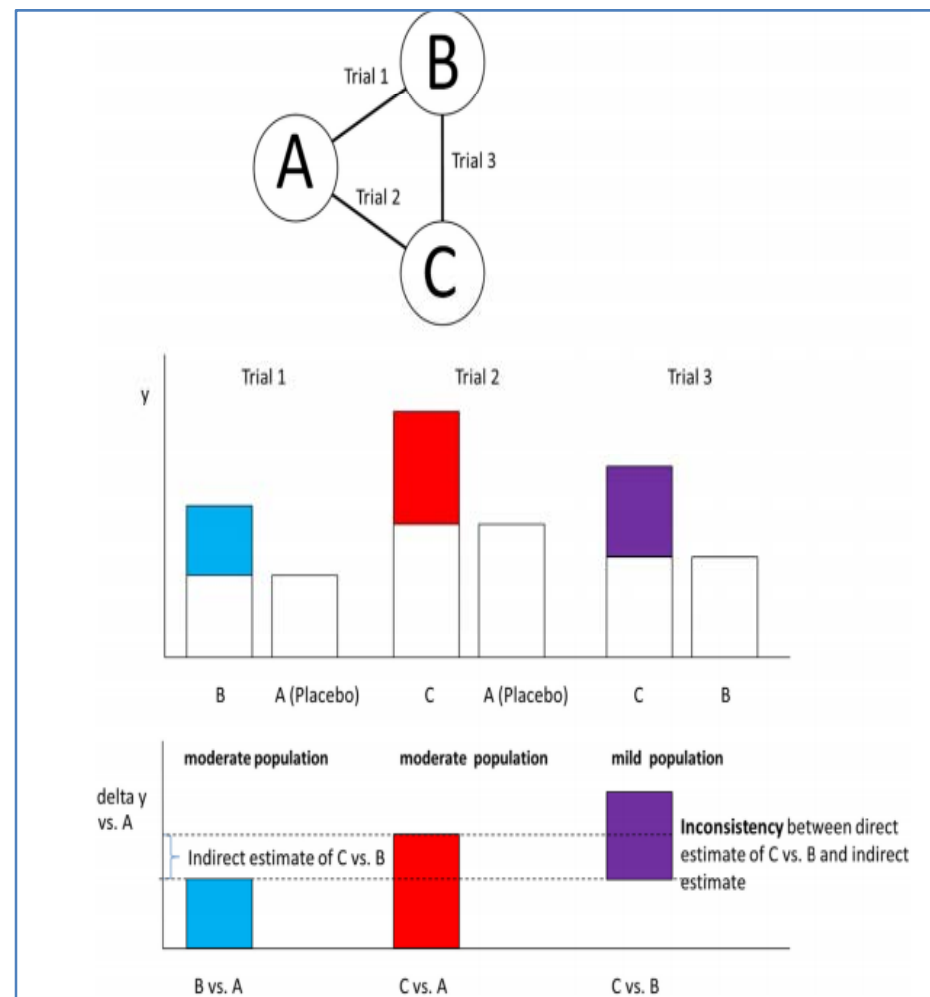
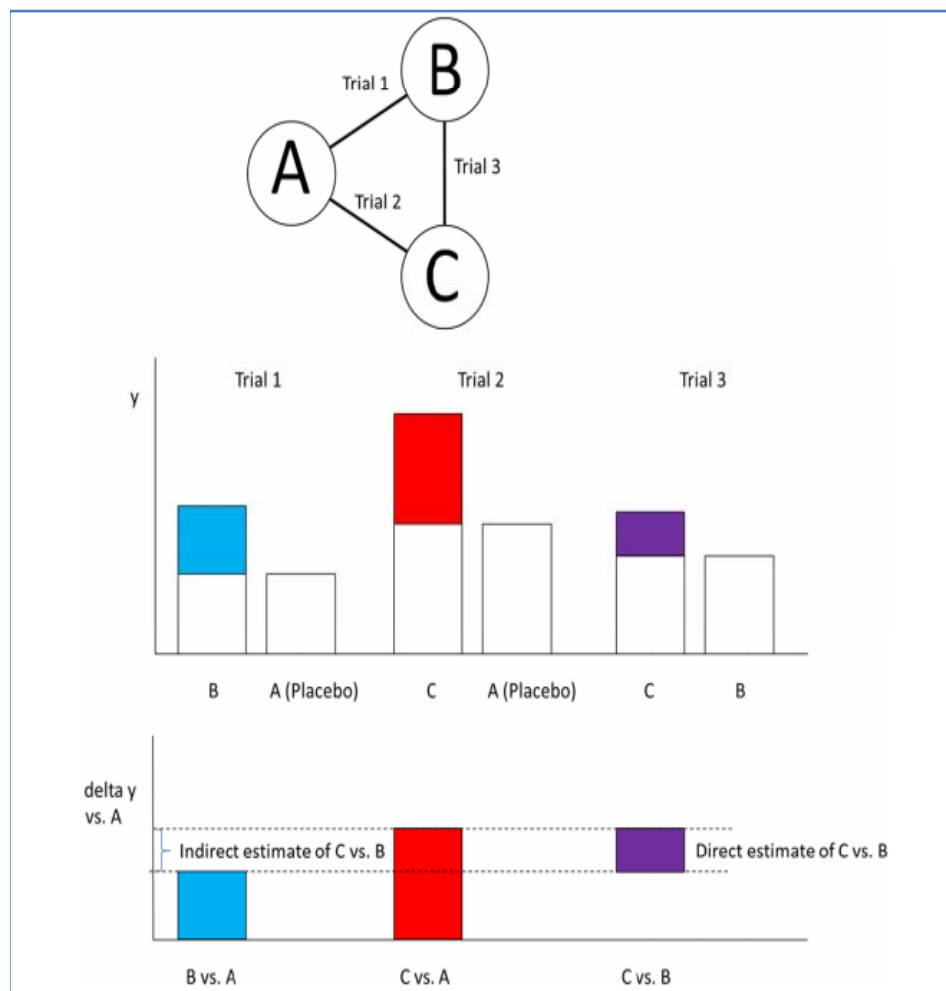
Bias in indirect comparisons

Each of the biasing issues above can affect the results of the direct comparisons on which the indirect comparisons are based

JAMA. 2012;308(12):1246-1253



Consistencia



Evaluation of inconsistency in networks of interventions

Areti Angeliki Veroniki,¹ Haris S Vasiliadis,^{2,3} Julian PT Higgins^{4,5} and Georgia Salanti^{1*}

¹Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, ²Department of Orthopaedics, School of Medicine, University of Ioannina, Greece, ³Molecular Cell Biology and Regenerative Medicine, Sahlgrenska Academy, University of Gothenburg, Sweden, ⁴MRC Biostatistics Unit, Cambridge, UK and ⁵Centre for Reviews and Dissemination, University of York, York, UK

*Corresponding author. Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, University Campus, Ioannina 45110, Greece. E-mail: gsalanti@cc.uoi.gr

Examinaron 315 bucles y hasta el 10% fueron inconsistentes

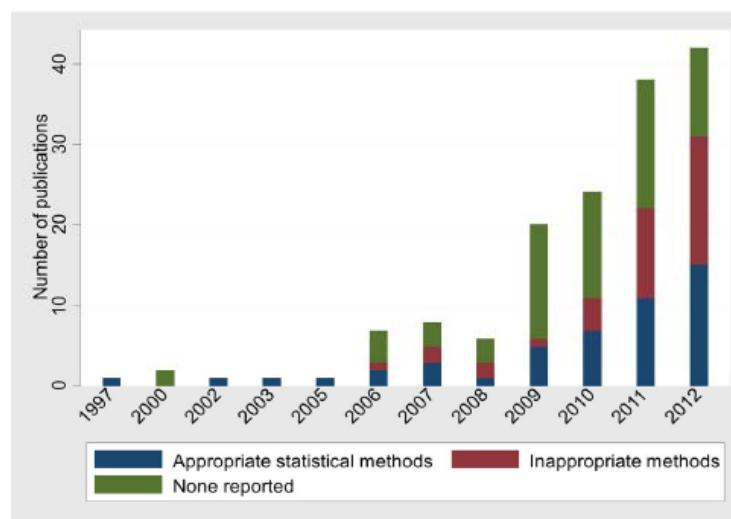
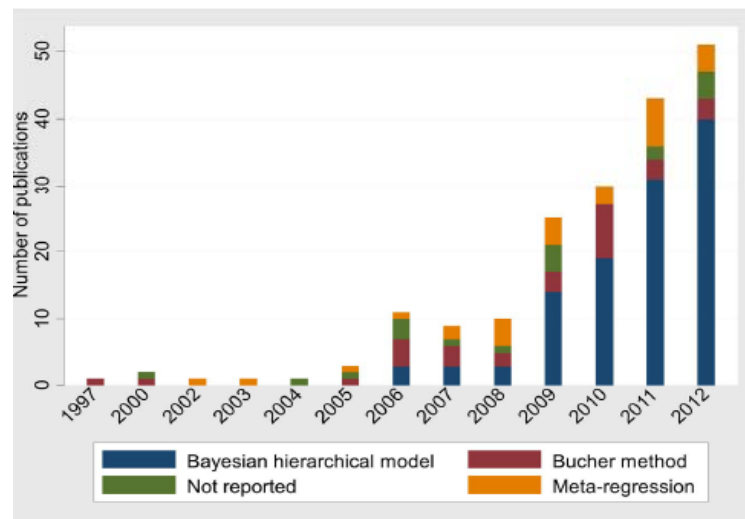
- Depende del estimador de la heterogeneidad
- La inconsistencia es más probable en bucles con comparaciones informadas por un solo estudio

Examinaron 40 networks y uno de cada 8 fue estadísticamente inconsistente





...En la evaluación de medicamentos



Variabilidad en los métodos empleados
Uso de métodos inapropiados para valorar la inconsistencia

Characteristics of Networks of Interventions: A
Description of a Database of 186 Published Networks

Adriani Nikolakopoulou¹, Anna Chaimani¹, Areti Angeliki Veroniki¹, Haris S. Vasiliadis^{2,3},
Christopher H. Schmid⁴, Georgia Salanti^{1*}

PLoS ONE | www.plosone.org

January 2014 | Volume 9 | Issue 1 | e80754



Table 2. Possible strategies that respond to heterogeneity and inconsistency.

Action	Heterogeneity	Inconsistency
Check the data	Studies that 'stand out' in the forest plot are checked for data extraction errors (Götzsche <i>et al.</i> , 2007; Horton <i>et al.</i> , 2010).	Simple loop inconsistency graphs can indicate studies with data extraction errors. Inconsistency in loops where a comparison is informed by a single study is particularly suspicious for data errors.
Try to bypass	There is empirical evidence that some measures are associated with larger heterogeneity than others (Deeks, 2002; Friedrich <i>et al.</i> , 2011).	Empirical evidence is needed to evaluate the impact of the effect measure on statistical inconsistency, particularly given the trade-off between heterogeneity and statistical inconsistency when random effects models are used.
Resign to it	Investigators may decide not to undertake meta-analysis in the presence of excessive heterogeneity.	Investigators may decide not to synthesize the network data in the presence of excessive inconsistency.
Encompass it	Apply random effects meta-analysis (DerSimonian and Laird, 1986; Higgins <i>et al.</i> , 2009).	Apply models that relax the consistency assumption by adding an 'extra' inconsistency-specific random effect (Lu and Ades, 2006; White <i>et al.</i> , 2012). However, as random effects are not a remedy for excessive heterogeneity and should be applied only for unexplained heterogeneity, inconsistency models should be employed to reflect inconsistency in the results, not to adjust for it.
Explore it	Use prespecified variables in a subgroup analysis or meta-regression (Thompson and Higgins, 2002).	Split the network into subgroups or use network meta-regression to account for differences across studies and comparisons. Specify the variables in the protocol, including bias-related characteristics. See for examples Cooper <i>et al.</i> (2009); Dias <i>et al.</i> (2010c); Salanti <i>et al.</i> (2009).

Salanti G Res Syn Meth 2012;3:80-97



Puntos claves

BMJ

BMJ 2013;347:f3675 doi: 10.1136/bmj.f3675 (Published 1 July 2013)

Page 1 of 12

RESEARCH

72% presentaba de manera inadecuada aspectos metodológicos

Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review

 OPEN ACCESS

Aida Bateta *PhD student*^{1,2}, Ludovic Trinquart *senior statistician*^{1,2,3,4}, Raphaële Seror *associate professor of rheumatology*^{1,2,4}, Philippe Ravaud *professor of epidemiology and director*^{1,2,3,4,5}

121 trabajos: 73% no indicaba la estrategia de búsqueda, 50% no incluía evaluación del riesgo de sesgo, el 79% no presentaba las características de los estudios individuales

Items	Overall (n=121)	General journals (n=55)	Specialty journals (n=66)
Other items			
Assumptions required in network meta-analysis			
Homogeneity assumption	107 (88)	50 (91)	57 (86)
Similarity assumption	41 (34)	18 (33)	23 (35)
Consistency or exchangeability assumption	58 (48)	32 (58)	26 (39)
Conflict of interest	95 (79)	47 (85)	48 (73)

Data are number (%) of reports featuring the corresponding item.

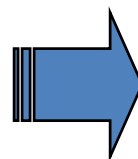


Puntos clave en la valoración crítica

1º punto fundamental TRANSPARENCIA

Datos gráficos y narrativos

Evidencia
Transitividad
Consistencia
Impacto clínico
Generalización
Aplicabilidad



Validez interna
Validez externa
Aplicabilidad

Puntos fuertes y débiles

Relevancia de cada uno de los puntos

Valoración cualitativa: Fiabilidad



FIABILIDAD ALTA
FIABILIDAD ACEPTABLE
FIABILIDAD DUDOSA
FIABILIDAD BAJA

Criterios GRADE???

GENESIS



Sociedad Española de
Farmacia Hospitalaria

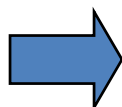




Puntos clave en la valoración crítica

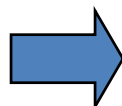
Definiciones

Heterogeneidad Clínica



Variación en las características de la población, condiciones coexistentes, cointervenciones y resultados evaluados en los estudios que pueden influir o modificar la magnitud del efecto de la medida de la intervención

Heterogeneidad Metodológica



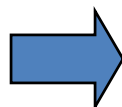
Diferencias entre estudios en los tamaños estimados del efecto que pueden ser atribuidos a la variabilidad y calidad y a los tipos de diseños de los estudios

Heterogeneidad Estadística



Variabilidad en los efectos observados de los tratamientos más allá de los esperado por el error aleatorio

Medida del Efecto



Expresión de las medidas de la variable (OR, RR, DR, RAR...)

Adaptada de West SL et al. AHRO. 2010

GENESIS



Sociedad Española de
Farmacia Hospitalaria





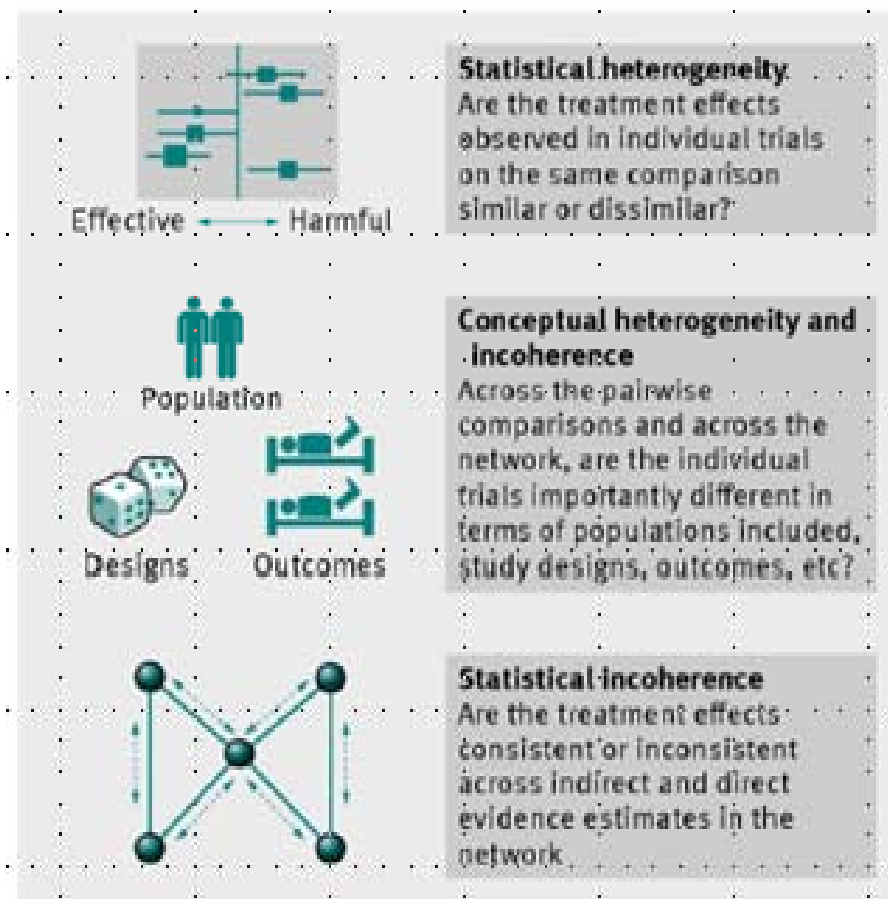
Puntos clave en la valoración crítica

Heterogeneidad

Excesiva discrepancia entre los efectos específicos entre los estudios

Inconsistencia

Excesiva discrepancia entre los efectos específicos de la fuente (directa e indirecta)





.... Dónde estamos en las comparaciones indirectas



- La información comparativa es crítica para la toma de decisiones
- Si no existen ECAs comparativos directos utilizar evidencia indirecta
- Utilizar CI ajustadas (NO USAR COMPARACIONES INDIRECTAS NAÏVE)
- Debe exigirse transparencia en la publicación (permitirá valorar la validez)
- Existen confusión en términos y se están desarrollando herramientas estadísticas más apropiadas
- La necesidad de evidencia robusta ha llevado a incorporar al NICE y agencias reguladoras la metodología de las comparaciones indirectas







An indirect comparison of dabigatran, rivaroxaban and apixaban for atrial fibrillation

Simon Mantha¹; Jack Ansell²

¹Department of Medicine, Lahey Clinic, Burlington, Massachusetts, USA; ²Department of Medicine, Lenox Hill Hospital, New York, New York, USA

Summary

New oral anticoagulant drugs are emerging as alternatives to warfarin for the prevention of stroke in patients with non-valvular atrial fibrillation. Two agents are direct factor Xa inhibitors (rivaroxaban and apixaban), and the third is a direct thrombin inhibitor (dabigatran). They have been separately compared to warfarin in large randomised trials. Our objective was to indirectly compare the three agents to each other for major efficacy and safety outcomes. Studies were assessed for comparability and the odds ratios of selected outcomes for each anticoagulant versus one another were estimated indirectly. The three cohorts differed significantly in terms of CHADS₂ score and the number of individuals with a past history of stroke, transient ischemic attack or systemic embolism. The estimated odds ratio of stroke or systemic embolism was 1.35 for rivaroxaban vs dabigatran 150 mg ($p=0.04$), 0.97 for rivaroxaban versus dabigatran 110 mg ($p=0.81$), 1.22 for apixaban versus dabigatran 150 mg ($p=0.18$), 0.88 for apixaban versus dabigatran 110 mg ($p=0.34$) and 0.90 for apixaban versus rivaroxaban ($p=0.43$). The esti-

ated odds ratio of major bleeding was 1.10 for rivaroxaban versus dabigatran 150 mg ($p=0.36$), 1.28 for rivaroxaban versus dabigatran 110 mg ($p=0.02$), 0.74 for apixaban versus dabigatran 150 mg ($p=0.004$), 0.87 for apixaban versus dabigatran 110 mg ($p=0.17$) and 0.68 for apixaban versus rivaroxaban ($p<0.001$). In conclusion, the available data indicate no significant difference in efficacy between dabigatran 150 mg and apixaban for the prevention of stroke or systemic embolism in patients with non-valvular atrial fibrillation. It appears however that apixaban is associated with less major bleeding than dabigatran 150 mg or rivaroxaban and that rivaroxaban is less effective than dabigatran 150 mg in preventing stroke or systemic embolism. Such an indirect comparison should be used only to generate hypotheses which need to be tested in a dedicated randomised trial comparing the three drugs directly.

Keywords

Atrial fibrillation, anticoagulants, stroke

Correspondence to:
Simon Mantha, MD, MPH
Department of Hematology/Oncology
Lahey Clinic Medical Center
41 Mall Road
Burlington, MA 01805, USA
Tel.: +1 781 744 7590, Fax: +1 781 744 5245
E-mail: simon.mantha@lahey.org

Received: February 16, 2012
Accepted after major revision: May 7, 2012
Prepublished online: June 28, 2012
doi:10.1160/TH12-02-0093
Thromb Haemost 2012; 108: 476–484



1. ¿Tiene interés leer una comparación indirecta de nuevos anticoagulantes orales (ACO) para su posicionamiento en terapéutica?

- ☐ Sí
- ☐ No

Si has contestado SI, explica por qué

2. ¿Cuáles son los resultados de eficacia del estudio?: ¿se observaron diferencias significativas entre los tres anticoagulantes?

3. La metodología y otras características del estudio, ¿nos permiten confiar en estos resultados y aplicarlos en la selección?



4. La comparación de ACO, ¿condicionan la selección de alguno de ellos como tratamiento de elección, al menos en algún grupo de pacientes?

5. Según estos datos, ¿podrían los nuevos ACO ser posicionados como alternativas terapéuticas equivalentes, seleccionando el fármaco en función de su precio?



Tabla 5.2.b.1.2

CUESTIONARIO SOBRE LA VALIDEZ DE UNA COMPARACIÓN INDIRECTA

	SI/NO	JUSTIFICAR
-¿Es adecuado el método utilizado para realizar la comparación indirecta? Rechazar comparaciones "naïve" realizadas con estimaciones puntuales derivadas de diferentes ensayos controlados o de los brazos activos de diferentes ensayos controlados		
-¿Se establece claramente cómo se realizó la búsqueda y la selección de los ensayos para la inclusión?		
-¿Se hace una descripción de los métodos de análisis/síntesis de la evidencia? Manejo de sesgos		
-¿Se analiza la homogeneidad de los ensayos y la estabilidad de los efectos? Todo los ensayos responden a la misma pregunta clínica en relación con el tipo de pacientes estudiado y los tratamientos comparados		
-¿Existe concordancia de los resultados? El resultado es el mismo independientemente de la cadena de comparaciones usadas para obtenerlo		
-Si existe una comparación directa ¿se observa consistencia en los resultados?		
-Interpretación de los resultados. ¿Permiten obtener conclusiones claras? Analizar diferencias significativas en los riesgos basales y las respuestas en el grupo placebo		
-Otros sesgos o limitaciones encontradas en el estudio		

Tabla 5.3.b.1.3

CUESTIONARIO SOBRE LA APLICABILIDAD DE UNA COMPARACIÓN INDIRECTA

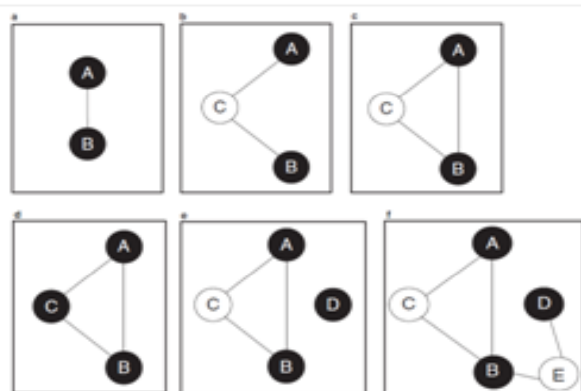
	SI/NO	JUSTIFICAR
¿Considera adecuado el/los comparadores? ¿Es el tratamiento del comparador común adecuado en nuestro medio?		
¿Son importantes clínicamente los resultados?		
¿Considera adecuada la variable de medida utilizada?		
¿Considera adecuados los criterios de inclusión y/o exclusión de los pacientes? Generalización de los hallazgos (Población de pacientes en los ensayos y entre los ensayos)		
¿Cree que los resultados pueden ser aplicados directamente a la práctica clínica?		
Otros sesgos o limitaciones encontradas en el estudio		





Lista de comprobación

LISTA DE COMPROBACIÓN DE COMPARACIONES INDIRECTAS (CI)¹



CITA BIBLIOGRÁFICA A VALORAR.

Manolis S, Ansell J. An indirect comparison of dabigatran, rivaroxaban and apixaban for atrial fibrillation. *Thromb Haemost* 2012;108(3):476-84.

¹ Ortega A., Riera M.D., Alegre-del-Rey E.J., Pulgarín-Latorre F., Porta A., Venizy P., Tenías J.M., Hawkins N.S., Caldwell D.M. A checklist for critical appraisal of indirect comparisons. *Int J Clin Pract* in press

El objetivo principal de esta lista guía es disponer de una herramienta de fácil uso para determinar si un estudio de comparación indirecta es adecuado para ayudarnos a tomar una decisión en la evaluación de medicamentos. El esquema propuesto está basado en las plantillas del grupo CASPe² y en artículos sobre metodología de las comparaciones indirectas^{3,4,5,6,7}.

El documento está estructurado en dos partes, la primera consta de tres preguntas de eliminación si tras su contestación el resultado es que no es adecuada la CI debemos plantearnos no seguir con el estudio y pasar a otro si existe. En el caso de que no exista ningún otro en función de la evidencia disponible podemos plantearnos seguir siempre que consideremos los sesgos, inconsistencias, etc que pueden producirse.

En la segunda parte se analiza con detalle el estudio, incidiendo en todos los aspectos recogidos en los estudios publicados sobre metodología que se han incluido en esta lista guía.

Por último se presenta un anexo con un glosario de términos.

La lectura crítica de estudios de comparaciones indirectas requiere que se tengan en cuenta tres aspectos generales:

- ¿Es válida esta comparación indirecta?
- ¿Cómo se evalúan y comparan los resultados?
- ¿Ayudarán los resultados a establecer una decisión terapéutica para nuestra población?

Para terminar, un código de colores nos indicará el grado de fiabilidad del estudio:

FIABILIDAD ALTA	FIABILIDAD DUDOSA
FIABILIDAD ACEPTABLE	FIABILIDAD BAJA

² Cabellro, J.B. por CASPe. Plantilla para ayudar a entender una Revisión Sistemática. En: CASPe. Guía CASPe: la Lectura Crítica de la Literatura Médica. Alicante: CASPe; 2005. Cuadro 1. p.12-17.

³ Song F, Loke YK, Walsh T, Glanzy AM, Barwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009 ;339:b1147. doi: 10.1136/bmj.b1147.

⁴ Song F, Altman DG, Glanzy AM, Desai JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003 Mar 1;326(7187):672.

⁵ Dias S, Welton NI, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29(7-8):932-44.

⁶ Caldwell DM, Welton NI, Ades AE. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *J Clin Epidemiol*. 2010 Jan 14.

⁷ Eckermann S, Coory M, Willan AR. Indirect comparison: relative risk fallacies and odds solution. *J Clin Epidemiol*. 2009;62(10):1071-6.





Lista de comprobación

Justificación

¿El método usado es apropiado?

Búsqueda

PARTE A/PREGUNTAS ELIMINATORIAS: ¿Son válidos los resultados de la comparación indirecta?

¿Está justificado realizar una comparación indirecta (CI)? ¿Existe una comparación directa? Las comparaciones directas son más fiables (generalmente asociadas a menores sesgos) que las indirectas. Si hay una comparación directa y una indirecta puede ser interesante valorar si llegan a la misma conclusión, es decir si hay consistencia (no hay diferencias) entre los resultados de ambas. Si no hay diferencias aumentará nuestra fiabilidad en el resultado y si hay diferencias debemos cuestionarnos a qué pueden deberse esas diferencias. Para valorar si se puede hacer una comparación indirecta se debe cumplir: ¿Están claros el objetivo, la población, los métodos (incluidas asunciones) y los resultados? (si no hay transparencia en todos estos aspectos no se puede o dificulta la evaluación de la CI) ¿Los estudios comparados son comparables clínicamente? (Características de la población estudiada, condiciones consistentes, intervenciones, dosificación, duración del tratamiento y resultados evaluados que influyen o modifican la magnitud del efecto de la intervención) ¿La medida de resultados es comparable en cuanto a su definición y su presentación? ¿La medida de resultados es clínicamente relevante?	☐ SI ☐ NO ☐ NO SÉ
¿El método de CI es apropiado? Las comparaciones indirectas no ajustadas (que no comparan el efecto relativo, sino que cogen un brazo de un estudio sin considerar al comparador) se consideran no apropiadas. Tipos (Ver anexo I): CI no ajustada, CI no ajustada cruda, CI no ajustada informal, CI ajustada, Metaanálisis de ratios de tratamiento o network-metaanálisis Métodos de comparaciones indirectas: Bucher, Metaanálisis de ratios de tratamiento, modelos de comparaciones indirectas, metaanálisis de múltiples tratamientos. ¿Existe una explicación narrativa de todas las comparaciones?	☐ SI ☐ NO ☐ NO SÉ
¿Se realizó una búsqueda adecuada de todos los estudios pertinentes? ¿Se dirige adecuadamente la pregunta para establecerla CI? (Nótese el orden PICO(Patient, Intervention, Comparison, Outcome)-ver anexo I) ¿Están incluidos todos los estudios relevantes? ¿El comparador es relevante? Tras la respuesta a estas tres preguntas, ¿consideras que debes continuar analizando este estudio? ☐ SI ☐ NO	☐ SI ☐ NO ☐ NO SÉ

☐ CONTINUAR ☐ STOP

3 CI-June 2014



GENESIS



Sociedad Española de Farmacia Hospitalaria



¿Está justificado realizar una comparación indirecta (CI)?

¿? ☒ SÍ ☐ NO
☐ NO SÉ

Heterogeneidad entre los ensayos incluidos tanto por las características de los pacientes (riesgo basal de ictus, control en el INR) y en la metodología del ensayos

La población no está muy claramente definida, existen algunos factores de confusión: diferencias en datos para comunicar la edad (medias y medianas), tratamientos concomitantes

En ROCKET AF no fueron incluidos (a diferencia de en ARISTOTLE y RE-LY) pacientes con CHADS2 menor a 2, por lo que sólo se incluyen dos subgrupos en el análisis: CHADS2 = 2 y CHADS2 ≥ 3.

ROCKET-AF y ARISTOTLE fueron diseñados como ensayos doble ciego. El tratamiento con warfarina no fue cegada en RE-LY, que podría ser una fuente potencial de sesgo. Mínimas diferencias en definiciones. Ictus y embolismo sistémico



Characteristic	RE-LY	ROCKET AF	ARISTOTLE	P-value
Total number of subjects enrolled	18,113	14,264	18,201	-
Individuals lost to follow-up (n)	20	32	69	<0.001
CHADS ₂ scores (n)				<0.001
0 or 1	5,775	0	6,183	
2	6,455	1,859	6,516	
3 to 6	5,882	12,402	5,502	
Age (years)	71 (mean)	73 (median)	70 (median)	_*
Female (n)	6,599	5,663	6,416	<0.001
Previously treated with a VKA (n)	8,989	8,904	10,401	<0.001
Previous embolic episode (n)	3,623 (stroke or TIA only)	7,845 (stroke, TIA or systemic embolism)	3,458 (stroke, TIA or systemic embolism)	<0.001
Mean time in the therapeutic range for the warfarin group (%)	64	55	62	_**
Aspirin use (n)	7,198	5,205	5,632	<0.001
Creatinine clearance [§]				<0.001
≤30 ml/min (n)	0	0	270	
>30 to 50 ml/min (n)	3,505	2,949	2,747	
>50 ml/min (n)	14,592	11,205	15,105	



Characteristic	RE-LY	ROCKET AF	ARISTOTLE	P-value
Total number of subjects enrolled	18,113	14,264	18,201	-
Individuals lost to follow-up (n)	20	32	69	<0.001
CHADS ₂ scores (n)				<0.001
0 or 1	5,775	0	6,183	
2	6,455	1,859	6,516	
3 to 6	5,882	12,402	5,502	
Age (years)	71 (mean)	73 (median)	70 (median)	_*
Female (n)	6,599	5,663	6,416	<0.001
Previously treated with a VKA (n)	8,989	8,904	10,401	<0.001
Previous embolic episode (n)	3,623 (stroke or TIA only)	7,845 (stroke, TIA or systemic embolism)	3,458 (stroke, TIA or systemic embolism)	<0.001
Mean time in the therapeutic range for the warfarin group (%)	64	55	62	_**
Aspirin use (n)	7,198	5,205	5,632	<0.001
Creatinine clearance [§]				<0.001
≤30 ml/min (n)	0	0	270	
>30 to 50 ml/min (n)	3,505	2,949	2,747	
>50 ml/min (n)	14,592	11,205	15,105	



- ¿El método de CI es apropiado?	☒ SÍ ☐ NO ☐ NO SÉ
- Las comparaciones indirectas no ajustadas (que no comparan el efecto relativo, sino que cogen un brazo de un estudio sin considerar el comparador) se consideran no apropiadas. Tipos (Ver anexo I): CI no ajustada, CI no ajustada cruda, CI no ajustada informal, CI ajustada, Metanálisis de redes de tratamiento o network-metanalysis - Métodos de comparaciones indirectas: Bucher, Metanálisis de redes de tratamiento, mezclas de comparaciones indirectas, metanálisis de múltiples tratamientos. - ¿Existe una explicación narrativa de todas las comparaciones?	



¿Se realizó una búsqueda adecuada de todos los estudios pertinentes?

¿? ☒ SÍ ☐ NO
☐ NO SÉ

Sólo se abordó la búsqueda en PubMed, no se especifican claramente los criterios, estrategia de búsqueda. No indican tiempo. Excluyen estudios piloto y fase II debido a la mayor heterogeneidad en el diseño y por el pequeño tamaño. Incluyen los 3 fase III, la actualización del RE-LY y consideraron los artículos sobre metodología de estos ensayos o análisis de subgrupos pertinentes. Solo consideran la información disponible públicamente

Debería incluir claramente los criterios PICO.

P: Pacientes con FA no valvular que requieren anticoagulación

I: NOAC

C: cumarinas (warfarina, acenocumarol)

O: Todas las causas de ictus/embolismo

Hemorragia mayor

Hemorragia intracraneal

Mortalidad cardiovascular

Todas las causas de mortalidad

Mortalidad cardiovascular

Hemorragia que amenace la vida

El comparador es relevante es el estándar en la actualidad

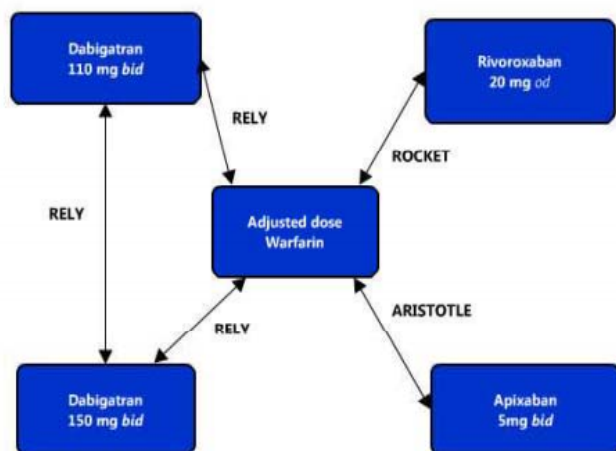
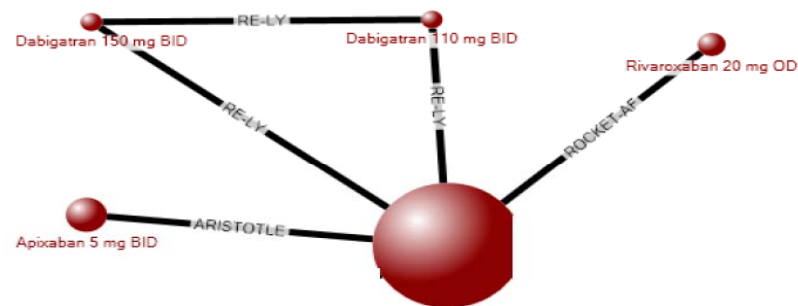
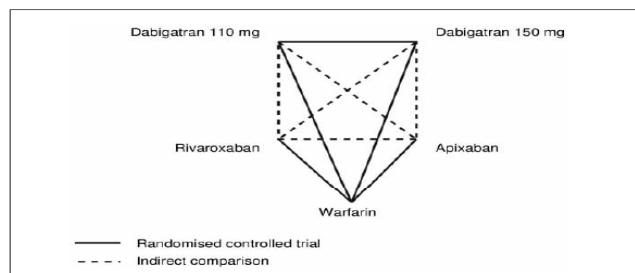


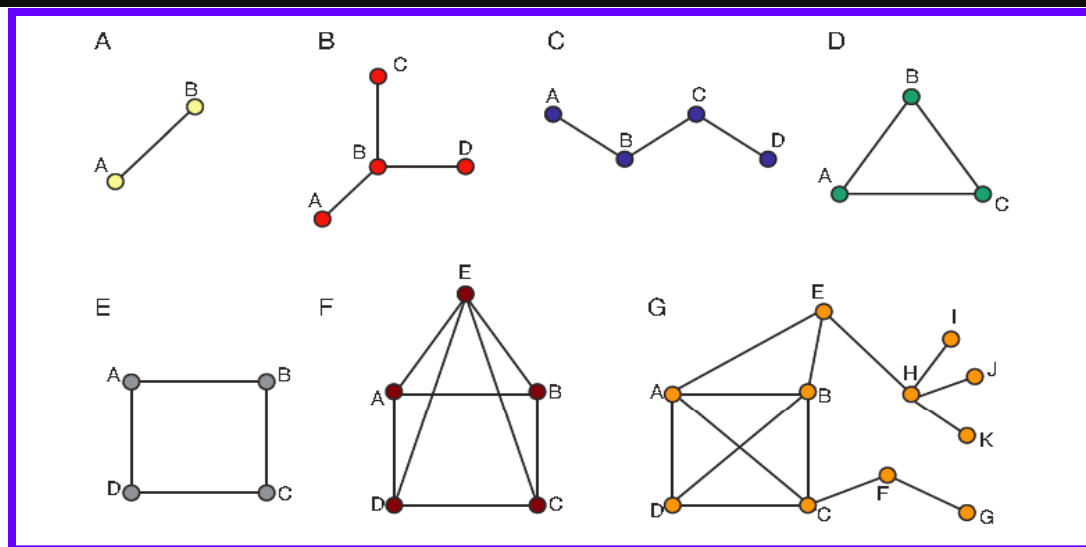
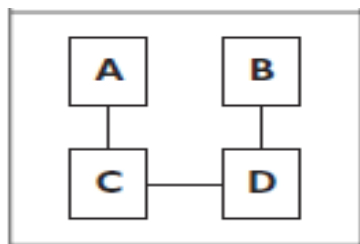
PARTE B/ PREGUNTAS DE DETALLE: GUÍA DE AYUDA PARA VALORAR UNA COMPARACIÓN INDIRECTA

Esta guía no es una relación exhaustiva de todos los aspectos a tener en cuenta en la CI, pretende poner de manifiesto algunos aspectos claves que nos permitan valorar críticamente este tipo de estudios. Lo ideal es que la mayoría de los ítem sean fiables o aceptables. Si algunos de los ítem es de fiabilidad baja, debemos cuestionarnos seriamente si utilizar este estudio.

ITEM	OPCIONES	FIABILIDAD	COMENTARIO
1 COMPARADOR COMUN Construir el gráfico del estudio. ¿Cuántos comparadores intermedios existen entre dos tratamientos a comparar A y B que se pretenden comparar indirectamente?	1	ALTA	Con más de 2 comparadores entre A y B, los intervalos de confianza (IC) de la diferencia estimada indirectamente son muy amplios y existe mucho riesgo al interpretar los resultados si no se consideran en todo momento los IC. La elaboración de un gráfico con todo los estudios incluidos, y aún mejor con los resultados de estos estudios proporciona una valiosa información para la interpretación de la evidencia disponible
	2	DUDOSA	
	más de 2	BAJA	
2 ESTRATEGIA DE BÚSQUEDA ¿Se utilizaron métodos sistemáticos para localizar todos los estudios relevantes? (incluyeron búsqueda en diferentes bases de datos así como la estrategia de búsqueda, siguieron las referencias, no usaron límites estrictos, valoraron sesgo de publicación, etc)	Si, cumple todo	ALTA	Para intentar localizar los artículos relevantes debería abordarse la búsqueda en diferentes bases de datos. Debe especificarse la estrategia para que pueda reproducirse y comprobar que cumple los criterios especificados.
	No se cumple totalmente pero se especifica el sesgo y como puede afectar	ACEPTABLE	
	No se cumple ni se indica el sesgo	BAJA	
3 CRITERIOS DE SELECCIÓN DE ESTUDIOS PARA LA COMPARACIÓN ¿Están definidos previamente? ¿Son razonables? ¿Se han aplicado de forma sistemática?	Si, cumple todo	ALTA	Una selección sesgada de estudios puede condicionar totalmente los resultados, por lo que los criterios de selección deben ser explícitos y estar definidos antes de la selección. Es importante poder comprobar que no hay una selección sesgada de los estudios, que se detalle si se han excluido estudios, el motivo, y este debe ser razonable. Si se ha realizado una revisión sistemática debería incluir un diagrama de flujo con la información localizada y seleccionada
	Alguno de estos puntos es mejorable, aunque aceptable	ACEPTABLE	
	No se cumple alguno de los puntos o se hace de forma inaceptable (sesgo importante)	BAJA	







Geometría de las redes

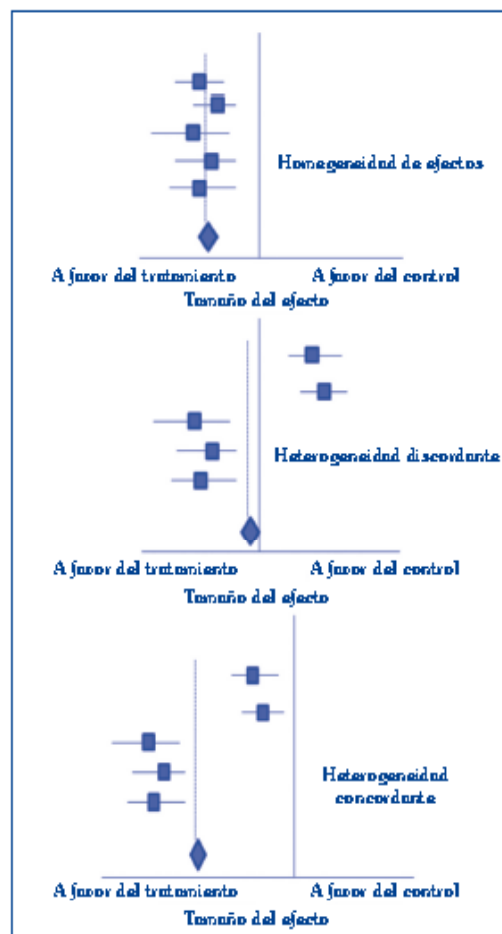
	Sitagliptin vs. placebo	Exenatide vs. placebo	Insulin vs. exenatide
Results of the head-to-head clinical trials: (absolute change in HbA1c)	0.57 (95% CI: 0.45, 0.72)	0.52 (95% CI: 0.32, 0.92)	0.40 (95% CI: 0.27, 0.60)
Adjusted indirect comparison: sitagliptin vs exenatide using placebo as common comparator	0.30 (95% CI: 0.17, 0.53)		
Adjusted indirect comparison: sitagliptin vs insulin using exenatide as common comparator	0.74 (95% CI: 0.40, 1.37)		



4	VARIABLE COMUN ¿Se utiliza en los estudios la misma variable definida de la misma forma (incluyendo mismo horizonte temporal) y presentada de la misma manera?	Si	ALTA	Comparar distintas variables o medidas en tiempos distintos, puede condicionar totalmente los resultados.
		No, pero la diferencia entre ambas variables es muy poco relevante. Se especifican y discuten las diferencias.	DUDOSA	
		No	BAJA	
5	RELEVANCIA DE LA VARIABLE La variable elegida para la comparación, ¿es clínicamente la más relevante posible? ¿es la variable principal de los estudios?	Si, es la más relevante y es la variable principal	ALTA	La elección de una variable poco relevante clínicamente cuando había otras opciones posibles constituye un sesgo inaceptable. En cambio, si la elección de una variable relevante, coincide con que sea la principal de los estudios, aporta una mayor validez interna. La selección de la variable debe estar justificada y la justificación ser razonable.
		Es la más relevante, pero no la principal de los estudios	ACEPTABLE	
		No es la más relevante posible, pero es razonable y es la mejor de la que se dispone	DUDOSA	
		Debería haberse elegido otra variable más relevante clínicamente, o bien no se disponía de ninguna variable medida con relevancia clínica suficiente	BAJA	
6	SIMILITUD CLÍNICA, SIMILITUD DE PACIENTES INCLUIDOS Mirar los estudios que se combinan: ¿los criterios de inclusión y exclusión son similares y razonables? Los grupos aleatorizados dentro de cada estudio: ¿eran comparables?	Si, con diferencias muy poco relevantes	ALTA	Es fundamental combinar poblaciones que sean, al menos en los factores que pueden modificar el efecto, similares. Se puede asumir, si es razonable que las diferencias no afectan ni al efecto relativo ni al riesgo basal. En caso de que afectaran al riesgo basal se podría intentar hacer análisis de subgrupos por riesgo basal, y si afecta al efecto relativo hacer análisis de subgrupos basado en variables que modifican el efecto, solo si clínicamente está justificado.
		Existen diferencias, pero la mayor parte de los pacientes son similares en los diversos estudios. Se compara un subgrupo y los pacientes de los subgrupos que se comparan son similares.	ACEPTABLE	
		Existen muchas diferencias de población o en los subgrupos comparados entre los estudios.	BAJA	
7	COMPARADOR COMUN El comparador común, ¿está usado en la misma posología y con la misma duración del tratamiento, y se permiten los mismos tratamientos concomitantes que pueden afectar al efecto?	Si	ALTA	La combinación de estudios con comparadores comunes que no son realmente iguales puede comprometer totalmente la validez de la comparación, si hay dudas que el efecto sea diferente habría que considerarlas como alternativas diferentes.
		Existen diferencias con muy poca influencia potencial en el resultado	ACEPTABLE	
		Existen diferencias relevantes	BAJA	



Figura 4. Homogeneidad y tipos de heterogeneidad



Fraga MD et al. Boletín SESCAM 2013

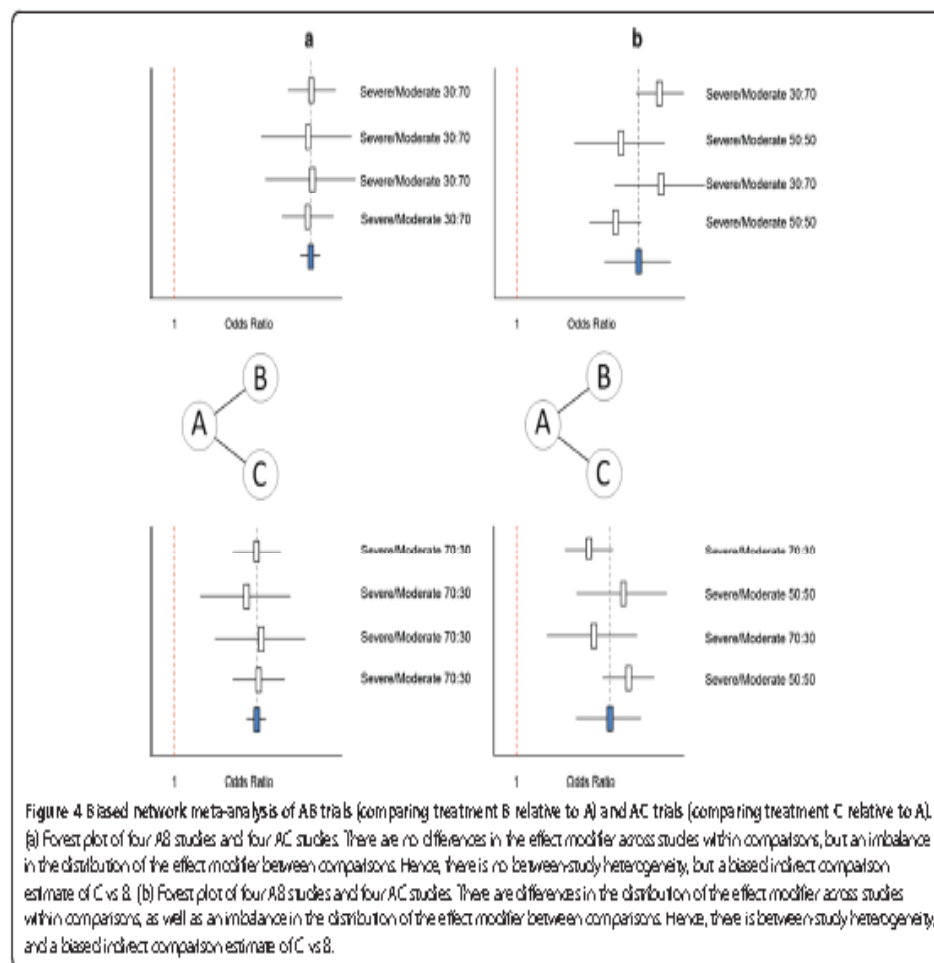


Figure 4 Biased network meta-analysis of AB trials (comparing treatment B relative to A) and AC trials (comparing treatment C relative to A). (a) Forest plot of four AB studies and four AC studies. There are no differences in the effect modifier across studies within comparisons, but an imbalance in the distribution of the effect modifier between comparisons. Hence, there is no between-study heterogeneity, but a biased indirect comparison estimate of C vs B. (b) Forest plot of four AB studies and four AC studies. There are differences in the distribution of the effect modifier across studies within comparisons, as well as an imbalance in the distribution of the effect modifier between comparisons. Hence, there is between-study heterogeneity, and a biased indirect comparison estimate of C vs B.

Jansen and Naci. BMC Medicine 2013



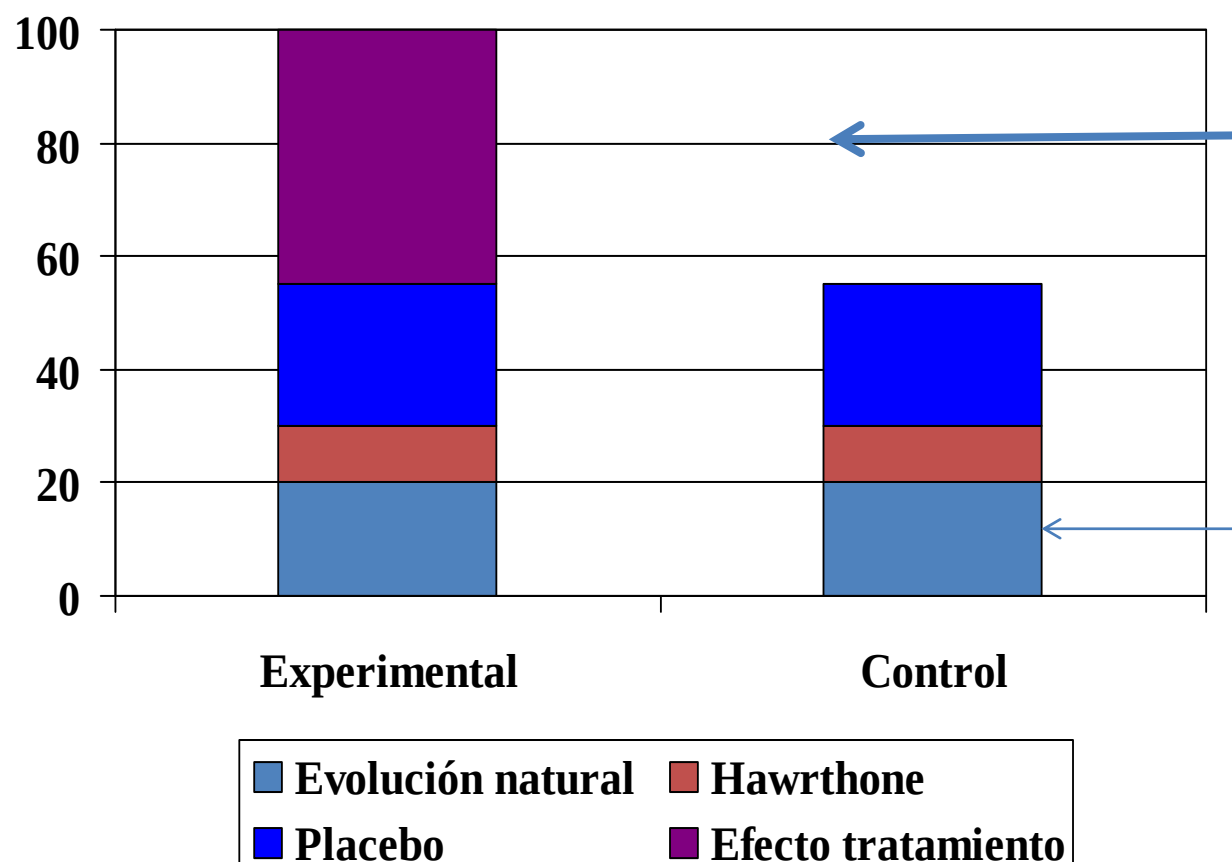
8	RESULTADO CUANTITATIVO DEL GRUPO CONTROL ¿Es similar en los estudios que se combinan en la CI?	Si. O no existen diferencias, o claramente carecen de relevancia clínica.	ALTA	Para poder combinar los estudios, el resultado del grupo control debería ser similar. No obstante, habría que buscar la causa y si las poblaciones son comparables, una diferencia en el grupo control puede paliarse por un análisis ajustado por el resultado del control (para ampliar información consultar ICWG Report http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\$File/ICWG%20Report%20FINAL2.pdf) o un análisis de subgrupos como se ha indicado en la pregunta 6.
		No. Existen diferencias que podrían ser clínicamente relevantes, pero resultan aceptables	ACEPTABLE	
		No. Existen diferencias que ponen en evidencia que las poblaciones y/o la intervención y/o la medida de resultados probablemente presentan diferencias importantes.	BAJA	
9	CALIDAD DE LOS ESTUDIOS COMBINADOS ¿Cuál es la calidad/validez interna de los estudios combinados? ¿Se dan los resultados de los estudios individuales (la transparencia es imprescindible para valorar la aplicabilidad)?	Los estudios que se combinan tienen baja probabilidad de sesgos, y son ensayos clínicos bien diseñados, o meta-análisis directos de ensayos clínicos homogéneos	ALTA	Los estudios observacionales tienen <i>per se</i> , una elevada probabilidad de sesgos. Combinarlos, además, con un comparador común, los multiplica y condiciona una baja validez interna. La transparencia en los estudios de partida es imprescindible para valoración de la comparación indirecta (se deberían dar no solo los valores relativos sino también los datos reales). Para valorar seguridad los estudios observacionales prospectivos pueden aportar datos más reales.
		Los estudios que se combinan incluyen algún estudio de cohortes con baja probabilidad de sesgos. Pero existe al menos un ensayo clínico bien diseñado de cada fármaco que se compara, y el peso de los estudios de cohortes (en nº de pacientes) no supera al de los ensayos clínicos para cada fármaco.	DUDOSA	
		No se puede incluir en ninguno de los dos perfiles anteriores	BAJA	



10	TRATAMIENTO DE LAS PERDIDAS En los estudios que se han combinado, ¿las pérdidas son similares y aceptables? ¿se trataban de la misma manera? ¿el tipo de análisis es el mismo (por protocolo, por intención de tratar-ITT-)?	Si, cumple todo, o bien existen diferencias irrelevantes	ALTA	Un tratamiento distinto de los pacientes que abandonan el estudio, si éstos son numerosos, puede condicionar en gran medida los resultados. Puede paliarse parcialmente mediante el análisis por intención de tratar y sería conveniente realizar un análisis de sensibilidad con los distintos escenarios
		Falta información, o bien existen diferencias relevantes en alguno de estos ítems, pero no anulan completamente la validez de la comparación. Sin embargo, el análisis es ITT en todos los estudios que se combinan.	DUDOSA	
		No se puede incluir en ninguno de los dos perfiles anteriores	BAJA	
11	FIABILIDAD DE LOS DATOS DE PARTIDA ¿Se han sometido a algún tipo de extrapolación/manipulación previa?	Los datos coinciden con los publicados	ALTA	La manipulación o extrapolación de datos supone ya una merma de validez que además se multiplica por la comparación indirecta.
		No coinciden, se han manipulado o extrapolado previamente	BAJA	
12	AJUSTE POR GRUPO CONTROL ¿Se trata de una comparación ajustada por el grupo control (método de Bucher, Network meta-análisis, etc.), es decir se comparan los efectos relativos respecto al grupo control de cada estudio?	Si	ALTA	Es imprescindible tener en cuenta los resultados en los grupos control para poder comparar los tratamientos.
		No	BAJA	
13	ANÁLISIS DE DATOS/ESTADÍSTICO El análisis ¿es adecuado? ¿son razonables las asunciones realizadas? Si es necesario el análisis de sensibilidad ¿es adecuado?	Si (Si se usan métodos bayesianos, están claramente descritos y las asunciones son razonables). Cuando solo hay 2 alternativas comparadas vs un comparador común con un ensayo cada una, el método de Bucher es adecuado.	ALTA	Las simulaciones a menudo se basan en asunciones subjetivas, y su validez interna es menor. La superposición de intervalos de confianza es una práctica errónea. Si se utilizan métodos bayesianos se debe describir los valores a priori utilizados, el número de iteraciones, la convergencia. Si hay varios estudios para una misma comparación se debe cuantificar la heterogeneidad entre ellos. Debería existir una descripción de los diferentes escenarios en el análisis de sensibilidad. Son más adecuados los análisis de sensibilidad probabilísticos en que se incluyen los parámetros relevantes.
		No.	BAJA	



¿Para que queremos un grupo de comparación?



Modificadores del efecto:

características de los pacientes del estudio que influyen en el efecto del tratamiento

Factores pronósticos:

características del estudio y de los pacientes que influyen en los resultados del brazo de la intervención y del brazo placebo de la misma manera



14	SIMILITUD DE LOS ESTUDIOS ¿Se detecta alguna otra característica diferencial de los estudios que se combinan, que haga inaceptable la comparación de los resultados?	No	ALTA	Puede haber otras causas de heterogeneidad específicas, no incluidas en ítems anteriores, que dificulten la comparabilidad de los distintos estudios utilizados.
		Se han detectado otras diferencias entre los estudios, pero no condicionan los datos de forma relevante	ACEPTABLE	
		Se ha detectado alguna diferencia que condiciona ampliamente los datos y hace incomparables los resultados de los estudios.	DUDOSA	
		Si	BAJA	
15	Contestar sólo si están publicadas comparaciones directas ¿EXISTE INCONSISTENCIA? ¿Se observan diferencias en las estimaciones de los efectos de los tratamientos entre las comparaciones directas e indirectas?	No. Se ha testado la inconsistencia y no es significativa. Existen métodos para testar la inconsistencia como la diferencia estandarizada o mediante métodos bayesianos como el modelo de inconsistencia o el node splitting.	ALTA	Debe distinguirse claramente si se observan diferencias significativas entre la comparación directa e indirecta. Si hay inconsistencia es importante estudiar la causa (características de los pacientes, características de los ensayos, otros factores...). Si la comparación indirecta está sesgada es porque lo están las comparaciones directas de las que parte o porque no hay similitud clínica entre los estudios combinados. En general, suelen estar menos sesgadas las comparaciones

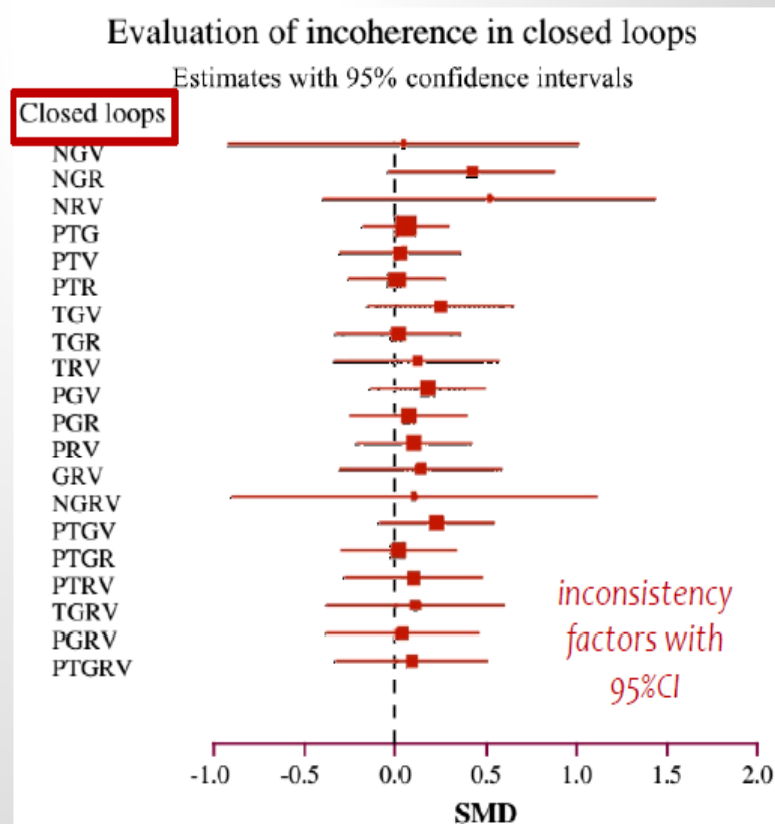


GRAPHS FOR INCONSISTENCY

in Network Meta-Analysis

Forest plot with inconsistency factors estimated by the loop- specific approach for inconsistency

[Example in Salanti et al. 2009]



GRAPHS FOR INCONSISTENCY

in Network Meta-Analysis

'Net-heat plot': a plot showing the contribution of the direct estimates to the network estimates and the change in inconsistency when detaching one comparison from the network

[Example in Krahn et al. 2013]

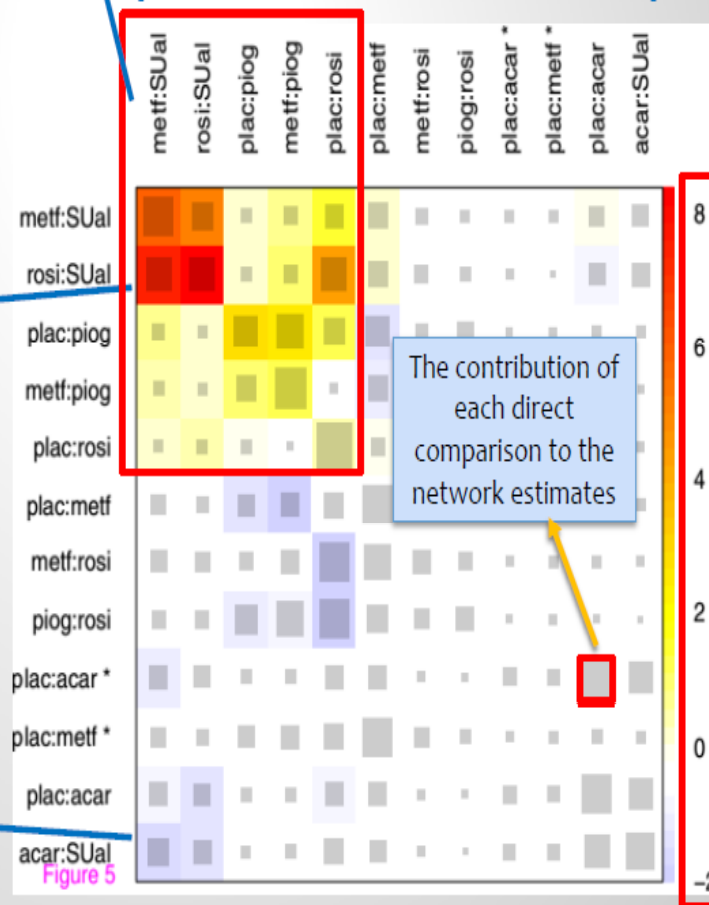
decrease in
inconsistency after
detaching a
comparison

observed
comparisons in the
network

increase in
inconsistency after
detaching a
comparison

possible important
sources of
inconsistency

detached comparisons



16	APLICABILIDAD Y VALIDEZ EXTERNA ¿Las conclusiones se deducen de los resultados? ¿Se pueden generalizar estos resultados a nuestra población?	Si	ALTA	Deben considerarse las posibles diferencias con la población a la que queremos aplicar estos resultados y considerar la extrapolación que estamos dispuestos a admitir
		No	BAJA	
17	CONFLICTO DE INTERESES ¿Existe declaración de conflicto de intereses?	Está declarado el conflicto de intereses y no se observa manipulación	ALTA	Una comparación indirecta está sometida a un riesgo de sesgos y manipulaciones, se debe valorar si el conflicto de intereses ha influido en la comparación
		Falta declaración de intereses o se observa manipulación ligada al conflicto de intereses.	BAJA	



How to Use an Article Reporting a Multiple Treatment Comparison Meta-analysis

Edward J. Mills, PhD, MSc

John P. A. Ioannidis, MD, DSc

Kristian Thorlund, PhD, MSc

Holger J. Schünemann, MD, PhD, MSc

Milo A. Puhan, MD, PhD

Gordon H. Guyatt, MD, MSc

CLINICAL SCENARIO

You are seeing a 45-year-old patient for whom, 6 weeks previously, you prescribed paroxetine, a selective serotonin reuptake inhibitor (SSRI), for treatment of generalized anxiety disorder (GAD). The patient reports reduced anxiety, but also insomnia and a reduced interest in sex. You wonder if there is another drug the patient might tolerate better while still maintaining treatment response. You retrieve a multiple treatment comparison (MTC) meta-analysis that evaluates response and tolerability of all available GAD drugs.¹ You are not familiar with the type of study, and you wonder if there are special issues to which you should attend in evaluating its methods and results.

MTC META-ANALYSES

Traditionally, a systematic review addresses the merits of one intervention vs another (eg, placebo, or another active intervention). Data are combined from all studies—often randomized clinical trials (RCTs)—that meet eligibility criteria in what we will term a pairwise meta-analysis. Compared with single RCTs, meta-analysis improves the power to detect differences and also facilitates the examination of the extent to which there are important differences in treatment effects across eligible RCTs—variability that is frequently called heterogeneity.^{2,3} Large,

Multiple treatment comparison (MTC) meta-analysis uses both direct (head-to-head) randomized clinical trial (RCT) evidence as well as indirect evidence from RCTs to compare the relative effectiveness of all included interventions. The methodological quality of MTCs may be difficult for clinicians to interpret because the number of interventions evaluated may be large and the methodological approaches may be complex. Clinicians and others evaluating an MTC should be aware of the potential biases that can affect the interpretation of these analyses. Readers should consider whether the primary studies are sufficiently homogeneous to combine; whether the different interventions are sufficiently similar in their populations, study designs, and outcomes; and whether the direct evidence is sufficiently similar to the indirect evidence to consider combining. This article uses the existing Users' Guides format to address study validity, interpretation of results, and application to a patient scenario.

JAMA. 2012;308(12):1246-1253

www.jama.com

unexplained heterogeneity may reduce a reader's confidence in estimates of treatment effects.

A drawback of pairwise meta-analysis is that it evaluates the effects of only 1 intervention vs 1 comparator and does not permit inferences about the relative effectiveness of several interventions unless all have been compared directly in head-to-head trials. Yet, for many medical conditions, there are numerous available interventions that have—unfortunately—most frequently been compared with placebo and seldom with one another.^{4,5} For example, despite 91 completed and ongoing RCTs addressing the effectiveness of drugs for the treatment of rheumatoid arthritis, only 5 compare directly against each other.⁶

Recently, another form of meta-analysis, the MTC meta-analysis (also known as network meta-analysis because it involves creating a network of treatments), has emerged.^{7,8} The MTC

approach provides estimates of effect sizes for all possible pairwise comparisons whether or not they have been compared head-to-head in RCTs. FIGURE 1 displays examples of common networks of treatments. The eAppendix (available at <http://www.jama.com>) provides a glossary of common nomenclature found in MTCs.

When 2 interventions, A and B (eg, paroxetine and citalopram in FIGURE 2A),

Author Affiliations: Faculty of Health Science, University of Ottawa, Ottawa, Ontario, Canada (Dr Mills); Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (Dr Mills, Thorlund, Schünemann, and Guyatt); Stanford Prevention Research Center, Department of Medicine and Health Research and Policy, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California (Dr Ioannidis); Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Dr Puhan).
Corresponding Author: Edward J. Mills, PhD, MSc, Faculty of Health Science, University of Ottawa, Room 001, Thompson Hall, 35 University Private, Ottawa, ON K1N 6N4, Canada (edward.mills@uottawa.ca).
Users' Guides to the Medical Literature Section Editor: Dr David Rennie, MD, Deputy Editor, JAMA.

Box 1. Critical Appraisal Guide to a Multiple Treatment Comparison Meta-analysis

A. Are the results of the study valid?

Did the review explicitly address a sensible clinical question?

Was the search for relevant studies exhaustive?

Were there major biases in the primary studies?

B. What are the results?

What was the amount of evidence in the network?

Were the results similar from study to study?

Were the results consistent in direct and indirect comparisons?

What were the overall treatment effects and their uncertainty, and how did the treatments rank?

Were the results robust to sensitivity assumptions and potential biases?

C. How can I apply the results to patient care?

Were all patient-important outcomes considered?

Were all potential treatment options considered?

Are any postulated subgroup effects credible?

What is the overall quality and what are limitations of the evidence?

JAMA. 2012;308(12):1246-1253





ELSEVIER

available at www.sciencedirect.com

SCIENCE @ DIRECT®

journal homepage: www.elsevier.com/locate/jval

SCIENTIFIC REPORT

Interpreting Indirect Treatment Comparisons and Network Meta-Analysis for Health-Care Decision Making: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 1

Jeroen P. Jansen, PhD^{1,*}, Rachael Fleurence, PhD², Beth Devine, PharmD, MBA, PhD³, Robbin Itzler, PhD⁴,
Annabel Barrett, BSc⁵, Neil Hawkins, PhD⁶, Karen Lee, MA⁷, Cornelis Boersma, PhD, MSc⁸, Lieven Annemans, PhD⁹,
Joseph C. Cappelleri, PhD, MPH¹⁰

¹Mapi Values, Boston, MA, USA; ²Oxford Outcomes, Bethesda, MD, USA; ³Pharmaceutical Outcomes Research and Policy Program, School of Pharmacy, School of Medicine, University of Washington, Seattle, WA, USA; ⁴Merck Research Laboratories, North Wales, PA, USA; ⁵Eli Lilly and Company Ltd., Warrington, Surrey, UK; ⁶Oxford Outcomes Ltd., Oxford, UK; ⁷Canadian Agency for Drugs and Technologies in Health (CADTH), Ottawa, ON, Canada; ⁸University of Groningen / HECTA, Groningen, The Netherlands; ⁹University of Ghent, Ghent, Belgium; ¹⁰Pfizer Inc., New London, CT, USA

Table 1 – Simplified checklist to assist decision makers in evaluating a reported network meta-analysis.

Report section	Checklist Item	What to look for in the paper
Introduction	Are the rationale for the study and the study objectives stated clearly?	A clear rationale for the review A clear objective or research question that pertains to the network meta-analysis
Methods	Does the methods section include the following? Description of eligibility criteria Information sources Search strategy Study selection process Data extraction (validity/quality assessment of individual studies)	A systematic review of the literature in accordance with Centre for Reviews and Dissemination (CRD) guidelines and PRISMA [52,54,55]
	Are the outcome measures described?	Justification of outcome measures selected for analysis
	Is there a description of methods for analysis/synthesis of evidence? Do the methods described include the following? Description of analyses methods/models Handling of potential bias/inconsistency Analysis framework	Description and justification of statistical model(s) used: multiple meta-analysis of pairwise comparisons vs. network meta-analysis models; fixed- vs. random-effects models; models without or with covariate (interactions) Description of whether analyses were performed with a frequentist or Bayesian approach Description of how possible bias/inconsistency was evaluated (either qualitative or quantitative, e.g., comparison of direct evidence with the indirect evidence). If meta-regression models are used, rationale for selection of covariates in models Description of relative-effect estimates used for presentation of findings (e.g., odds ratio, relative risk, hazard ratio, difference in change from baseline) Description of whether relative-effect measures were transformed into expected (absolute) outcomes (e.g.,



Sección	Pregunta
1. Introducción	¿Está justificada suficientemente la realización de un metaanálisis en red y se describen los objetivos de forma clara y precisa?
2. Material y métodos	¿Son relevantes los objetivos para la práctica clínica?
	¿Se especifican claramente la población diana, los tratamientos evaluados (competidores) y la(s) variable(s) de resultado del efecto de los tratamientos?
	¿Se ha realizado una revisión sistemática (siguiendo la declaración PRISMA)?
	¿Se describen con detalle los métodos de búsqueda, las estrategias de búsqueda complementarias, los criterios de inclusión/exclusión y los métodos de extracción de estudios?
	¿Se tienen en cuenta la calidad y validez de los artículos siguiendo criterios explícitos y reproducibles?
	¿Se discute la transitividad?
	¿Se describen claramente los métodos utilizados para los análisis estadísticos incluyendo la justificación del modelo estadístico para comparaciones directas (por ejemplo, modelos de efectos fijos frente a efectos aleatorios) e indirectas (por ejemplo, metarregresión, modelos bayesianos, modelos de efectos fijos frente a efectos aleatorios)?
	¿Se presentan estimaciones de comparaciones directas, indirectas o la combinación de ambas (si es posible y apropiado)?
	¿Se presentan los efectos de los tratamientos (por ejemplo, riesgo relativo combinado) para cada alternativa evaluada frente al tratamiento de referencia de acuerdo con la red de evidencia definida?
	¿Se describen métodos para abordar posibles sesgos (por ejemplo, selección, publicación)?
	¿Se describen los métodos para evaluar la heterogeneidad y la (in)consistencia?
	¿Cómo se aplicaron todos estos criterios? ¿Se establecieron a priori?
	¿Se presentaron análisis de sensibilidad para valorar el efecto de la incertidumbre?
3. Resultados	¿Se presentan los resultados de los estudios en una red de evidencia (con información sobre el número de estudios y el número de pacientes para cada comparación)? ¿Se presenta una figura con la red de evidencia?
	¿Se presenta (en un anexo, si es necesario) la información de cada estudio por separado?
	¿Se presentan los resultados del metaanálisis en red de forma clara, objetiva y transparente?
	¿Se discuten las limitaciones?
4. Discusión	¿Se hace una descripción o resumen de los principales hallazgos?
	¿Se discuten las limitaciones del metaanálisis en red?
	¿Se discute la validez interna y externa de los análisis?
	¿Se discuten las consecuencias para la toma de decisiones en la práctica clínica?
	¿Se justifican de manera razonada las conclusiones del trabajo?
5. Otros	¿Se declaran los conflictos de interés y la fuente de financiación del trabajo?

Catalá-López F, Tobías A. Med Clin (Barc): 2013;140(4):182-7



The Quality of Reporting Methods and Results in Network Meta-Analyses: An Overview of Reviews and Suggestions for Improvement

. PLoS ONE 9(3): e92508.

doi:10.1371/journal.pone.0092508

Table 2. Summary of Considerations for Methodologic Reporting.

Literature Search:

- Specification of efforts to search for direct and indirect evidence of relevance
- Presentation of search terms and the full search strategy (or strategies if separate searches undertaken for each comparison) for one electronic database
- Provision of information regarding involvement of primary literature searching versus use of existing systematic reviews as a means of including studies
- If existing reviews were used, specification of how these were located and description of their related inclusion criteria

Specification of Eligibility Criteria (and Planned Network Structure):

- Specification of PICOS eligibility criteria for the review, including specification of all treatments included in the planned meta-analysis
- Specification of how related but different implementations of the same agent (e.g. varied doses of pharmacologic treatments) are to be handled with associated rationale (i.e. address 'lumping and splitting' of interventions)

Specification of Assumptions for Network Meta-Analysis:

- Specification of the assumptions of homogeneity, similarity and consistency (or related terminology used, e.g. transitivity, exchangeability)
- Specification of efforts taken by reviewers to evaluate the assumption appropriateness
- Specification of what information is being provided to readers to allow them to consider the validity of the assumptions

Statistical Approach to Analysis:

- Specification of details of the approach to statistical analysis taken: hierarchical model, adjusted indirect approach or meta-regression, Frequentist or Bayesian framework, fixed or random effects model, homogeneous or heterogeneous between-trial variance structure
- Specification of methods used to assess the degree of statistical heterogeneity and the potential for publication bias within the treatment network
- Specification of methods used to evaluate for the presence of disagreement between direct and indirect evidence in the treatment network
- Description of statistical methods used to address clinical and methodologic homogeneity in the analyses (e.g. subgroups, meta-regression including adjustments for baseline risk and the impact of risk of bias variations)

doi:10.1371/journal.pone.0092508.t002



The Quality of Reporting Methods and Results in Network Meta-Analyses: An Overview of Reviews and Suggestions for Improvement

. PLoS ONE 9(3): e92508.
doi:10.1371/journal.pone.0092508

Table 3. Summary of Considerations for Reporting and Interpretation of Results.

Presentation of Evidence and Its Characteristics in the Treatment Network

- Presentation of network diagram to summarize identified evidence
- Reporting information reflecting the amount of information in the network, e.g. sample sizes, numbers of studies per comparison and the presence of multi-arm studies
- Presentation of information allowing readers to assess clinical and methodological heterogeneity within the treatment network: e.g. information tables listing effect modifiers across studies and comparisons, These can include patient characteristics and risk of bias assessments

Assessment of Assumptions for Network Meta-Analysis

- Information to summarize evaluations of statistical heterogeneity within the treatment network
- Information and approach to summarize analyses to assess agreement of direct and indirect sources of evidence (and efforts to improve agreement if discrepancies are found)

Presentation of Summary Treatment Effect Estimates and Related Measures

- What estimates to report: all possible pairwise comparisons? Only those which are comparisons against the chosen reference group or a treatment of primary focus?
- Should findings from traditional pairwise analyses also be provided?
- Presentation of summary estimates and corresponding uncertainty (i.e. credible/confidence intervals)
- Presentation of summary estimates from sensitivity and subgroup analyses
- Optimal use of tables and figures to most easily convey results to readers
- Presentation of treatment rankings and corresponding probabilities: Should they be included? What should be presented?

Discussion and Conclusions

- Commentary on the clinical and biologic plausibility of the observed findings
- Commentary relevant to any important concerns regarding the assumptions underlying the joint synthesis that may play an important role in strength of interpretations drawn

doi:10.1371/journal.pone.0092508.t003



CI propuesta metodológica proporciona estimaciones del efecto de los tratamientos respecto a las alternativas disponibles

¿Están incluidos todos los tratamientos relevantes?

¿Están incluidos todos los ensayos relevantes?
(nuevo vs nuevo, nuevo vs estándar, estándar vs estándar, estándar vs antiguo, antiguo vs antiguo), considerar la posibles parejas para las comparaciones indirectas. Hacer gráficos.

Herramientas de ayuda:

Declaración PRISMA, análisis Forest Plot, calculadoras (Wells...)

¿Es más adecuada una CI ajustada o una mezcla de múltiples tratamientos?

¿Existe heterogeneidad o inconsistencia?

Nuevas herramientas en desarrollo
Extensión del PRISMA, COPAS (sesgo de publicación)
Gráficos para inconsistencia



analyses (91) analysis (67) characteristics (46) clinical (53)
 cochrane (34) combined (32) comparisons (33) consider (33) continuous (42) control (35)
 data (161) dichotomous (45) difference (156)
 effect (244) error (41) estimate (45) events (114)
 example (65) fixed-effect (35) heterogeneity (107)
 important (36) intervention (159) investigate (56) log (30)
 measures (44) meta-analyses (34) meta-analysis (168)
 meta-regression (31) methods (75) number (60) observed (43)
 odds (106) outcome (74) participants (51) random-
 effects (48) ratio (168) results (104) review (103)
 revman (37) risk (183) scales (31) section (114)
 standard (86) statistical (110)
 studies (218) subgroup (52) summary (55) systematic (34)
 treatment (37) trials (48) value (31)

