Generalità e requisiti

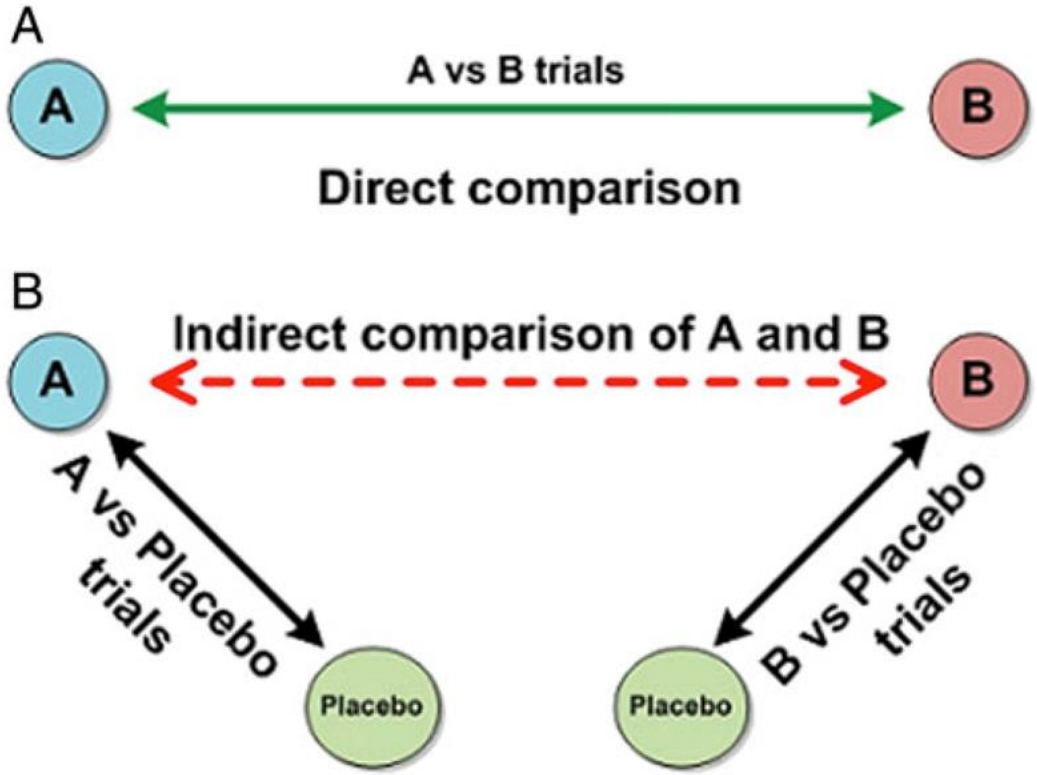
(G.L. Pappagallo)

Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.



through a

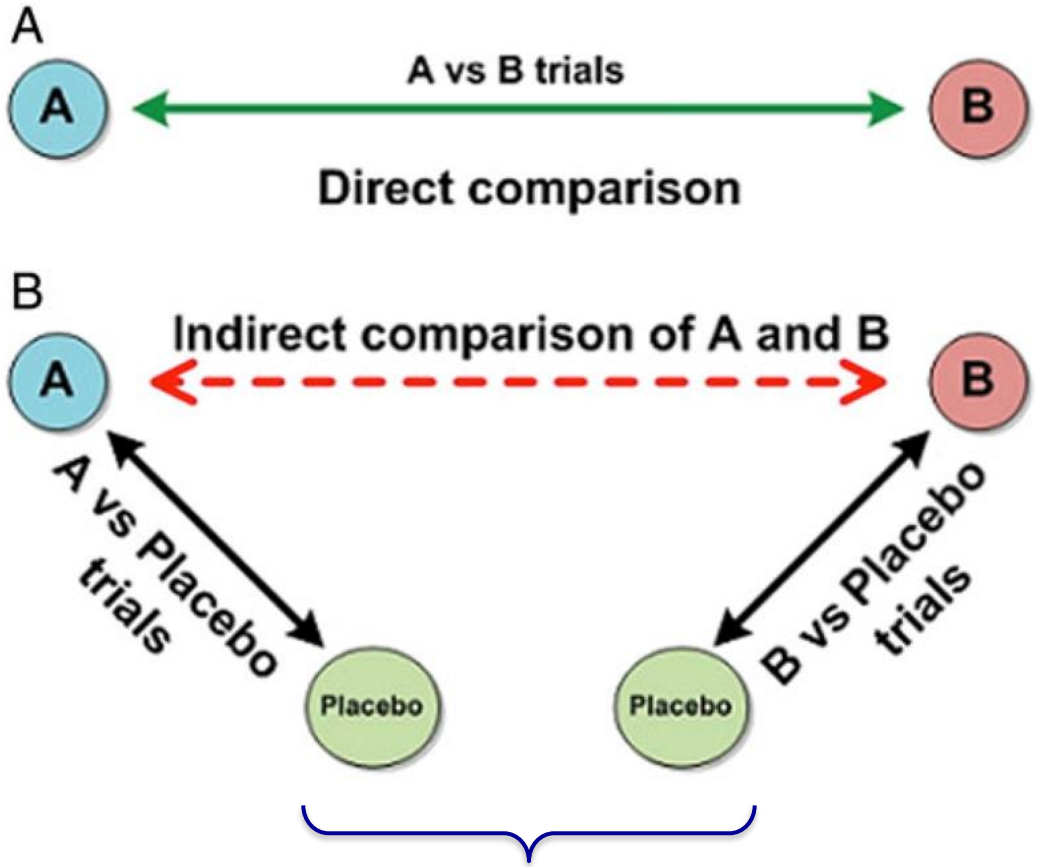
Common Comparator

Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.



Similarity Assumption

trials must be comparable on effect modifiers to obtain an unbiased pooled estimate.

Indirect comparisons of competing interventions

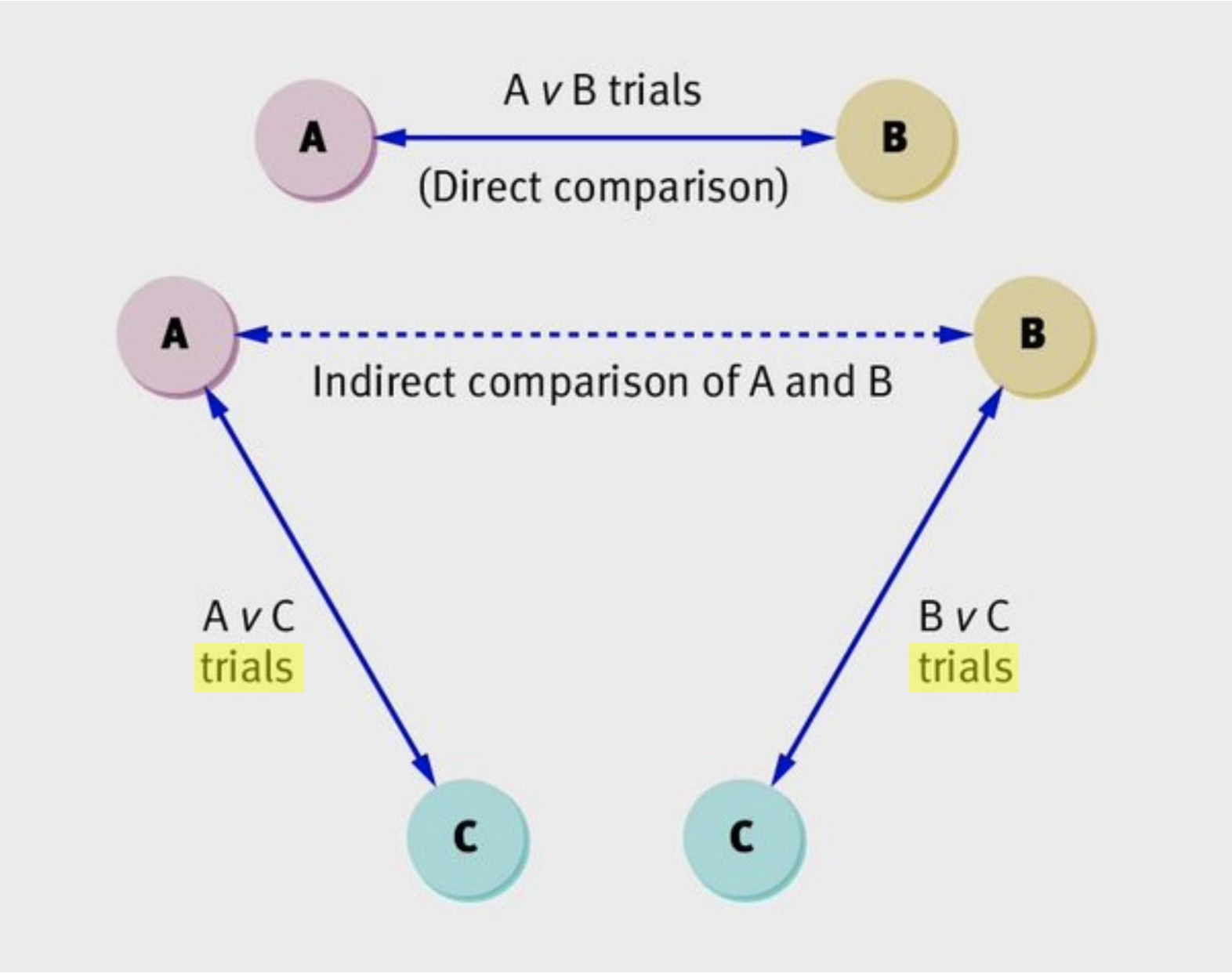
AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

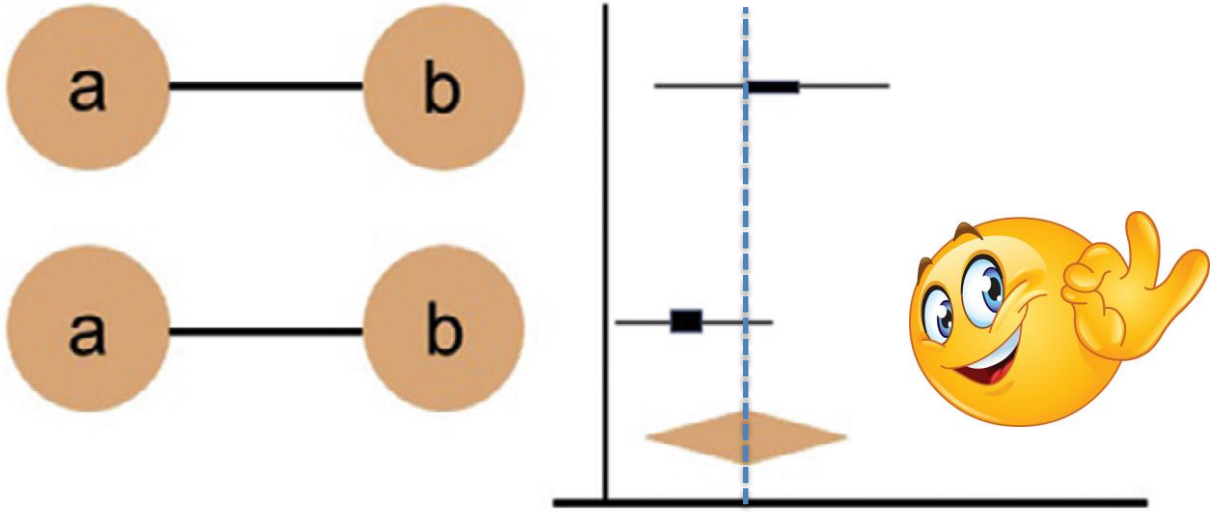
When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.

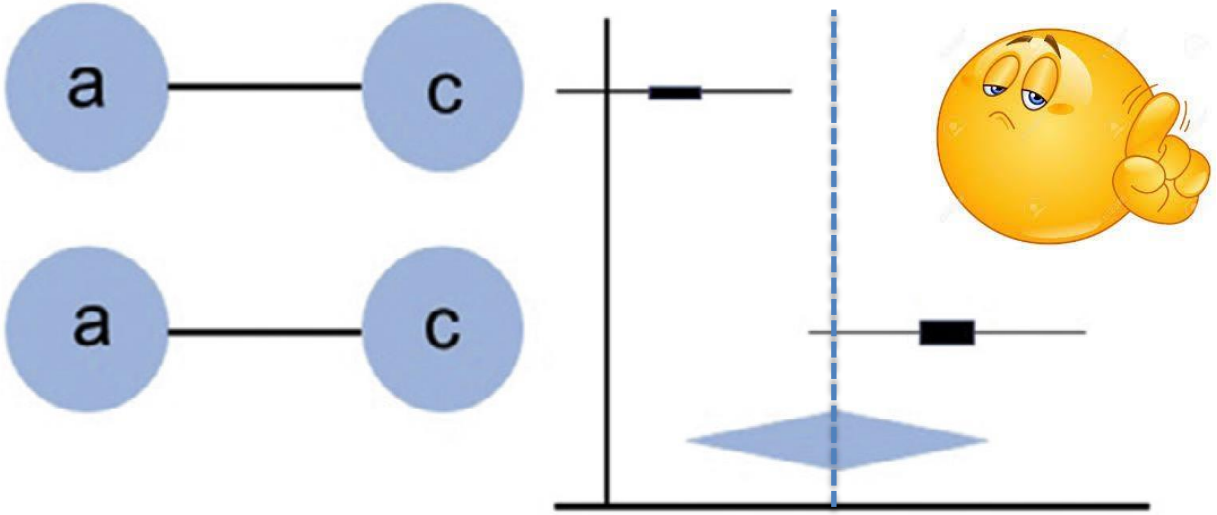
WHAT FACTORS DETERMINE SIMILARITY?

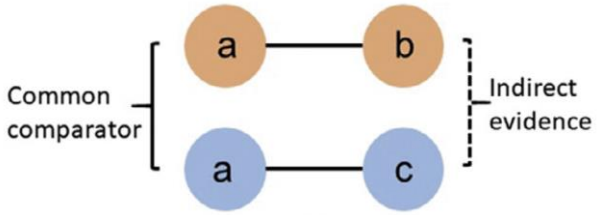
- Included trials should be “comparable” in terms of key factors that could affect the outcome of treatment
- If differences in patient or study characteristics would not be expected to influence treatment effect, the assumption of similarity is not violated
- **There are no statistical methods to test for similarity**
- Must use clinical knowledge and best judgement to assess appropriate comparability



Quando
le evidenze dirette
sono costituite
da più trials…

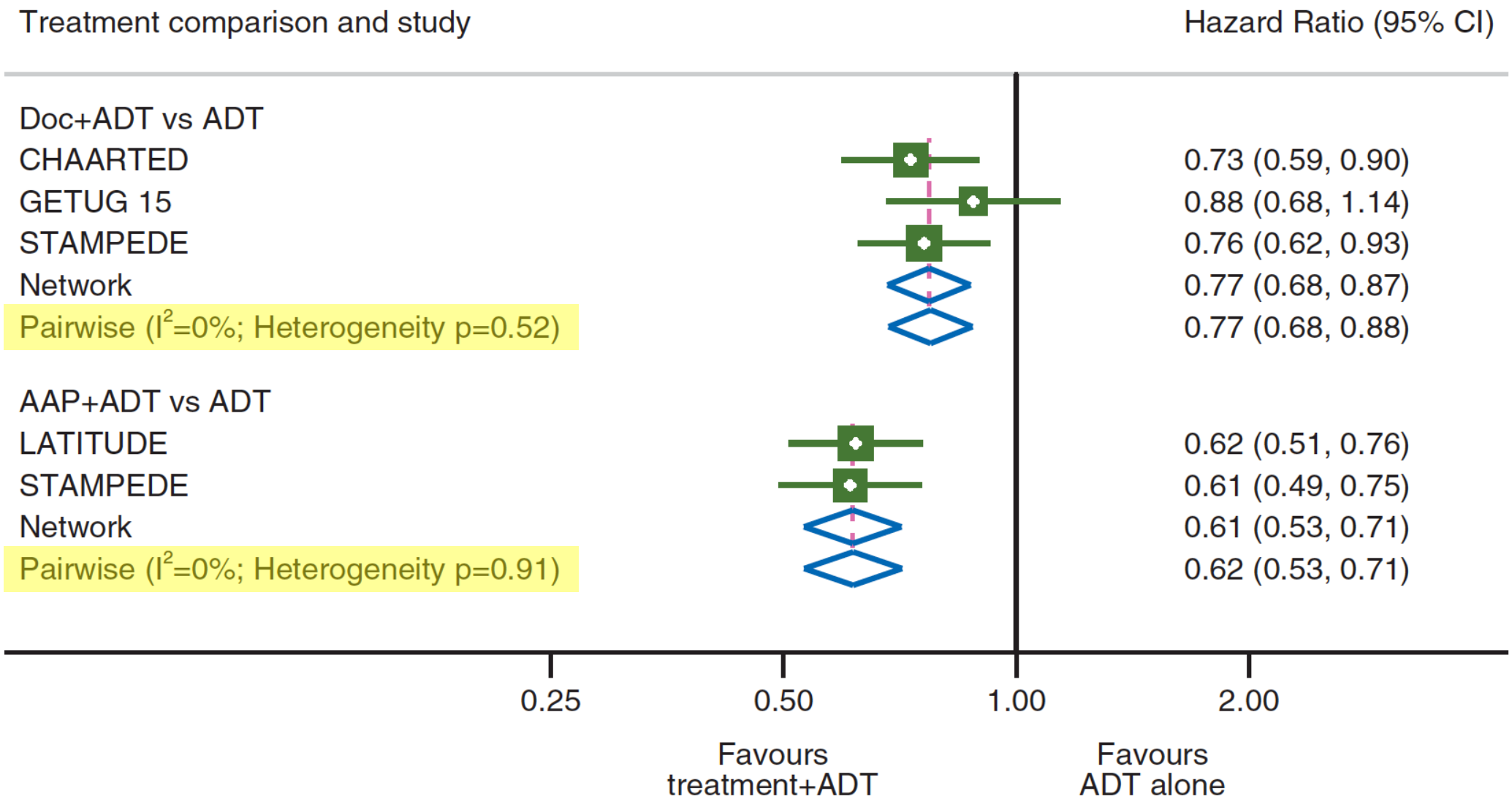






Homogeneity Assumption

there must be
no relevant heterogeneity
between trial results in
pairwise comparisons



Annals of Oncology 29: 1249–1257, 2018

Commonly applied methods

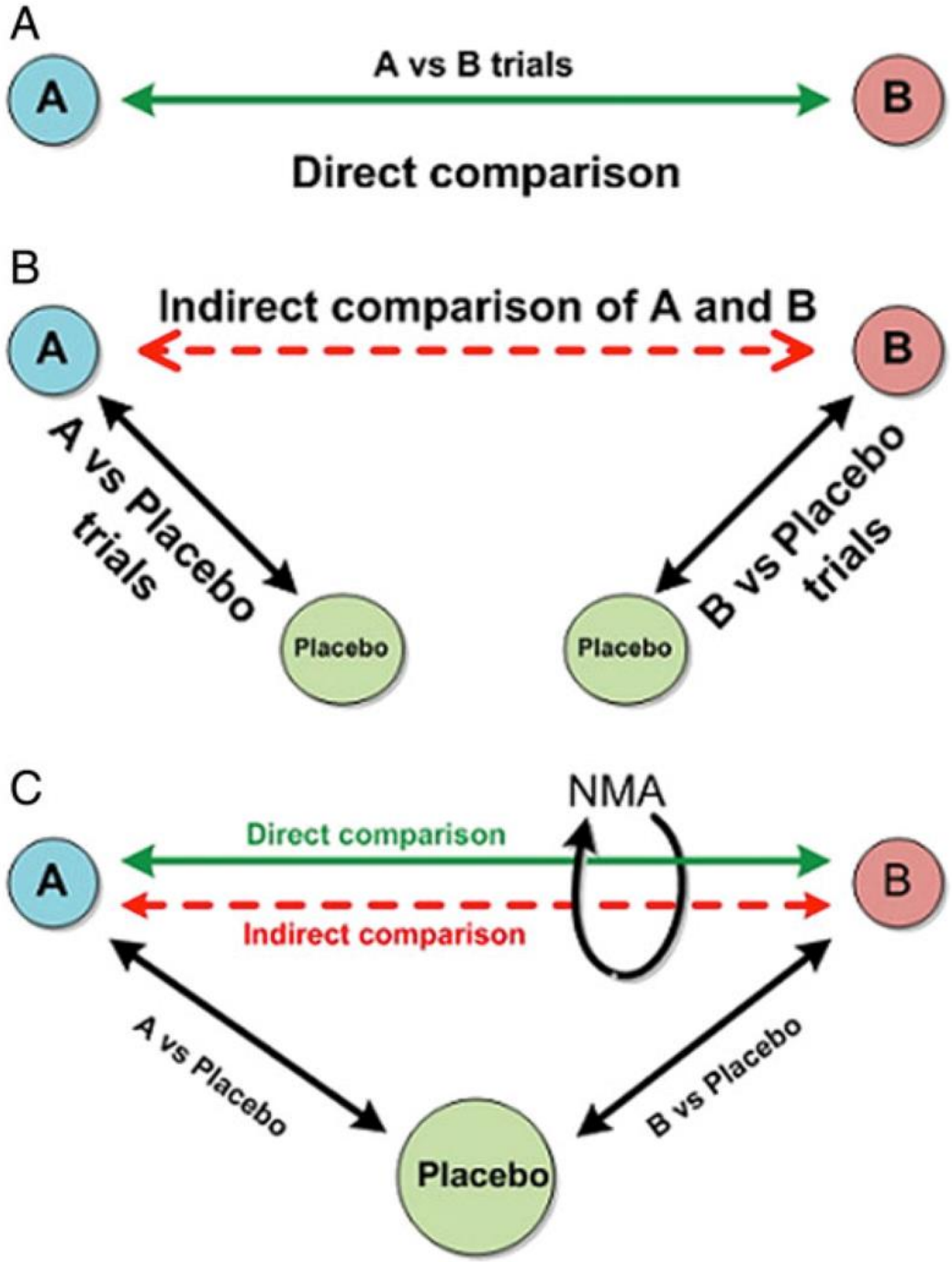
- **ITC (Bucher)**
  - IPD not required
  - treatment effects calculated for each trial separately
  - within study randomization preserved
- **Population-adjusted indirect comparison (MAIC)**
  - IPD required for at least 1 trial
  - to match the IPD to the AgD of the other trial
- **Network Meta-Analysis (NMA)**
  - comparing interventions simultaneously in a single analysis by combining both direct and indirect evidence across a network of studies.

Commonly applied methods

- **ITC (Bucher)**
  - IPD not required
  - treatment effects calculated for each trial separately
  - within study randomization preserved
- **Population-adjusted indirect comparison (MAIC)**
  - IPD required for at least 1 trial
  - to match the IPD to the AgD of the other trial
- **Network Meta-Analysis (NMA)**
  - comparing interventions simultaneously in a single analysis by combining both direct and indirect evidence across a network of studies.

Commonly applied methods

- **ITC (Bucher)**
  - IPD not required
  - treatment effects calculated for each trial separately
  - within study randomization preserved
- **Population-adjusted indirect comparison (MAIC)**
  - IPD required for at least 1 trial
  - to match the IPD to the AgD of the other trial
- **Network Meta-Analysis (NMA)**
  - comparing interventions simultaneously in a single analysis by combining both direct and indirect evidence across a network of studies.



Indirect comparisons of competing interventions

AM Glenny,[1*] DG Altman,[2] F Song,[3] C Sakarovitch,[2] JJ Deeks,[2] R D'Amico,[2] M Bradburn[2] and AJ Eastwood[4]

Health Technology Assessment 2005; Vol. 9: No. 26

When conducting systematic reviews to evaluate the effectiveness of interventions, direct evidence from good-quality RCTs should be used wherever possible. If little or no such evidence exists, it may be necessary to look for indirect comparisons from RCTs. The reviewer needs, however, to be aware that the results may be susceptible to bias.

Consistency Assumption

there must be no relevant discrepancy between direct and indirect evidence

Indirect Treatment Comparison (Bucher)

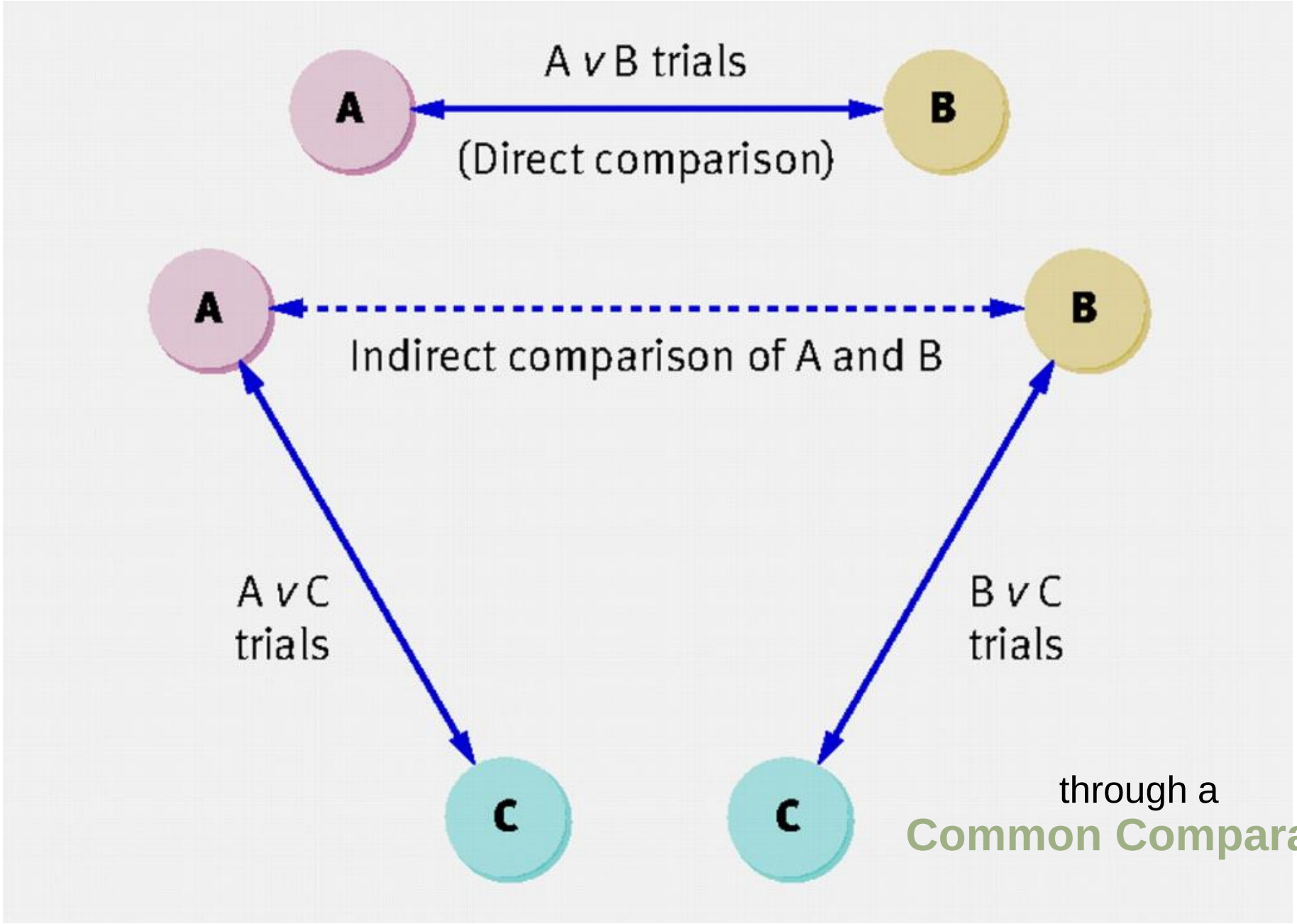
(M. Cinquini)

Introduction – Indirect Treatment Comparisons (ITC)

• **Systematic reviews** of randomized controlled trials (RCTs) are a standard method of analyzing information in the health-care setting.
– **ITCs** are often necessary in order to combine this information and answer many research questions of interest.

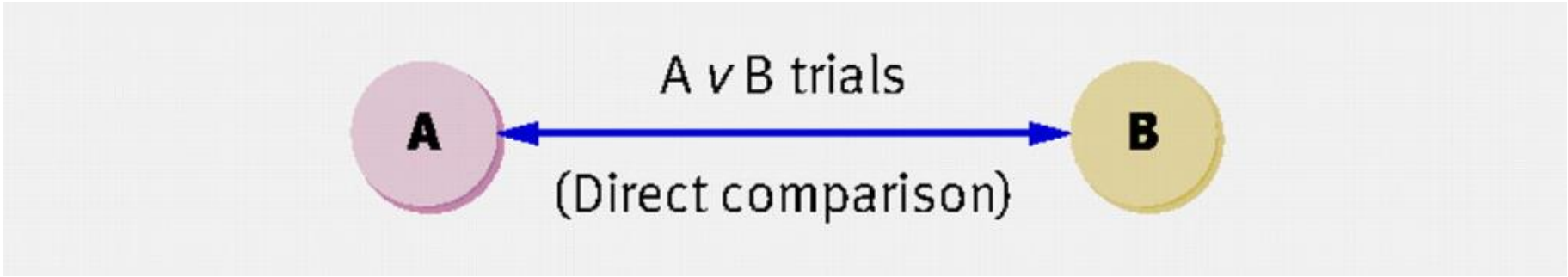
• This is particularly important in the comparative effectiveness landscape where head-to-head comparisons of interest are often unavailable.

• Approach:
– ITCs often use the relative effects of the treatments versus their common comparator (e.g., placebo) in order to assess the head-to-head comparison of interest



through a

Common Comparator



The best?

No head-to-head comparison

Indirect Comparisons

Basic assumptions underlying indirect comparisons include:

- ✓ **similarity** assumption for adjusted indirect comparison,
- ✓ **homogeneity** assumption for standard meta-analysis and
- ✓ **consistency** assumption for the combination of direct and indirect evidence. It is essential to fully understand and appreciate these basic assumptions in order to use adjusted indirect and mixed treatment comparisons appropriately.

What is indirect comparison? Fujian Song BMed MMed PhD Reader in Research Synthesis, Faculty of Health, University of East Anglia
www.whatisseries.co.uk http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_ind_comp.pdf

SIMILARITY (TRANSITIVITY) ASSUMPTION

- For an adjusted indirect comparison (A vs B) to be valid, a similarity assumption is required in terms of moderators of relative treatment effect.
- That is, patients included should be sufficiently similar in the two sets of control arms (C_1 from the trial comparing A vs C_1, and C_2, from the trial comparing B vs C_2).
- This is crucial as only a large theoretical overlap between patients enrolled in C_1 and C_2 enables the relative effect estimated by trials of A versus C_1 to be generalizable to patients in trials of B versus C_1, and the relative effect estimated by trials of B versus C_2 to be generalizable to patients in trials of A versus C_2.

Song, What is …? 2009

WHAT FACTORS DETERMINE SIMILARITY?

- Included trials should be “comparable” in terms of key factors that could affect the outcome of treatment
- If differences in patient or study characteristics would not be expected to influence treatment effect, the assumption of similarity is not violated
- **There are no statistical methods to test for similarity**
- Must use clinical knowledge and best judgement to assess appropriate comparability

IDENTIFY KNOWN TREATMENT EFFECT MODIFIERS

- Effect modifiers include more than patient characteristics
- Trial and outcome properties are also important
- The PICOS framework can be used to identify these variables

Would the treatment be expected to work equally in all patients included into the meta-analysis?

	Description	Sample Variab
P	Patient Population	Demographics, baseline clinical characteristics, disease severity
I	Intervention	Dose, mode of admin, duration
C	Comparator	Active treatment, placebo, concomitant meds
O	Outcomes	Definitions, thresholds, ITT vs. PP, EOS vs. EOT, LOCF vs. NC=F,
S	Setting	Study design, study duration, location/country, method of outcome assessment

IDENTIFY KNOWN TREATMENT EFFECT MODIFIERS

- Effect modifiers include more than patient characteristics
- Trial and outcome properties are also important
- The PICOS framework can be used to identify these variables

Dosing and duration may or may not be important to treatment outcome.

	Description	Sample Variab
P	Patient Population	Demographics, baseline clinical characteristics, disease severity
I	Intervention	Dose, mode of admin, duration
C	Comparator	Active treatment, placebo, concomitant meds
O	Outcomes	Definitions, thresholds, ITT vs. PP, EOS vs. EOT, LOCF vs. NC=F,
S	Setting	Study design, study duration, location/country, method of outcome assessment

IDENTIFY KNOWN TREATMENT EFFECT MODIFIERS

- Effect modifiers include more than patient characteristics
- Trial and outcome properties are also important
- The PICOS framework can be used to identify these variables

	Description	Sample Variables
P	Patient Population	Demographics, baseline clinical characteristics, disease severity
I	Intervention	Dose, mode of admin, duration
C	Comparator	Active treatment, placebo, concomitant meds
O	Outcomes	Definitions, thresholds, ITT vs. PP, EOS vs. EOT, LOCF vs. NC=F,
S	Setting	Study design, study duration, location/country, method of outcome assessment

In pair-wise meta-analyses the comparator must be the same for each trial. In NMA, the comparators need not be equal, but it must fit within the network.

IDENTIFY KNOWN TREATMENT EFFECT MODIFIERS

- Effect modifiers include more than patient characteristics
- Trial and outcome properties are also important
- The PICOS framework can be used to identify these variables

	Description	Sample Variables
P	Patient Population	Demographics, baseline clinical characteristics, disease severity
I	Intervention	Dose, mode of admin, duration
C	Comparator	Active treatment, placebo, concomitant meds
O	Outcomes	Definitions, thresholds, ITT vs. PP, EOS vs. EOT, LOCF vs. NC=F,
S	Setting	Study design, study duration, location/country, method of outcome assessment

How outcomes are calculated can influence observed treatment effect.

IDENTIFY KNOWN TREATMENT EFFECT MODIFIERS

- Effect modifiers include more than patient characteristics
- Trial and outcome properties are also important
- The PICOS framework can be used to identify these variables

	Description	Sample Variables
P	Patient Population	Demographics, baseline clinical characteristics, disease severity
I	Intervention	Dose, mode of admin, duration
C	Comparator	Active treatment, placebo, concomitant meds
O	Outcomes	Definitions, thresholds, ITT vs. PP, EOS vs. EOT, LOCF vs. NC=F,
S	Setting	Study design, study duration, location/country, method of outcome assessment

Some general study characteristics can be important. Eg, timing of assessments, study locations with different standards of care, patient vs. physician-reported outcomes.

Indirect Comparisons

Basic assumptions underlying indirect comparisons include:

- ✓ **similarity** assumption for adjusted indirect comparison,
- ✓ **homogeneity** assumption for standard meta-analysis and
- ✓ **consistency** assumption for the combination of direct and indirect evidence. It is essential to fully understand and appreciate these basic assumptions in order to use adjusted indirect and mixed treatment comparisons appropriately.

What is indirect comparison? Fujian Song BMed MMed PhD Reader in Research Synthesis, Faculty of Health, University of East Anglia
www.whatisseries.co.uk http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_ind_comp.pdf

HOMOGENEITY ASSUMPTION

- When multiple trials are available for a given comparison, the results from multiple trials can be pooled in meta-analyses before an adjusted indirect comparison is conducted.
- For a meta-analysis to be valid, it is commonly established that results from different trials should be sufficiently homogeneous from a clinical and statistical perspective.
- This is usually demonstrated by a 2-tailed p value for homogeneity at Pearson chi-squared test or Cochran Q test > 0.10 and a I^2 (inconsistency) < 50%.
- When homogeneity is unlikely (e.g. I^2>50%) than heterogeneity and inconsistency are likely.

Song, What is …? 2009; Higgins et al, BMJ 2003

ELSEVIER

Critical Reviews in Oncology/Hematology 94 (2015) 213–227

CRITICAL REVIEWS IN
Oncology
Hematology
Incorporating Geriatric Oncology

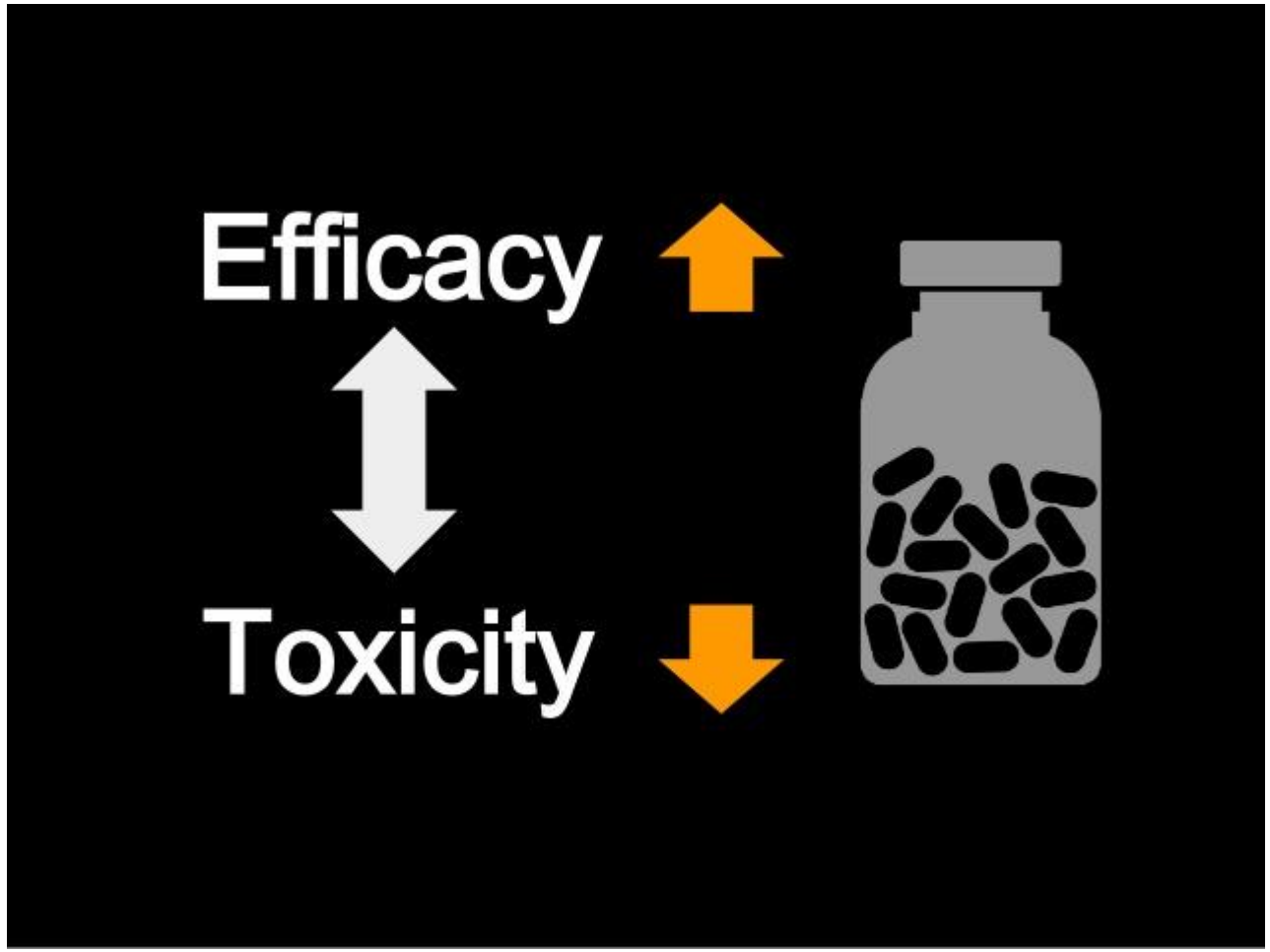
www.elsevier.com/locate/critrevonc

Is there evidence for different effects among EGFR-TKIs? Systematic review and meta-analysis of EGFR tyrosine kinase inhibitors (TKIs) *versus* chemotherapy as first-line treatment for patients harboring EGFR mutations

Eva Regina Haspinger[a], Francesco Agustoni[a], Valter Torri[b], Francesco Gelsomino[a], Marco Platania[a], Nicoletta Zilembo[a], Rosaria Gallucci[a], Marina Chiara Garassino[a,*], Michela Cinquini[b]

[a] *Medical Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
[b] *Fondazione IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy*

Accepted 11 November 2014


Efficacy
Toxicity

Population:

- ✓ previously untreated
- ✓ any age and race
- ✓ histologically proven NSCLC harbouring activating EGFR-mutation

Intervention:

- ✓ EGFR-TKIs (Erlotinib, Gefitinib, Afatinib)

Comparison:

- ✓ Platinum-based chemotherapy

Outcomes:

- ✓ PFS (whenever possible independently reviewed data)
- ✓ PFS in exon 19 deletion
- ✓ PFS in L858R mutation
- ✓ OS
- ✓ ORR (complete and/or partial and/or stable)
- ✓ Treatment related toxic events

Search strategy

PubMed, Cancer-Lit, Embase-databases and Cochrane-Library were searched for RCTs up to June 2014 with no language or publication status restrictions. Search terms were “TKI” [Substance Name] and “Carcinoma, NSCLC”[Substance Name]. The proceedings of the 2008–2014 conferences of the American Society of Clinical Oncology(ASCO), European Society of Medical Oncology (ESMO)and International Association for the Study of Lung Cancer (IASLC), World Conference of Lung Cancer were also searched for relevant abstracts. Any unpublished RCTs were considered for inclusion.

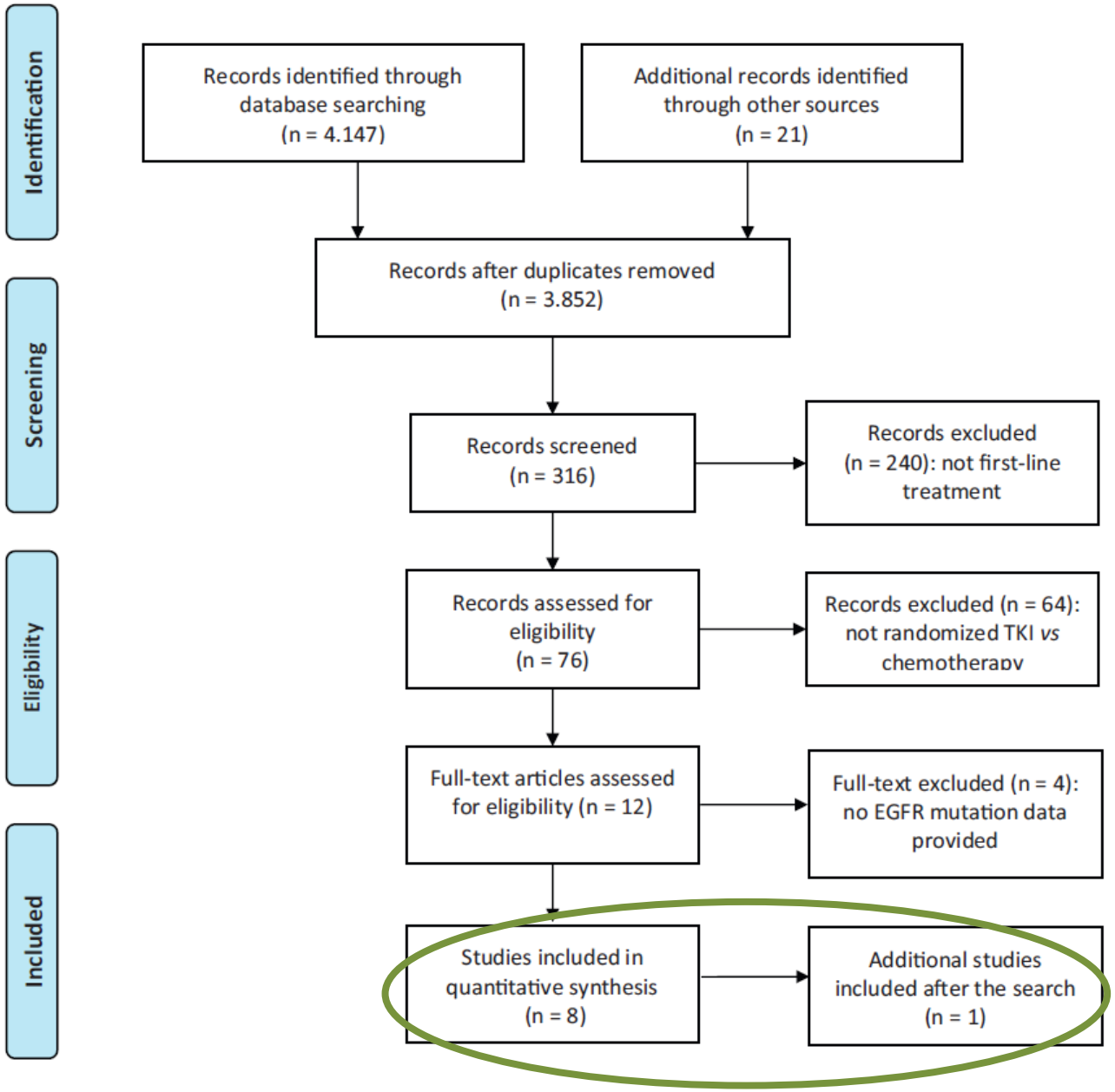


Fig. 1. Flow diagram for the selection of studies included in this meta-analysis.
From: Moher D, Liberati A, Telzlaff J, Altman DG, The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the Prisma statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097.

Indirect Comparisons

Basic assumptions underlying indirect comparisons include:

- ✓ **similarity** assumption for adjusted indirect comparison,
- ✓ **homogeneity** assumption for standard meta-analysis and
- ✓ **consistency** assumption for the combination of direct and indirect evidence. It is essential to fully understand and appreciate these basic assumptions in order to use adjusted indirect and mixed treatment comparisons appropriately.

What is indirect comparison? Fujian Song BMed MMed PhD Reader in Research Synthesis, Faculty of Health, University of East Anglia
www.whatisseries.co.uk http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_ind_comp.pdf

Table 1
Characteristics of the 9 clinical trials included in the meta-analysis.

Trial	Primary end-point	TKI	Chemotherapy	Patients (TKI/CT)	EGFR + patients (%)	Asiatic patients (%)	Crossover (%)[a]
IPASS Mok, 2009	Progression-free survival	Gefitinib	Carboplatin + paclitaxel	1.217 (609/608)	21.4	99.8	39.5
WJTOG3405 Mitsudomi, 2010	Progression-free survival	Gefitinib	Cisplatin + paclitaxel	177 (88/89)	100	100	59.3
NEJ002 Maemondo, 2010	Progression-free survival	Gefitinib	Carboplatin + paclitaxel	228 (114/114)	100	100	94.6
First-SIGNAL Han, 2012	Overall survival	Gefitinib	Cisplatin + gemcitabine	309 (159/150)	13.6	100	75.0
TORCH Gridelli, 2012	Overall survival	Erlotinib	Cisplatin + gemcitabine	760 (380/380)	5.1	0	60.9
OPTIMAL Zhou, 2011	Progression-free survival	Erlotinib	Carboplatin + gemcitabine	154 (82/72)	100	100	NA
EURTAC Rosell, 2011	Progression-free survival	Erlotinib	Cisplatin/carboplatin + docetaxel/gemcitabine	173 (86/87)	100	0	76.0
LUX-Lung 3 Sequist, 2012	Progression-free survival	Afatinib	Cisplatin + pemetrexed	345 (230/115)	100	100	75.0
LUX-Lung 6 Wu, 2013	Progression-free survival	Afatinib	Cisplatin + gemcitabine	364 (242/122)	100	100	56.0

[a] Patients who have been treated with crossover from chemotherapy to TKI in second-line.

Study		
FIRST-SIGNAL	Cisplatin 75 mg/m2 Gemcitabine 1	ery 3 weeks cles
IPASS	Carboplati mg/mi Pac	weeks up to 6 weeks
NEJG002		
WJTOG3405		to 6 weeks
EURTAC		
OPTI		.v. 4 cycles
TORC	g/m2	i.v. every 3 weeks up to 6 weeks
LUX-LUNG I	m2 Pemetrexed	i.v. 6 cycles
LUX-LUNG VI	75 mg/m2 Gemcitabine mg/m2 day 1&8	i.v. Up to 6 cycles

STANDARD CHEMOTHERAPY

Indirect Comparisons

Basic assumptions underlying indirect comparisons include:

- ✓ **similarity** assumption for adjusted indirect comparison,
- ✓ **homogeneity** assumption for standard meta-analysis and
- ✓ **consistency** assumption for the combination of direct and indirect evidence. It is essential to fully understand and appreciate these basic assumptions in order to use adjusted indirect and mixed treatment comparisons appropriately.

What is indirect comparison? Fujian Song BMed MMed PhD Reader in Research Synthesis, Faculty of Health, University of East Anglia
www.whatisseries.co.uk http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_ind_comp.pdf

Data synthesis:

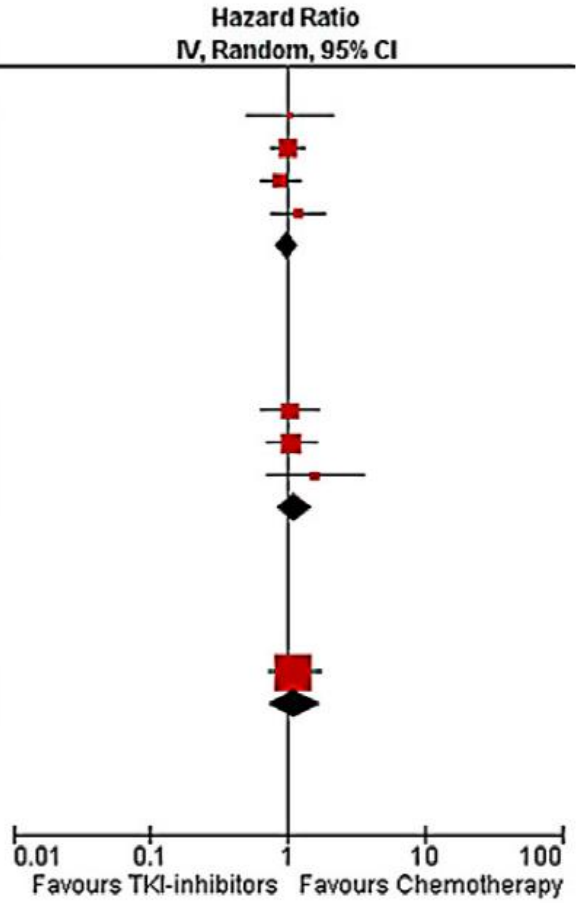
- ✓ HR for OS and PFS
- ✓ RR for the Others

OS

Panel B

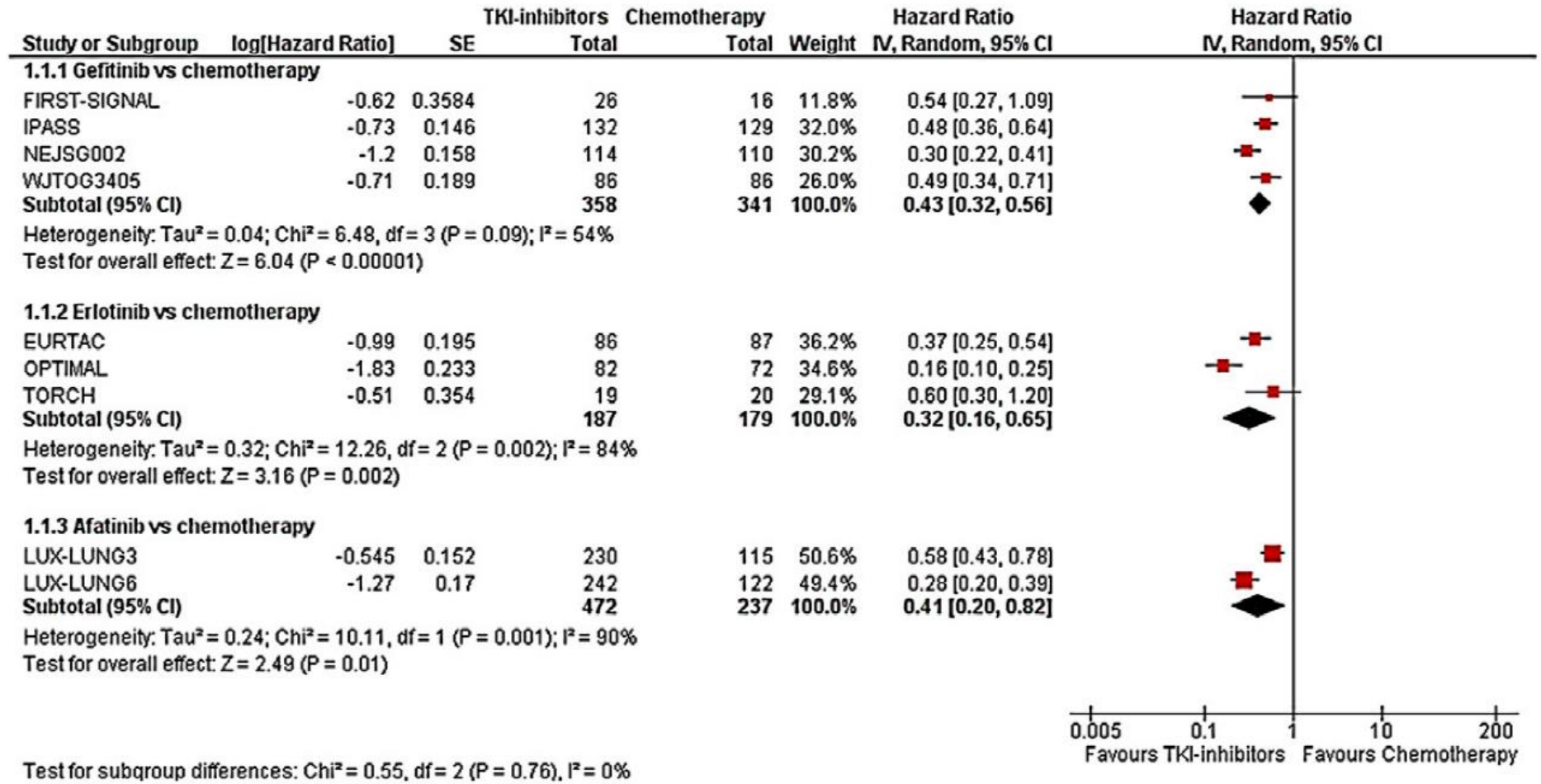
Study or Subgroup	log[Hazard Ratio]	SE	TKI-inhibitors Total	Chemotherapy Total	Weight	Hazard Ratio IV, Random, 95% CI
1.2.1 Gefitinib vs chemotherapy						
FIRST-SIGNAL	0.0392	0.3755	26	16	6.4%	1.04 [0.50, 2.17]
IPASS	0	0.143	132	129	44.3%	1.00 [0.76, 1.32]
NEJSG002	-0.12	0.171	114	110	31.0%	0.89 [0.63, 1.24]
WJTOG3405	0.17	0.223	86	86	18.2%	1.19 [0.77, 1.84]
Subtotal (95% CI)			**358**	**341**	**100.0%**	**1.00 [0.83, 1.20]**
Heterogeneity: Tau² = 0.00; Chi² = 1.08, df = 3 (P = 0.78); I² = 0%						
Test for overall effect: Z = 0.04 (P = 0.97)						
1.2.2 Erlotinib vs chemotherapy						
EURTAC	0.039	0.24	86	87	39.5%	1.04 [0.65, 1.66]
OPTIMAL	0.0677	0.219	82	72	47.4%	1.07 [0.70, 1.64]
TORCH	0.457	0.416	19	20	13.1%	1.58 [0.70, 3.57]
Subtotal (95% CI)			**187**	**179**	**100.0%**	**1.11 [0.83, 1.50]**
Heterogeneity: Tau² = 0.00; Chi² = 0.82, df = 2 (P = 0.66); I² = 0%						
Test for overall effect: Z = 0.71 (P = 0.48)						
1.2.3 Afatinib						
LUX-LUNG3	0.11	0.22	230	115	100.0%	1.12 [0.73, 1.72]
Subtotal (95% CI)			**230**	**115**	**100.0%**	**1.12 [0.73, 1.72]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.50 (P = 0.62)						

Test for subgroup differences: Chi² = 0.51, df = 2 (P = 0.77), I² = 0%

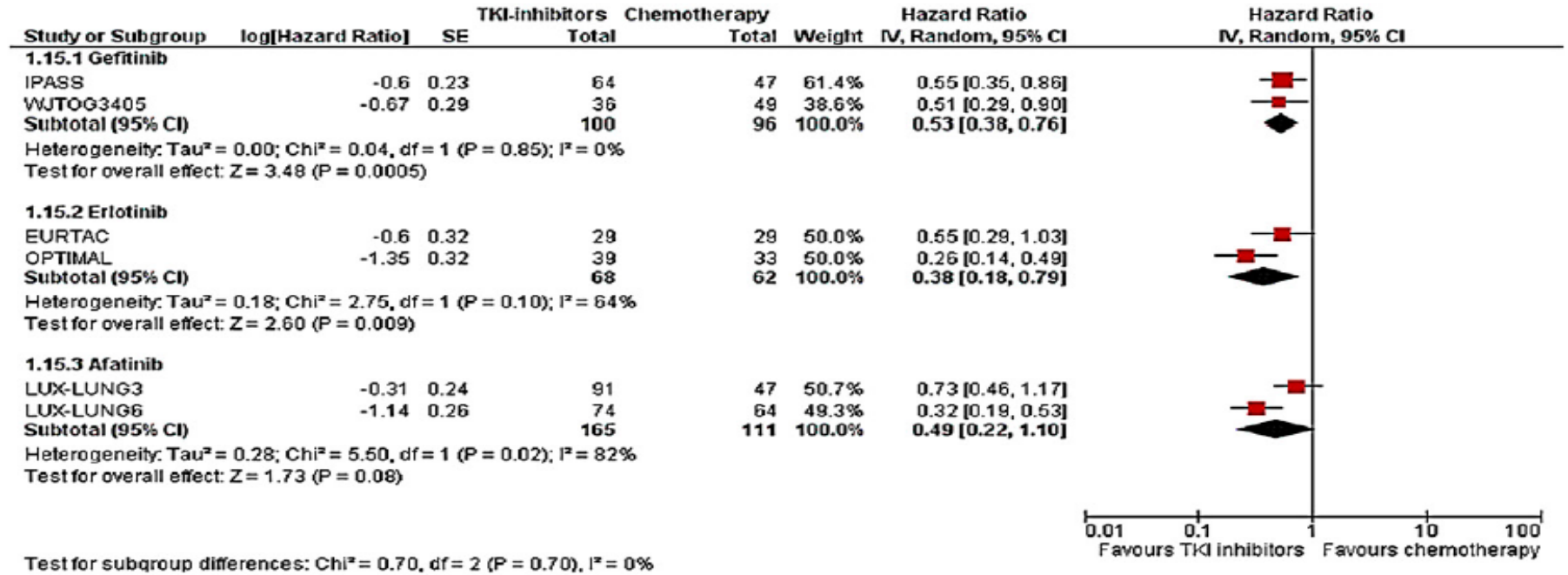


PFS

Panel A

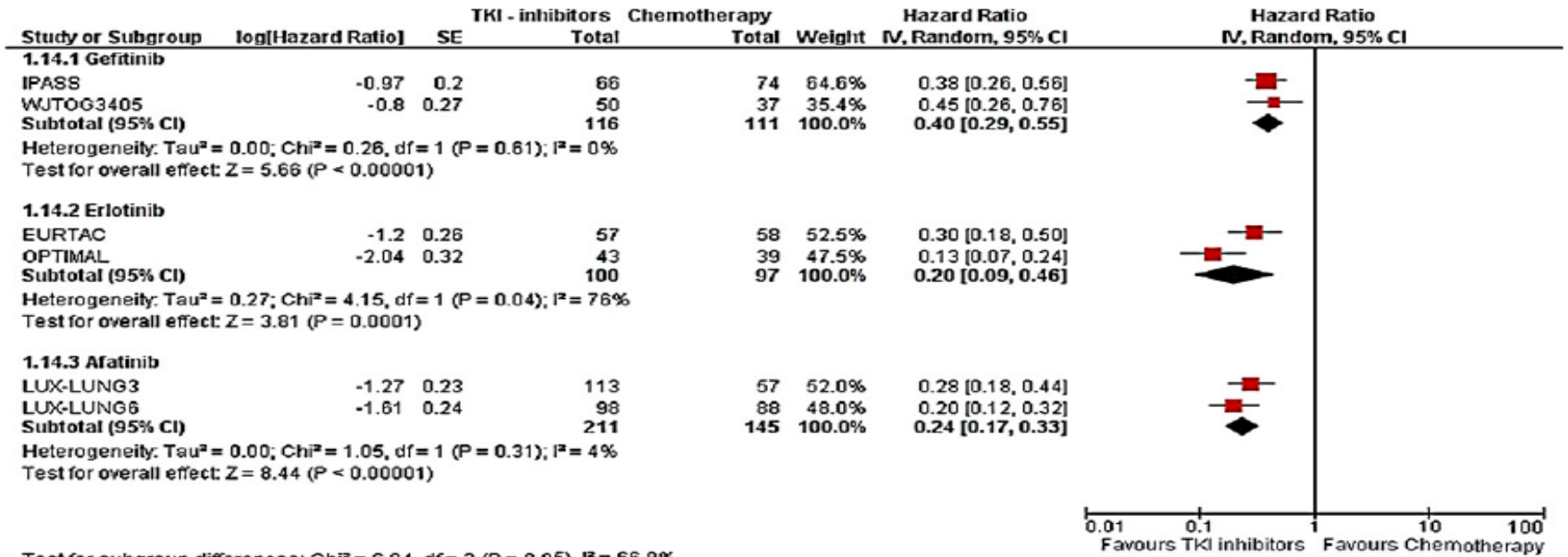

Study or Subgroup log[Hazard Ratio] SE TKI-inhibitors Total Chemotherapy Total Weight Hazard Ratio IV, Random, 95% CI
1.1.1 Gefitinib vs chemotherapy
FIRST-SIGNAL -0.62 0.3584 26 16 11.8% 0.54 [0.27, 1.09]
IPASS -0.73 0.146 132 129 32.0% 0.48 [0.36, 0.64]
NEJSG002 -1.2 0.158 114 110 30.2% 0.30 [0.22, 0.41]
WJTOG3405 -0.71 0.189 86 86 26.0% 0.49 [0.34, 0.71]
Subtotal (95% CI) 358 341 100.0% 0.43 [0.32, 0.56]
Heterogeneity: Tau² = 0.04; Chi² = 6.48, df = 3 (P = 0.09); I² = 54%
Test for overall effect: Z = 6.04 (P < 0.00001)
1.1.2 Erlotinib vs chemotherapy
EURTAC -0.99 0.195 86 87 36.2% 0.37 [0.25, 0.54]
OPTIMAL -1.83 0.233 82 72 34.6% 0.16 [0.10, 0.25]
TORCH -0.51 0.354 19 20 29.1% 0.60 [0.30, 1.20]
Subtotal (95% CI) 187 179 100.0% 0.32 [0.16, 0.65]
Heterogeneity: Tau² = 0.32; Chi² = 12.26, df = 2 (P = 0.002); I² = 84%
Test for overall effect: Z = 3.16 (P = 0.002)
1.1.3 Afatinib vs chemotherapy
LUX-LUNG3 -0.545 0.152 230 115 50.6% 0.58 [0.43, 0.78]
LUX-LUNG6 -1.27 0.17 242 122 49.4% 0.28 [0.20, 0.39]
Subtotal (95% CI) 472 237 100.0% 0.41 [0.20, 0.82]
Heterogeneity: Tau² = 0.24; Chi² = 10.11, df = 1 (P = 0.001); I² = 90%
Test for overall effect: Z = 2.49 (P = 0.01)
Test for subgroup differences: Chi² = 0.55, df = 2 (P = 0.76), I² = 0%
Hazard Ratio IV, Random, 95% CI
0.005
0.1
1
10
200
Favours TKI-inhibitors
Favours Chemotherapy

Exon 21

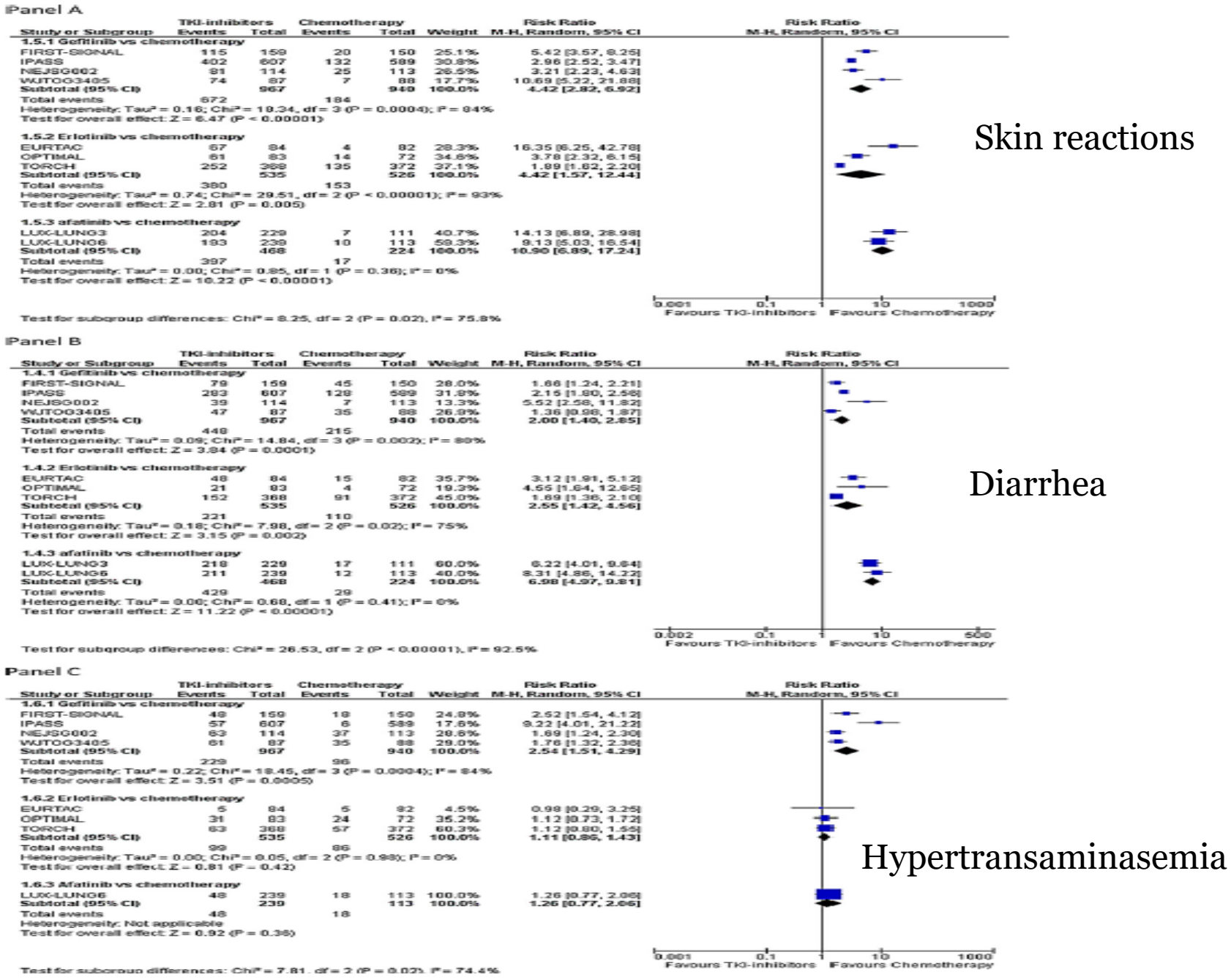


Test for subgroup differences: Chi² = 0.70, df = 2 (P = 0.70), I² = 0%

Exon 19



Test for subgroup differences: Chi² = 6.04, df = 2 (P = 0.05), I² = 66.9%


Panel A
TKI-inhibitors
Chemotherapy
Risk Ratio
Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Random, 95% CI
M-H, Random, 95% CI
1.5.1 Gefitinib vs chemotherapy
FIRST-SIGNAL 115 159 20 150 25.1% 5.42 [3.57, 8.25]
IPASS 402 607 132 589 30.8% 2.96 [2.52, 3.47]
NEJSG002 81 114 25 113 26.5% 3.21 [2.23, 4.63]
WJTOG3405 74 87 7 88 17.7% 10.69 [5.22, 21.88]
Subtotal (95% CI) 967 940 100.0% 4.42 [2.82, 6.92]
Total events 672 184
Heterogeneity: Tau² = 0.16; Chi² = 19.24, df = 3 (P = 0.0004); I² = 84%
Test for overall effect: Z = 6.47 (P < 0.00001)
1.5.2 Erlotinib vs chemotherapy
EURTAC 67 84 4 82 28.3% 16.35 [6.25, 42.78]
OPTIMAL 61 83 14 72 34.6% 3.78 [2.32, 6.15]
TORCH 252 368 135 372 37.1% 1.89 [1.62, 2.20]
Subtotal (95% CI) 535 526 100.0% 4.42 [1.57, 12.44]
Total events 380 153
Heterogeneity: Tau² = 0.74; Chi² = 29.51, df = 2 (P < 0.00001); I² = 93%
Test for overall effect: Z = 2.81 (P = 0.005)
1.5.3 afatinib vs chemotherapy
LUX-LUNG3 204 229 7 111 40.7% 14.13 [6.89, 28.98]
LUX-LUNG6 193 239 10 113 59.3% 9.13 [5.03, 16.54]
Subtotal (95% CI) 468 224 100.0% 10.90 [6.89, 17.24]
Total events 397 17
Heterogeneity: Tau² = 0.00; Chi² = 0.85, df = 1 (P = 0.36); I² = 0%
Test for overall effect: Z = 10.22 (P < 0.00001)
Test for subgroup differences: Chi² = 8.25, df = 2 (P = 0.02), I² = 75.8%
0.001 0.1 1 10 1000
Favours TKI-inhibitors Favours Chemotherapy
Skin reactions
Panel B
TKI-inhibitors
Chemotherapy
Risk Ratio
Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Random, 95% CI
M-H, Random, 95% CI
1.4.1 Gefitinib vs chemotherapy
FIRST-SIGNAL 79 159 45 150 28.0% 1.66 [1.24, 2.21]
IPASS 283 607 128 589 31.8% 2.15 [1.80, 2.56]
NEJSG002 39 114 7 113 13.3% 5.52 [2.58, 11.82]
WJTOG3405 47 87 35 88 26.9% 1.36 [0.98, 1.87]
Subtotal (95% CI) 967 940 100.0% 2.00 [1.40, 2.85]
Total events 448 215
Heterogeneity: Tau² = 0.09; Chi² = 14.84, df = 3 (P = 0.002); I² = 80%
Test for overall effect: Z = 3.94 (P = 0.0001)
1.4.2 Erlotinib vs chemotherapy
EURTAC 48 84 15 82 35.7% 3.12 [1.91, 5.12]
OPTIMAL 21 83 4 72 19.3% 4.55 [1.64, 12.65]
TORCH 152 368 91 372 45.0% 1.69 [1.36, 2.10]
Subtotal (95% CI) 535 526 100.0% 2.55 [1.42, 4.56]
Total events 221 110
Heterogeneity: Tau² = 0.18; Chi² = 7.98, df = 2 (P = 0.02); I² = 75%
Test for overall effect: Z = 3.15 (P = 0.002)
1.4.3 afatinib vs chemotherapy
LUX-LUNG3 218 229 17 111 60.0% 6.22 [4.01, 9.64]
LUX-LUNG6 211 239 12 113 40.0% 8.31 [4.86, 14.22]
Subtotal (95% CI) 468 224 100.0% 6.98 [4.97, 9.81]
Total events 429 29
Heterogeneity: Tau² = 0.00; Chi² = 0.68, df = 1 (P = 0.41); I² = 0%
Test for overall effect: Z = 11.22 (P < 0.00001)
Test for subgroup differences: Chi² = 26.53, df = 2 (P < 0.00001), I² = 92.5%
0.002 0.1 1 10 500
Favours TKI-inhibitors Favours Chemotherapy
Diarrhea
Panel C
TKI-inhibitors
Chemotherapy
Risk Ratio
Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Random, 95% CI
M-H, Random, 95% CI
1.6.1 Gefitinib vs chemotherapy
FIRST-SIGNAL 48 159 19 150 24.8% 2.52 [1.54, 4.12]
IPASS 57 607 6 589 17.6% 9.22 [4.01, 21.22]
NEJSG002 63 114 37 113 28.6% 1.69 [1.24, 2.30]
WJTOG3405 61 87 35 88 29.0% 1.76 [1.32, 2.36]
Subtotal (95% CI) 967 940 100.0% 2.54 [1.51, 4.29]
Total events 229 96
Heterogeneity: Tau² = 0.22; Chi² = 18.45, df = 3 (P = 0.0004); I² = 84%
Test for overall effect: Z = 3.51 (P = 0.0005)
1.6.2 Erlotinib vs chemotherapy
EURTAC 5 84 5 82 4.5% 0.98 [0.29, 3.25]
OPTIMAL 31 83 24 72 35.2% 1.12 [0.73, 1.72]
TORCH 63 368 57 372 60.3% 1.12 [0.80, 1.56]
Subtotal (95% CI) 535 526 100.0% 1.11 [0.86, 1.43]
Total events 99 86
Heterogeneity: Tau² = 0.00; Chi² = 0.05, df = 2 (P = 0.98); I² = 0%
Test for overall effect: Z = 0.81 (P = 0.42)
1.6.3 Afatinib vs chemotherapy
LUX-LUNG6 48 239 18 113 100.0% 1.26 [0.77, 2.06]
Subtotal (95% CI) 239 113 100.0% 1.26 [0.77, 2.06]
Total events 48 18
Heterogeneity: Not applicable
Test for overall effect: Z = 0.92 (P = 0.36)
Test for subgroup differences: Chi² = 7.81, df = 2 (P = 0.02), I² = 74.4%
0.001 0.1 1 10 1000
Favours TKI-inhibitors Favours Chemotherapy
Hypertransaminasemia

So, who's the best?

COMPUTATIONS
based on Bucher et al. method

- The log relative risk of the adjusted indirect comparison of A and B ($\ln RR_{A \text{ vs } B}$) can be estimated by:

$$\ln RR_{A \text{ vs } B} = \ln RR_{A \text{ vs } C1} - \ln RR_{B \text{ vs } C2}$$

- and its standard error is:

$$SE(\ln RR_{A \text{ vs } B}) = \sqrt{[SE(\ln RR_{A \text{ vs } C1})^2 + SE(\ln RR_{B \text{ vs } C2})^2]}$$

- Similar computations can be envisioned for odds ratio, absolute risk reductions, weighted mean differences, and standardized mean differences.

Higgins et al, BMJ 2003; Song, What is …? 2009;
http://www.metcardio.org/macros/IMT.xls

Panel A

Study or Subgroup	log[Hazard Ratio\Risk ratio]	SE	Weight	Hazard Ratio\Risk ratio IV, Random, 95% CI
Progression-free survival	0.295	0.385		1.34 [0.63, 2.86]
PFS-exon 19	0.693	0.447		2.00 [0.83, 4.80]
PFS-L858R	0.332	0.417		1.39 [0.62, 3.16]
Overall survival	-0.104	0.177		0.90 [0.64, 1.27]
Objective response rate	-0.036	0.168		0.96 [0.69, 1.34]
Diarrhea	-0.223	0.121		0.80 [0.63, 1.01]
Rash	0	0.101		1.00 [0.82, 1.22]
Hypertransaminasemia	0.83	0.175		2.29 [1.63, 3.23]
Treatment discontinuation	-0.019	0.384		0.98 [0.46, 2.08]
Treatment-related death	1.05	1.295		2.86 [0.23, 36.17]

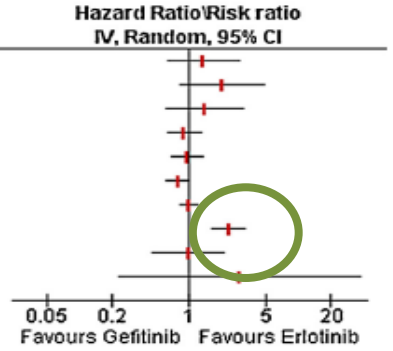
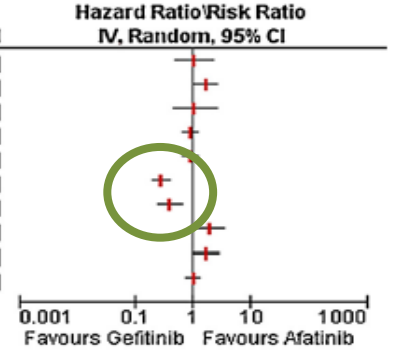


Image of Fig. 5

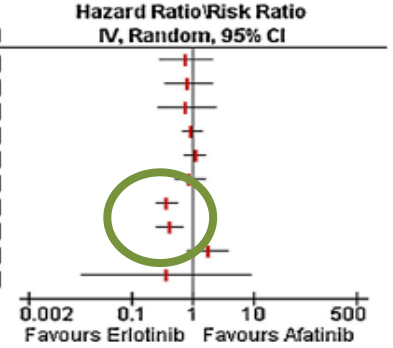
Panel B

Study or Subgroup	log[Hazard Ratio\Risk Ratio]	SE	Weight	Hazard Ratio\Risk Ratio IV, Random, 95% CI
Progression-free survival	0.048	0.387		1.05 [0.49, 2.24]
PFS-exon 19	0.511	0.235		1.67 [1.05, 2.64]
PFS-L858R	0.078	0.447		1.08 [0.45, 2.60]
Overall survival	-0.099	0.167		0.91 [0.65, 1.26]
Objective response rate	-0.097	0.157		0.91 [0.67, 1.23]
Diarrhea	-1.25	0.187		0.29 [0.20, 0.41]
Rash	-0.903	0.244		0.41 [0.25, 0.65]
Hypertransaminasemia	0.701	0.276		2.02 [1.17, 3.46]
Treatment discontinuation	0.531	0.273		1.70 [1.00, 2.90]
Treatment-related death	0.022	0.136		1.02 [0.78, 1.33]



Panel C

Study or Subgroup	log[Hazard Ratio\Risk Ratio]	SE	Weight	Hazard Ratio\Risk Ratio IV, Random, 95% CI
Progression-free survival	-0.248	0.507		0.78 [0.29, 2.11]
PFS-exon 19	-0.182	0.449		0.83 [0.35, 2.01]
PFS-L858R	-0.254	0.558		0.78 [0.26, 2.32]
Objective response rate	-0.061	0.186		0.94 [0.65, 1.35]
Overall survival	0.094	0.204		1.10 [0.74, 1.64]
Hypertransaminasemia	-0.127	0.285		0.88 [0.50, 1.54]
Diarrhea	-1.01	0.2		0.36 [0.25, 0.54]
Rash	-0.903	0.245		0.41 [0.25, 0.66]
Treatment discontinuation	0.55	0.395		1.73 [0.80, 3.76]
Treatment-related death	-1.03	1.637		0.36 [0.01, 8.83]



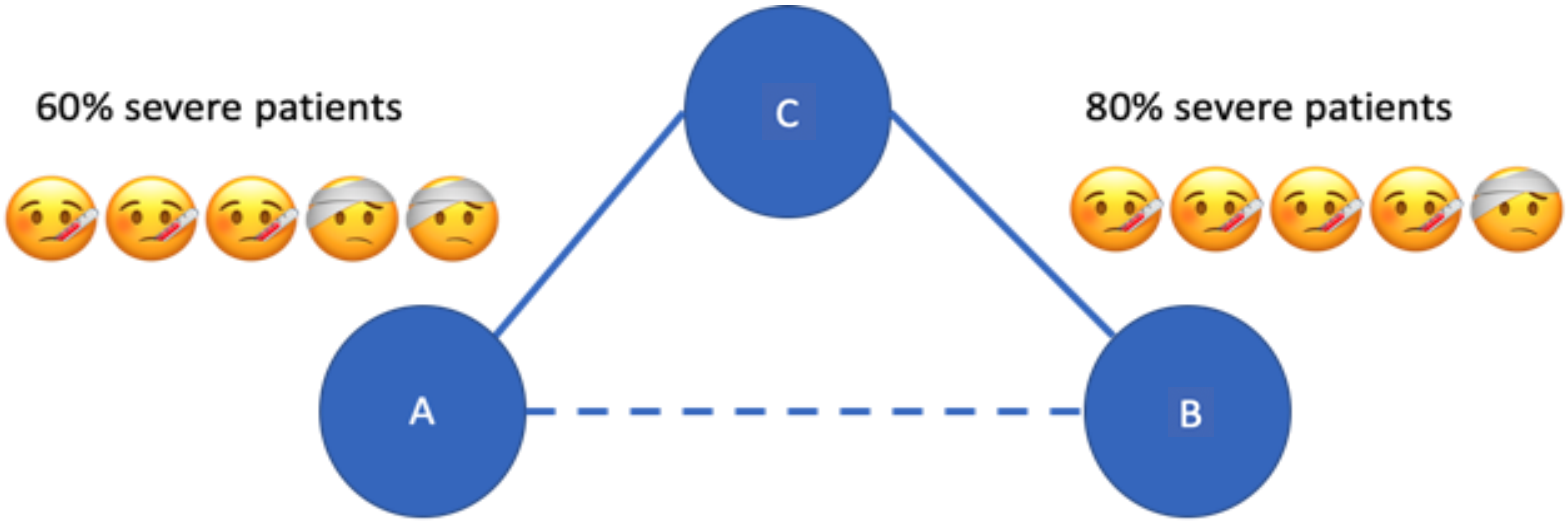
TAKE HOME MESSAGES

- Adjusted indirect comparison meta-analysis represents a simple yet robust tool to make statistical and clinical inference despite the lack of conclusive evidence from head-to-head randomized clinical trials.

- Despite being not at the uppermost level of the hierarchy of evidence based medicine, it can often provide results equivalent to those of subsequent direct comparisons.

Population-Adjusted Indirect Comparison

(G.L. Pappagallo)



WHAT FACTORS DETERMINE SIMILARITY?

- Included trials should be “comparable” in terms of key factors that could affect the outcome of treatment
- If differences in patient or study characteristics would not be expected to influence treatment effect, the assumption of similarity is not violated
- **There are no statistical methods to test for similarity**
- Must use clinical knowledge and best judgement to assess appropriate comparability

Population-adjusted Indirect Comparisons

- Population-adjusted Indirect Comparisons use patient-level data from a trial of a given treatment (referred to as the index trial) to derive a comparison of outcomes with competing treatments, based on published information from similarly designed studies, after adjusting for differences in the characteristics of the populations.
- *in other words*: individual patient data (IPD) in one or more trials are used to adjust for between-trial differences in the distribution of variables that influence outcome
- "anchored" indirect comparison (common comparator arm in each trial) Vs "unanchored" indirect comparison (disconnected treatment network or single-arm studies)
  - *an unanchored comparison assumes that all effect modifiers and prognostic factors are accounted for*
- unanchored methods for population adjustment are problematic and should not be used when anchored methods can be applied

https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

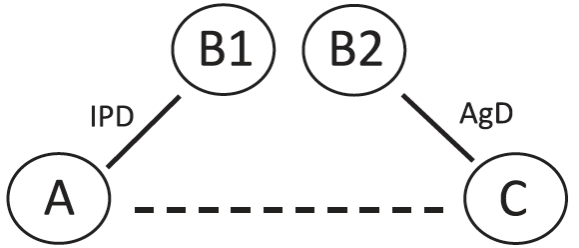
Population-adjusted Indirect Comparisons

- Population-adjusted Indirect Comparisons use patient-level data from a trial of a given treatment (referred to as the index trial) to derive a comparison of outcomes with competing treatments, based on published information from similarly designed studies, after adjusting for differences in the characteristics of the populations.
- *in other words*: individual patient data (IPD) in one or more trials are used to adjust for between-trial differences in the distribution of variables that influence outcome
- "anchored" indirect comparison (common comparator arm in each trial) Vs "unanchored" indirect comparison (disconnected treatment network or single-arm studies)
  - *an unanchored comparison assumes that all effect modifiers and prognostic factors are accounted for*
- unanchored methods for population adjustment are problematic and should not be used when anchored methods can be applied

https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

Matching-adjusted indirect comparison (MAIC)

- needs IPD for at least 1 trial, because the aim is to match the IPD to the AgD of the other trial



- the matching procedure selects a weight for each patient to reach similarity in the summary measures of the baseline characteristics of the IPD and AgD trial and follows the idea of *propensity score matching*
- the odds between being a patient in trial AB *Vs* trial CB provides the weights for balancing the populations

Comparative Effectiveness Research in Oncology Methodology: Observational Data

Dawn L. Hershman and Jason D. Wright

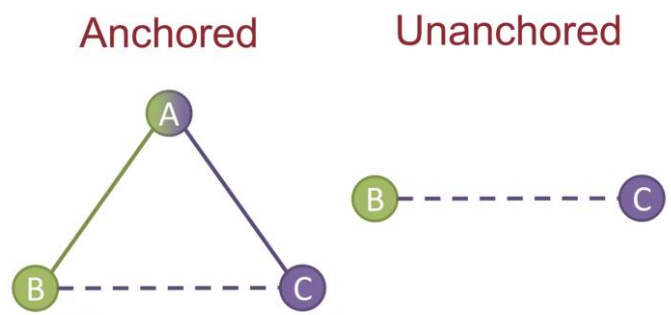
J Clin Oncol 30:4215-4222. © 2012 by American Society of Clinical Oncology

Propensity Score Analysis

Propensity score analyses attempt to balance covariates between experimental groups. Using multivariable modeling, the characteristics of a cohort are used to calculate the probability of receiving the intervention. This probability is the propensity score.

Population-adjusted Indirect Comparisons

- Population-adjusted Indirect Comparisons use patient-level data from a trial of a given treatment (referred to as the index trial) to derive a comparison of outcomes with competing treatments, based on published information from similarly designed studies, after adjusting for differences in the characteristics of the populations.
- *in other words*: individual patient data (IPD) in one or more trials are used to adjust for between-trial differences in the distribution of variables that influence outcome
- "anchored" indirect comparison (common comparator arm in each trial) Vs "unanchored" indirect comparison (disconnected treatment network or single-arm studies)
  - *an unanchored comparison assumes that all effect modifiers and prognostic factors are accounted for*
- unanchored methods for population adjustment are problematic and should not be used when anchored methods can be applied

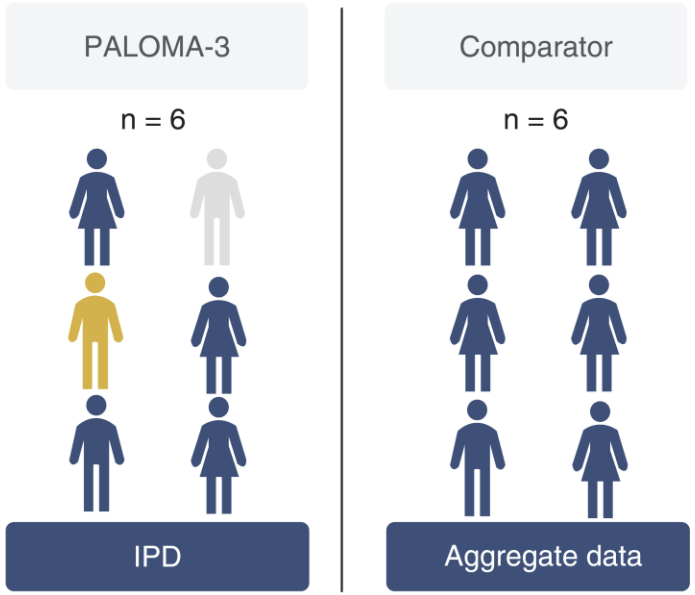


https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

Matching-adjusted indirect comparison of palbociclib versus ribociclib and abemaciclib in hormone receptor-positive/HER2-negative advanced breast cancer

Hope S Rugo*,[1], Anja Haltner[2], Lin Zhan[3], Anh Tran[2], Eustratios Bananis[3], Becky Hooper[2], Debanjali Mitra[3] & Chris Cameron[2]

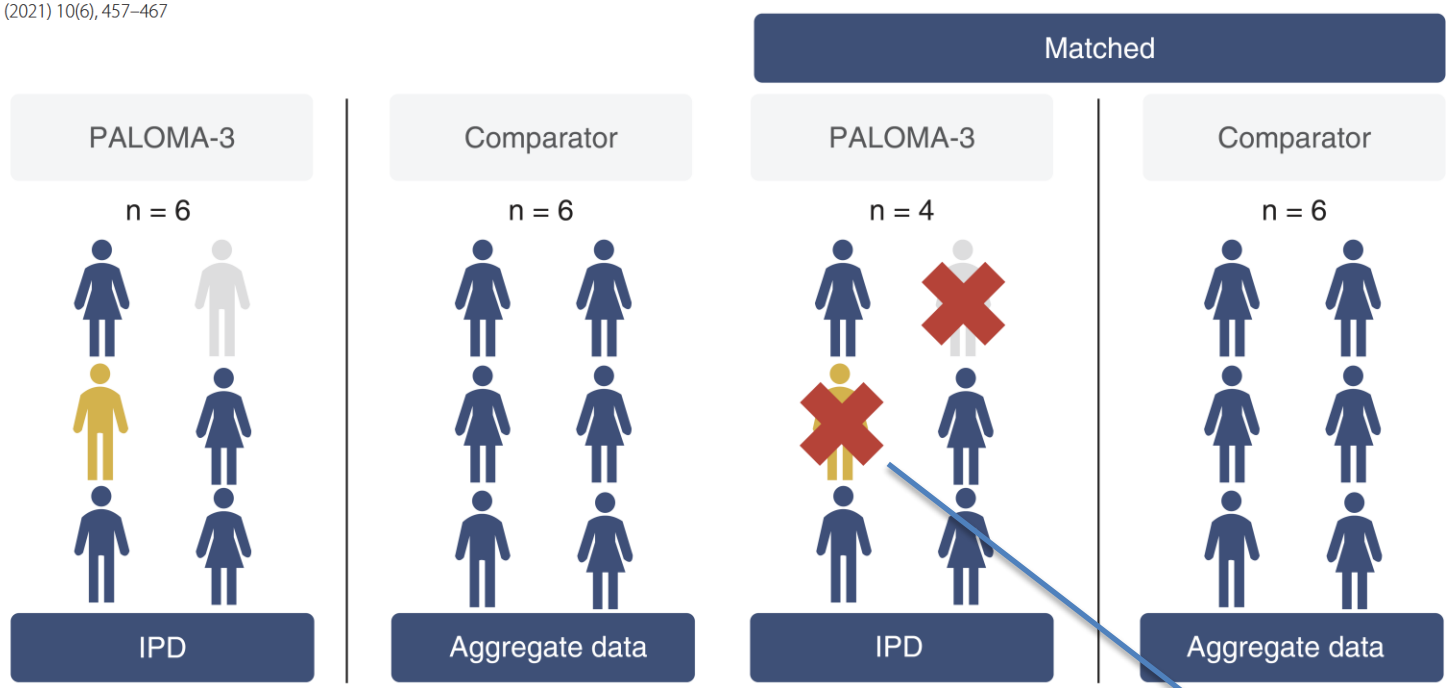
J. Comp. Eff. Res. (2021) 10(6), 457–467



Matching-adjusted indirect comparison of palbociclib versus ribociclib and abemaciclib in hormone receptor-positive/HER2-negative advanced breast cancer

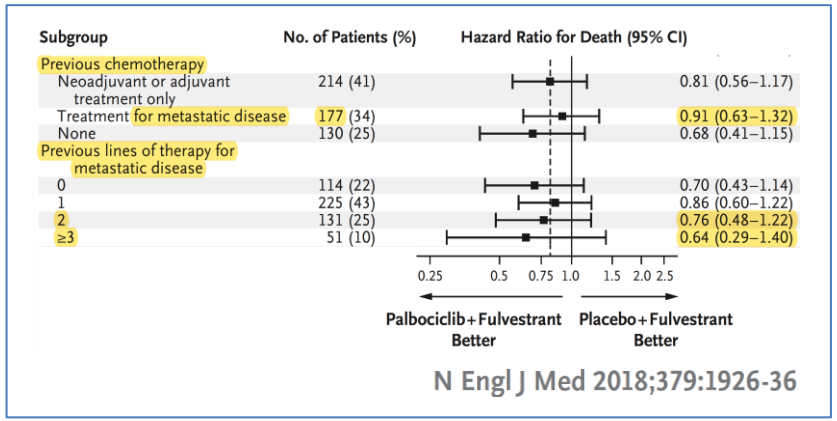
Hope S Rugo*,[1], Anja Haltner[2], Lin Zhan[3], Anh Tran[2], Eustratios Bananis[3], Becky Hooper[2], Debanjali Mitra[3] & Chris Cameron[2]

J. Comp. Eff. Res. (2021) 10(6), 457–467



Published trials differ on eligibility criteria and patient characteristics

Patients who would not have been eligible for enrolment in the comparator trial are excluded (e.g., patients with prior chemotherapy for MBC, or had ≥2 prior lines of ET for MBC)

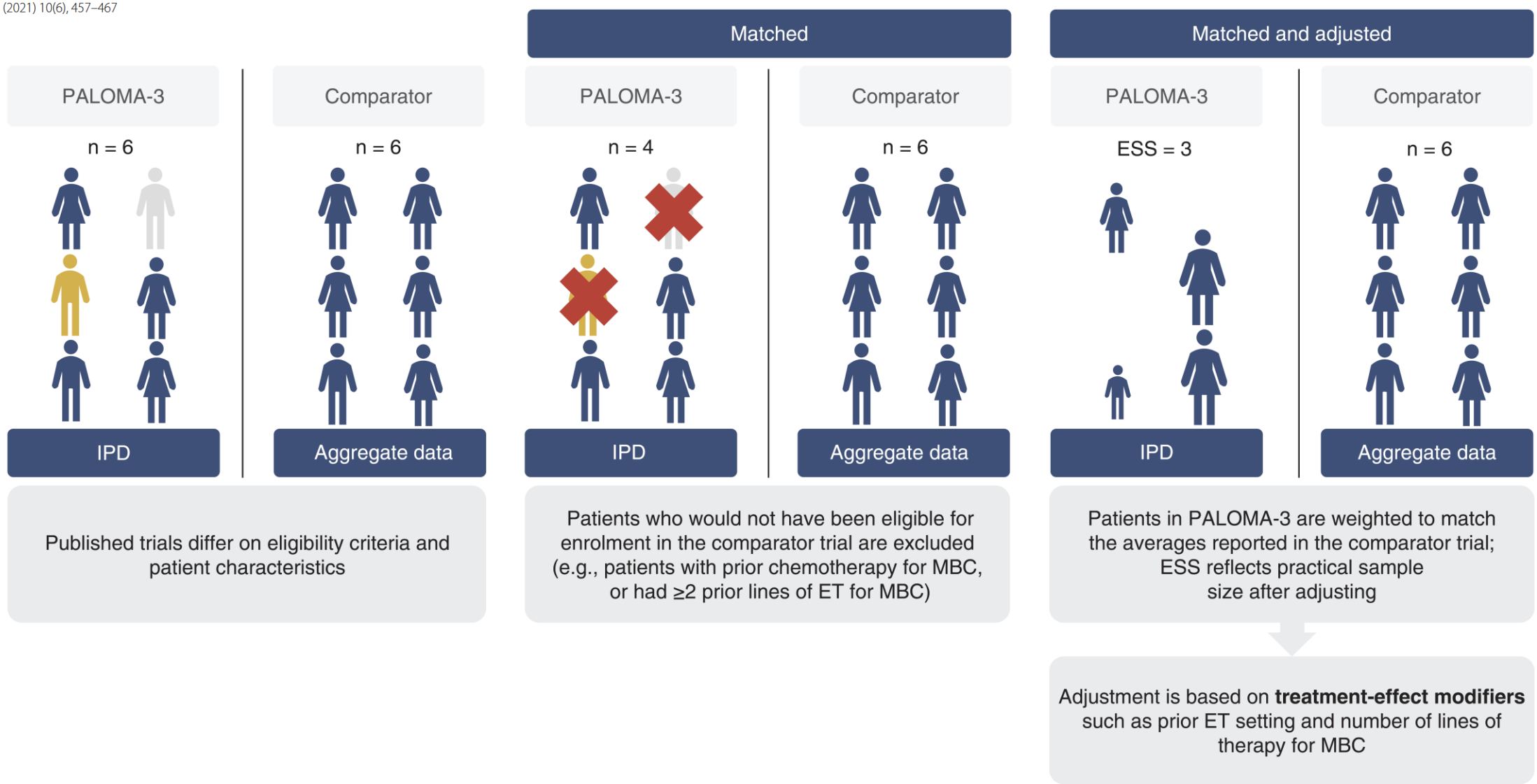


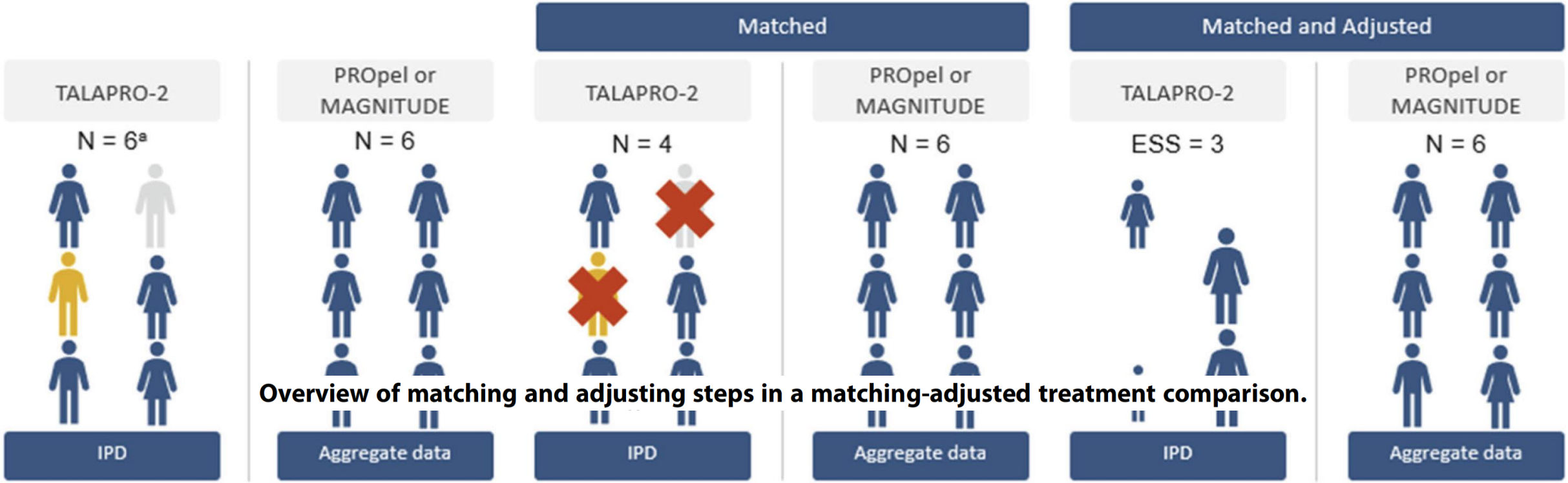
N Engl J Med 2018;379:1926-36

Matching-adjusted indirect comparison of palbociclib versus ribociclib and abemaciclib in hormone receptor-positive/HER2-negative advanced breast cancer

Hope S Rugo*,[1], Anja Haltner[2], Lin Zhan[3], Anh Tran[2], Eustratios Bananis[3], Becky Hooper[2], Debanjali Mitra[3] & Chris Cameron[2]

J. Comp. Eff. Res. (2021) 10(6), 457–467



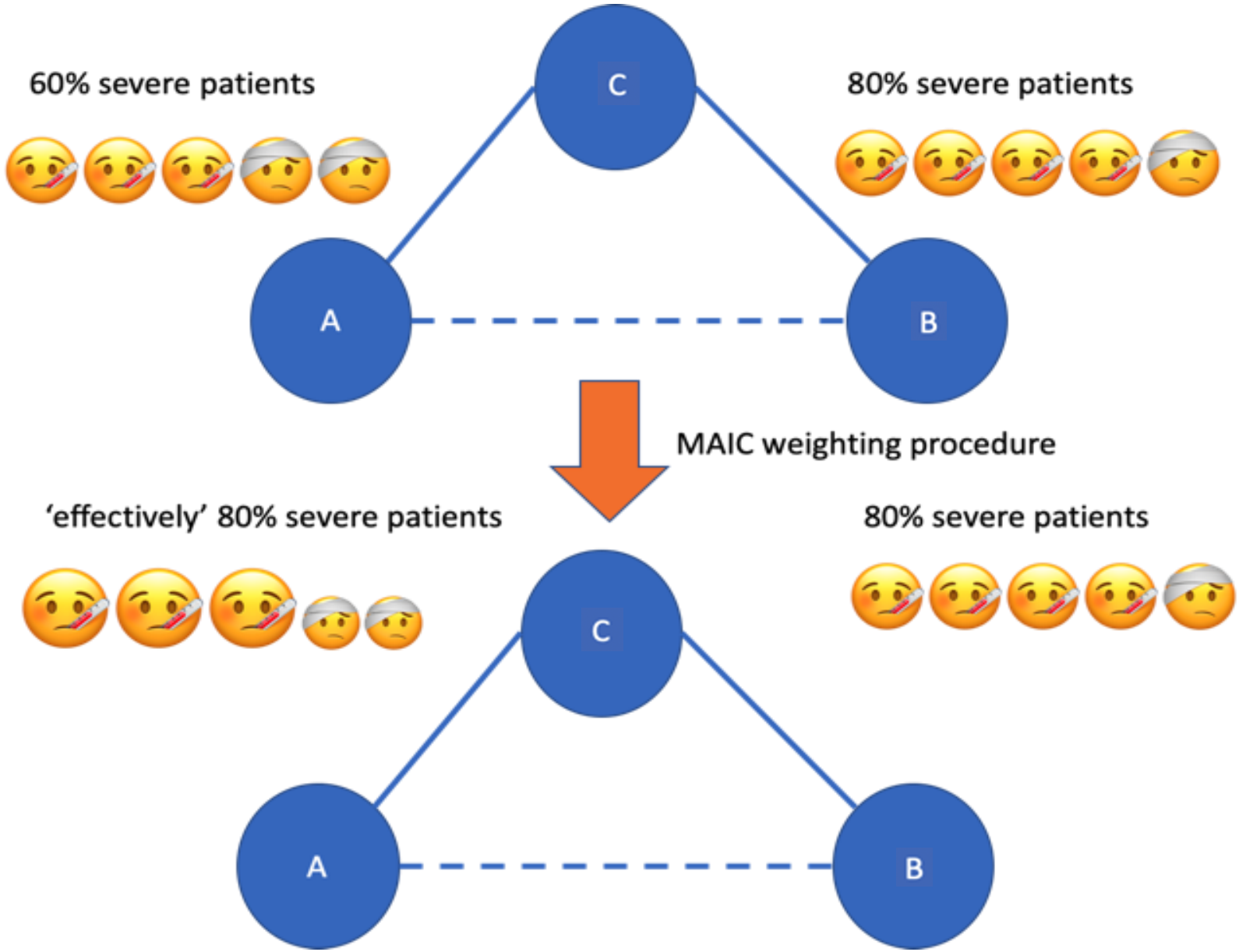


Overview of matching and adjusting steps in a matching-adjusted treatment comparison.

Talazoparib plus enzalutamide versus olaparib plus abiraterone acetate and niraparib plus abiraterone acetate for metastatic castration-resistant prostate cancer: a matching-adjusted indirect comparison

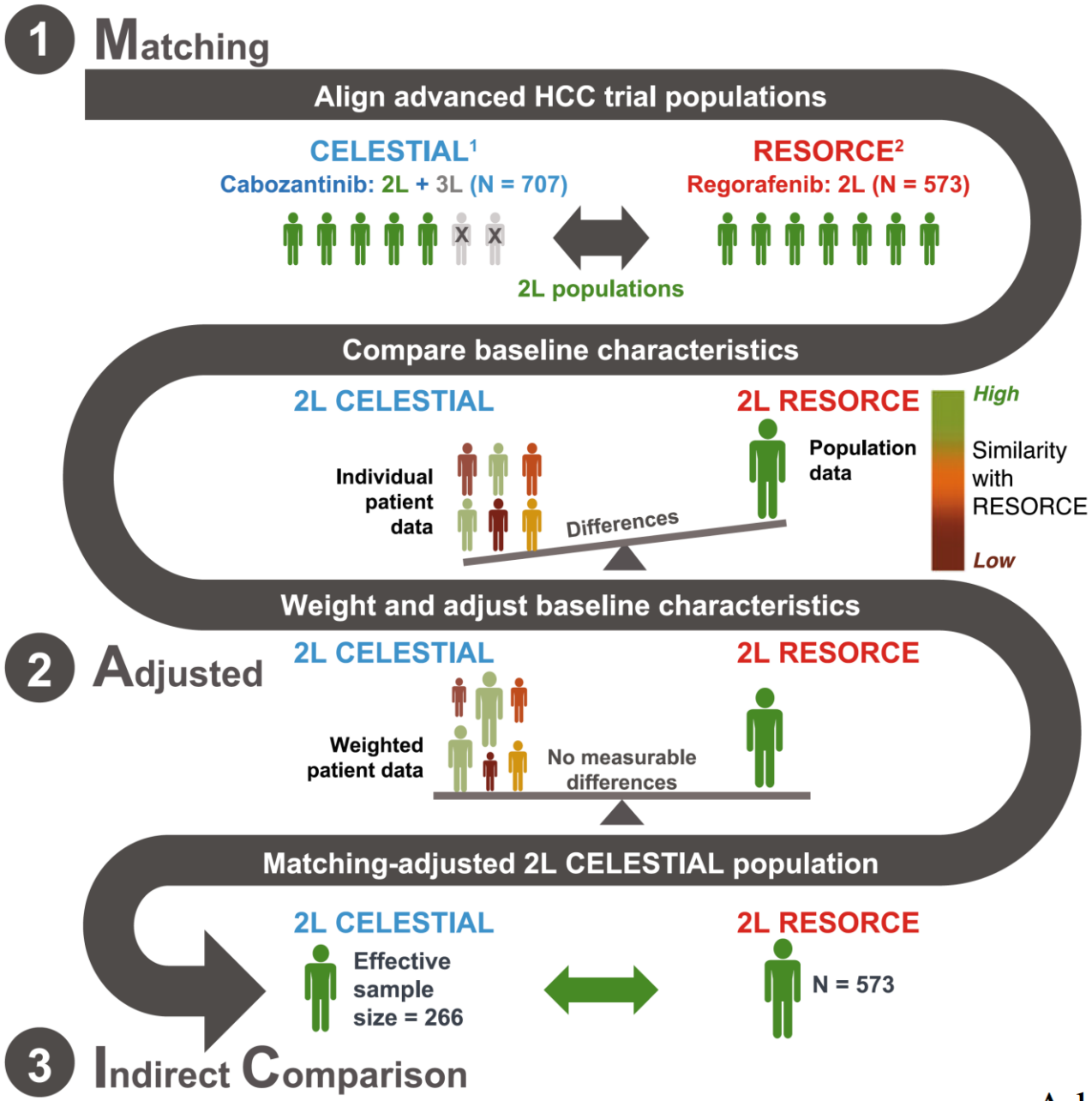
Elena Castro[1✉], Di Wang[2,5], Sarah Walsh[2,5], Samantha Craigie[2,5], Anja Haltner[3,5], Jonathan Nazari[4], Alexander Niyazov[4] and Imtiaz A. Samjoo[2,5✉]

Prostate Cancer and Prostatic Diseases; https://doi.org/10.1038/s41391-024-00924-x


60% severe patients
C
80% severe patients
A
B
MAIC weighting procedure
'effectively' 80% severe patients
80% severe patients
C
A
B

The main limitation relates to the inherent challenge of MAIC in that it is only possible to adjust baseline variables that are mutually reported between trials, and therefore it cannot address the potential unmeasurable differences between the trials.

LEUKEMIA & LYMPHOMA
https://doi.org/10.1080/10428194.2024.2313628



Adv Ther (2020) 37:2678–2695

Matching cannot account for all differences between trial populations, and it is possible that the results of this MAIC are affected by some residual between-trial differences, as evidenced by **the difference in survival outcomes for the placebo arms despite matching and adjustment**.

Table 3 Median survival estimates for the matching-adjusted second-line CELESTIAL population and the RESORCE population: weighted Kaplan-Meier estimates

		KM-derived estimate, months (median [95% CI])	*p* value
Overall survival			
Active treatment	Cabozantinib (ESS = 187)	11.4 (8.9–17.0)	0.3474[a]
	Regorafenib (*n* = 379)	10.6 (9.1–12.1)	
Placebo	CELESTIAL (ESS = 81)	7.2 (6.1–10.8)	NE
	RESORCE (*n* = 194)	7.8 (6.3–8.8)	
Progression-free survival			
Active treatment	Cabozantinib (ESS = 187)	5.6 (4.9–7.3)	0.0005[a]
	Regorafenib (*n* = 379)	3.1 (2.8–4.2)	
Placebo	CELESTIAL (ESS = 81)	1.9 (1.9–2.1)	NE
	RESORCE (*n* = 194)	1.5 (1.4–1.6)	

CI confidence interval, *ESS* effective sample size, *KM* Kaplan-Meier, *NE* not evaluated

[a] Log-rank test

Safety outcomes of darolutamide vs. apalutamide and enzalutamide in non-metastatic castration-resistant prostate cancer (nmCRPC): Matching-adjusted indirect comparisons

Shan Jiang, PhD; Emi Terasawa, PhD; Viviana Garcia-Horton, PhD; Rajeev Ayyagari, PhD; Reg Waldeck, PhD; Susan Halabi, PhD; Neal Shore, MD

ASCO20 Virtual

Figure 1. Differences in risk differences for DARO vs APA[a] safety outcomes

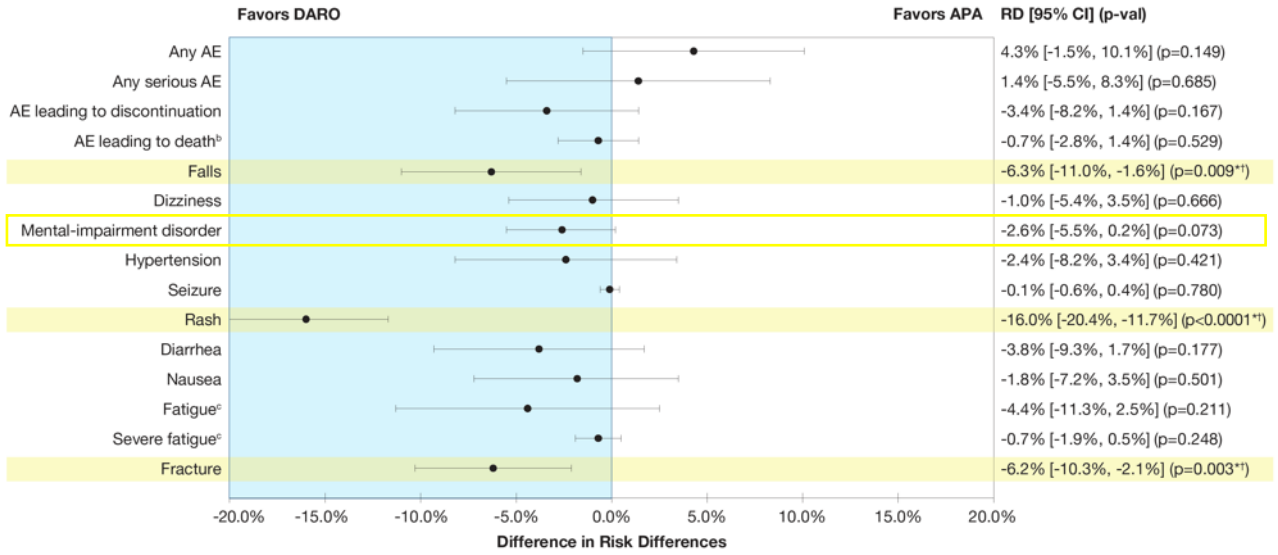
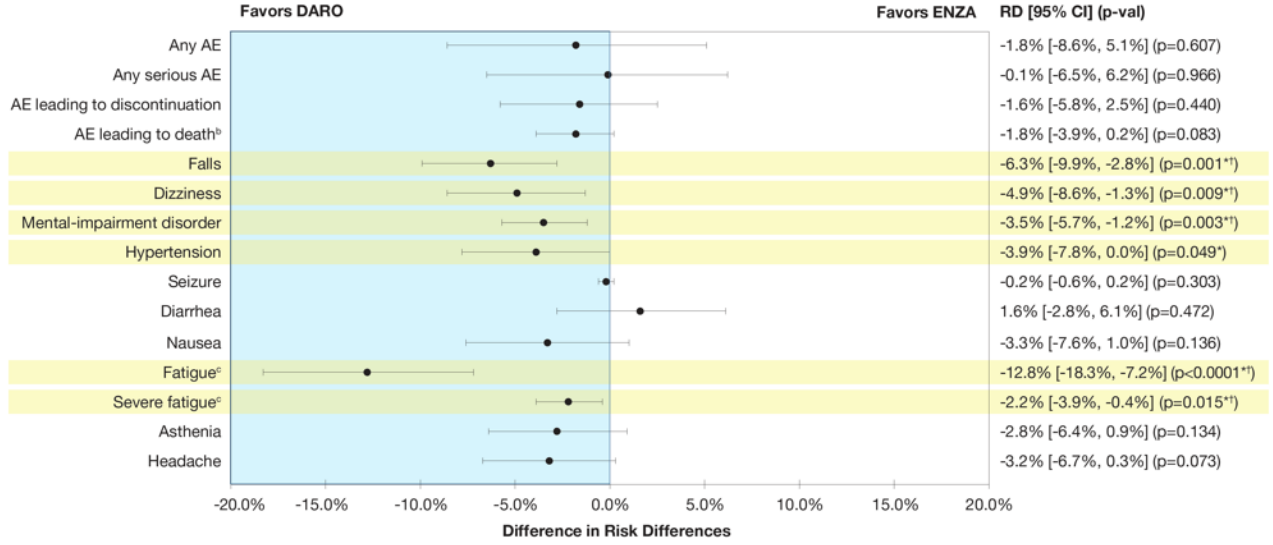


Figure 3. Differences in risk differences for DARO vs ENZA[a] safety outcomes



LIMITATIONS

- Only known baseline factors that were consistently reported across trials were included among the matching covariates in the MAICs.
- As with any comparison of non-randomized treatment groups, such comparisons are subject to potential bias due to unobserved or unmeasurable confounding factors.

… together with information bias due to different monitoring schedules and duration of follow-up.

Overall survival and adverse events after treatment with darolutamide vs. apalutamide vs. enzalutamide for high-risk non-metastatic castration-resistant prostate cancer: a systematic review and network meta-analysis

Mike Wenzel[1,2] · Luigi Nocera[2,3] · Claudia Collà Ruvolo[2,4] · Christoph Würnschimmel[2,5] · Zhe Tian[2] · Shahrokh F. Shariat[6,7,8,9,10,11] · Fred Saad[2] · Derya Tilki[5,12] · Markus Graefen[5] · Luis A. Kluth[1] · Alberto Briganti[3] · Philipp Mandel[1] · Francesco Montorsi[3] · Felix K. H. Chun[1] · Pierre I. Karakiewicz[2]

Conclusion. The current network meta-analysis suggests the highest OS efficacy and lowest grade 3+ toxicity for darolutamide. It is noteworthy that study design, study population, and follow-up duration represent some of the potentially critical differences that distinguish between the three studies and remained statistically unaccounted for using the network meta-analysis methodology. Those differences should be strongly considered in the interpretation of the current and any network meta-analyses.

Prostate Cancer and Prostatic Diseases
https://doi.org/10.1038/s41391-021-00395-4

Indirect Comparison of Darolutamide versus Apalutamide and Enzalutamide for Nonmetastatic Castration-Resistant Prostate Cancer

Susan Halabi,* Shan Jiang,† Emi Terasawa,‡ Viviana Garcia-Horton,‡ Rajeev Ayyagari,‡ A. Reginald Waldeck† and Neal Shore||,§

THE JOURNAL OF UROLOGY® Vol. 206, 298-307, August 2021

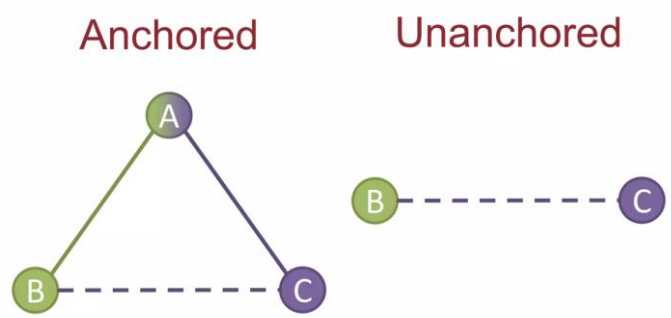
ARAMIS and SPARTAN differed in trial duration (eg median OS followup: 17.9 vs 20.3 months) and AE assessment schedules (every 4 months vs every month). To assess the potential impact of these differences relative to SPARTAN, additional adjustments of the ARAMIS safety outcomes were performed (see supplementary Methods and supplementary figure, https://www.jurology.com).

Results from the primary analysis of darolutamide vs apalutamide persisted even after adjustments for followup time differences: darolutamide had lower risk of fractures, falls and rash than apalutamide as measured via RDs (supplementary table 4, https://www.jurology.com). The lower odds of fractures for darolutamide vs apalutamide as measured via OR, however, did not persist (supplementary table 5, https://www.jurology.com).

All outcome comparisons between darolutamide vs apalutamide were evaluated under two scenarios, using different sets of matching covariates (primary and sensitivity analyses). Importantly, the most pronounced effects were conserved across both analyses, lending further credence to reliability of the results.

Population-adjusted Indirect Comparisons

- Population-adjusted Indirect Comparisons use patient-level data from a trial of a given treatment (referred to as the index trial) to derive a comparison of outcomes with competing treatments, based on published information from similarly designed studies, after adjusting for differences in the characteristics of the populations.
- *in other words*: individual patient data (IPD) in one or more trials are used to adjust for between-trial differences in the distribution of variables that influence outcome
- "anchored" indirect comparison (common comparator arm in each trial) Vs "unanchored" indirect comparison (disconnected treatment network or single-arm studies)
  - *an unanchored comparison assumes that all effect modifiers and prognostic factors are accounted for*
- unanchored methods for population adjustment are problematic and should not be used when anchored methods can be applied



https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

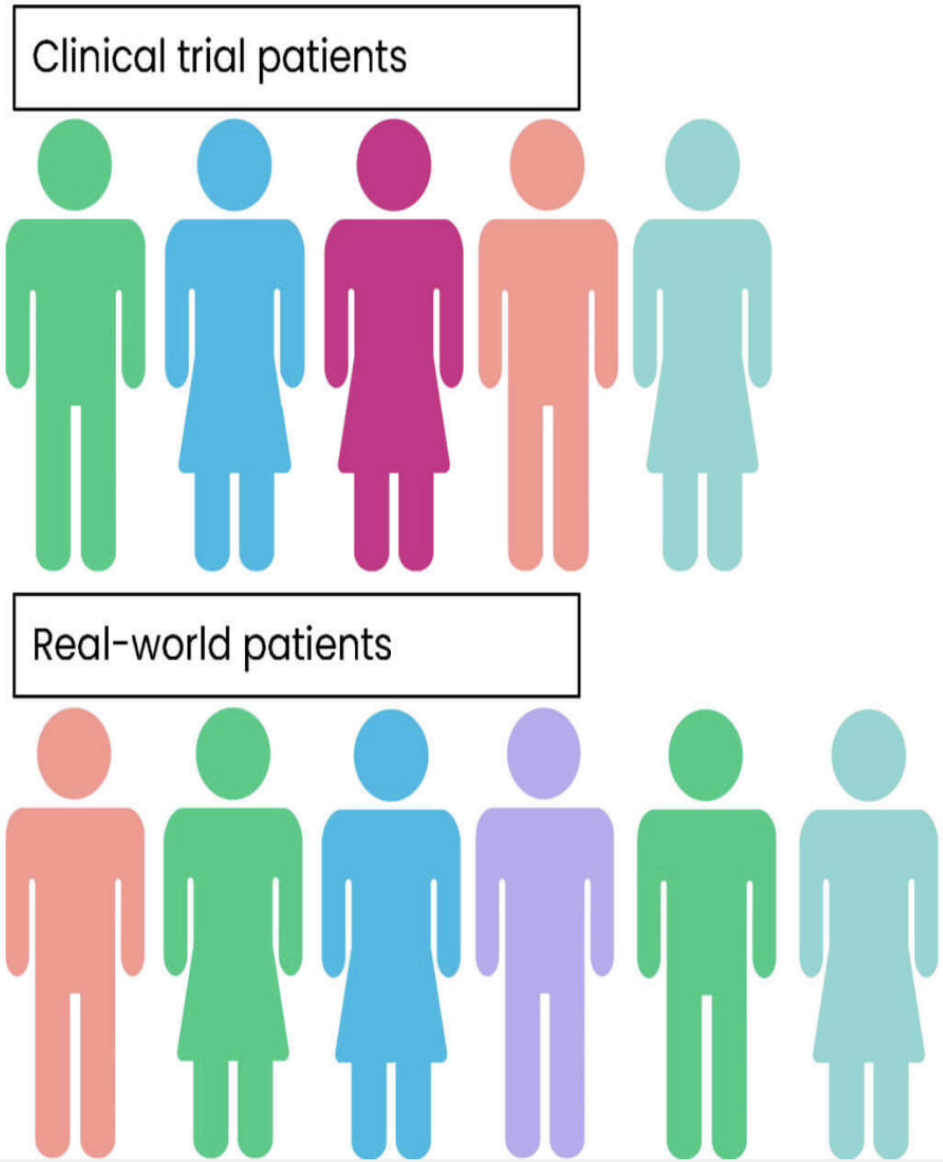
Elranatamab efficacy in MagnetisMM-3 compared with real-world control arms in triple-class refractory multiple myeloma

Luciano J Costa[1], Thomas W LeBlanc[2], Hans Tesch[3], Pieter Sonneveld[4], Ryan P Kyle[5], Liliya Sinyavskaya[5], Patrick Hlavacek[6], Aster Meche[6], Jinma Ren[7], Alex Schepart[6], Didem Aydin[6], Guido Nador[8] & Marco daCosta DiBonaventura*[,6]

Future Oncol. 2024 Feb 28. doi: 10.2217/fon-2023-0995. Epub ahead of print. PMID: 38415370

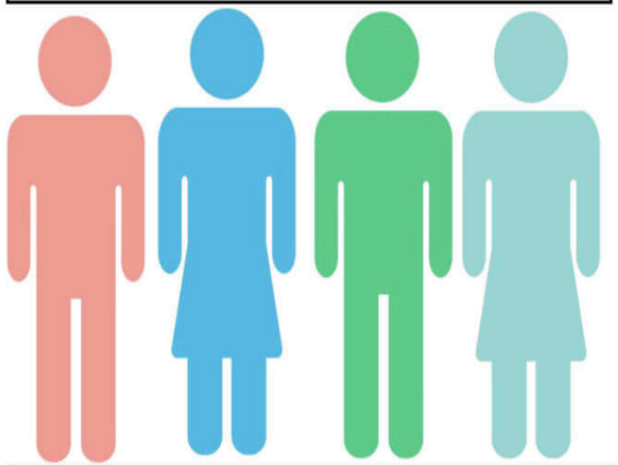
Elranatamab efficacy in the single-arm, registrational MagnetisMM-3 trial (NCT04649359) was compared with that of physician's choice of treatment (PCT) for triple-class refractory multiple myeloma. MagnestisMM-3 eligibility criteria were applied to two USA-based oncology electronic health record databases, COTA and Flatiron Health (FH), to identify cohorts for this study (NCT05932290). Applied statistical techniques accounted for cohort imbalances. MagnetisMM-3 (BCMA-naive; n = 123) outcomes were compared with those from COTA (n = 239) and FH (n = 152).

Original, unadjusted population

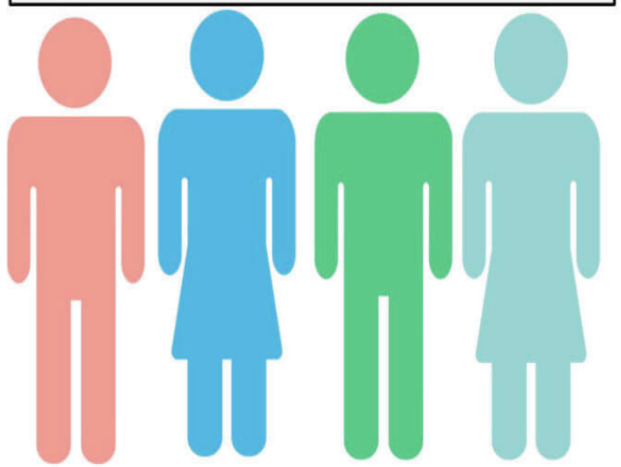


Adjusted population after PS matching

Retained clinical trial patients

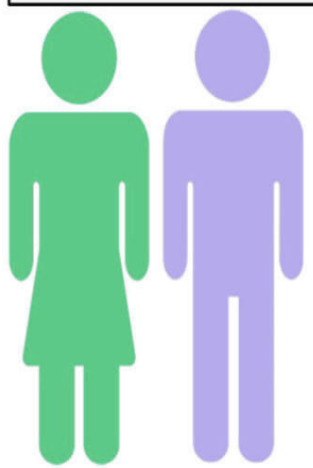


Retained real-world patients



Dropped clinical trial patient

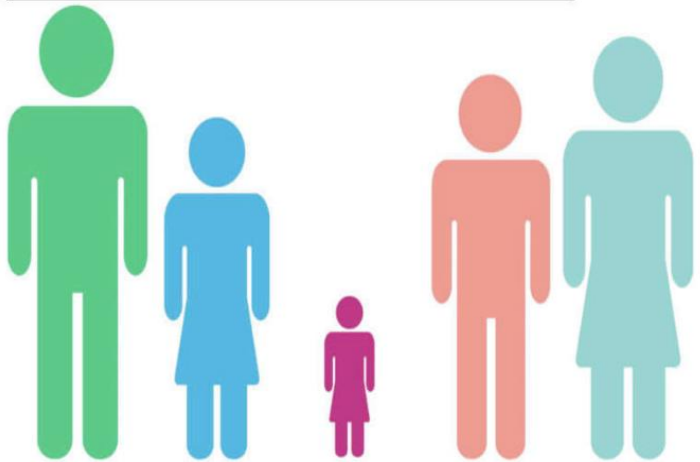
Dropped real-word patients



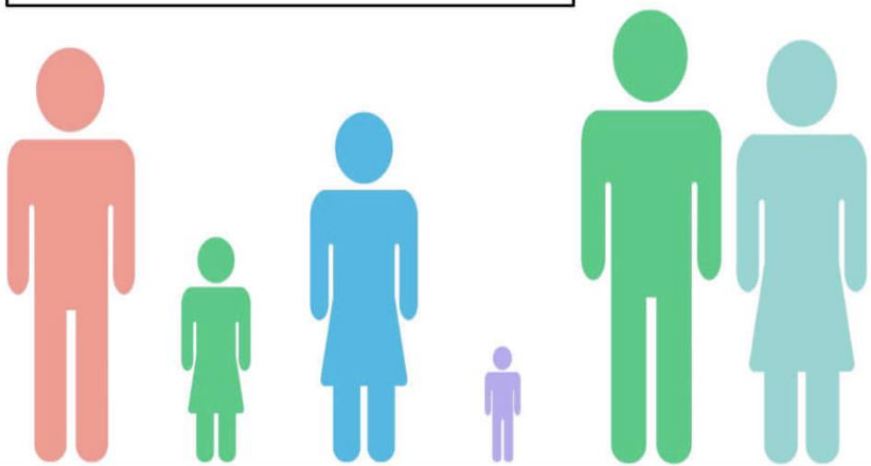
Patients in the clinical trial get "paired up" with a similar real-world patient. Patients that don't find a match get dropped.

Adjusted population after PS weighting

Clinical trial patients



Real-world patients



More similar patients in each group get up-weighted, while the less similar patients get down-weighted. All patients are retained.

Table 1. Baseline demographics and disease characteristics of the elranatamab arm from MagnetisMM-3 Cohort A and the physician's choice arms from COTA before and after inverse probability of treatment weighting.

	Unweighted data before MICE				IPTW sample after MICE†		
	MagnetisMM-3 cohort A (N = 123)	COTA cohort (N = 239)	p-value	SMD	MagnetisMM-3 Cohort A (N = 113)	COTA cohort (N = 253.2)	SMD
Age at index, mean (SD)	67.1 (9.4)	68.0 (9.4)	0.359	0.102	68.5	67.3	0.125
Male, n (%)	68 (55.3)	130 (54.4)	0.872	0.018	49.9	53.7	0.076
White, n (%)	72 (58.5)	175 (73.2)	0.004	**0.314**	67.0	69.7	0.059
ISS disease stage, n (%)							
I	35 (28.5)	31 (13.0)	0.000	**1.258**	20.7	20.4	0.085
II	45 (36.6)	26 (10.9)			20.8	17.9	
III	25 (20.3)	22 (9.2)			12.0	11.9	
Unknown or not assessed	18 (14.6)	160 (66.9)			46.5	49.8	
ECOG performance status, n (%)							
0	45 (36.6)	71 (29.7)	0.013	**0.371**	34.0	31.2	0.066
1	71 (57.7)	129 (54.0)			53.1	55.6	
2	7 (5.7)	39 (16.3)			12.9	13.2	
Time from diagnosis to index, mean (SD), years	6.6 (3.8)	5.4 (4.4)	0.010	**0.280**	6.2	5.7	0.135
Bone lesions during the baseline period or on the index date, n (%)	34 (27.6)	121 (50.6)	0.000	**0.485**	43.8	44.9	0.022
Extramedullary disease, n (%)	38 (30.9)	32 (13.4)	0.000	**0.431**	31.0	22.6	0.192
High-risk cytogenetics (t[4;14], t[14;16], or del[17p]), n (%)	31 (25.2)	49 (20.5)	0.307	0.112	15.3	22.2	0.177
CCI score, n (%)							
2	83 (67.5)	200 (83.7)	0.012	**0.429**	81.6	80.7	0.078
3	21 (17.1)	22 (9.2)			10.1	11.4	
4	11 (8.9)	11 (4.6)			4.9	4.7	
5	6 (4.9)	4 (1.7)			2.8	2.6	
≥6	2 (1.6)	2 (0.8)			0.7	0.7	
Number of LOTs used prior to index date, mean (SD)	5.2 (2.6)	4.9 (2.4)	0.269	0.124	5.3	4.9	0.130
Penta-drug refractory, n (%)	52 (42.3)	45 (18.8)	0.000	**0.526**	32.3	24.6	0.170
SCT during the baseline period, n (%)	87 (70.7)	137 (57.3)	0.013	**0.282**	65.4	62.5	0.060
Aspartate aminotransferase, mean (SD), microkat/l	0.4 (0.2)	0.4 (0.4)	0.799	0.027	0.4	0.4	0.093
Alanine aminotransferase, mean (SD), microkat/l	0.3 (0.3)	0.4 (0.4)	0.024	**0.243**	0.3	0.4	0.198
Hemoglobin, mean (SD), g/l	104.0 (17.1)	105.1 (19.4)	0.581	0.061	103.7	104.4	0.038
Creatinine clearance, mean (SD), ml/min	74.2 (30.8)	71.5 (42.5)	0.520	0.072	72.9	76.2	0.086
Calcium in serum or plasma, mean (SD), mmol/l	2.3 (0.2)	2.3 (0.2)	0.144	0.169	2.3	2.3	0.008
Bilirubin, mean (SD), μmol/l	9.0 (7.0)	8.6 (5.7)	0.577	0.066	9.6	9.1	0.072
Serum albumin, mean (SD), g/dl	36.1 (5.4)	34.5 (5.9)	0.012	**0.280**	35.4	34.9	0.088

Bolded SMD values indicate those over the a priori defined threshold of 0.20.
†The MICE procedure generated five unique data sets with alternative imputed values. The descriptive statistics post IPTW and post MICE reflect the average of the descriptive statistics across these five multiple imputation data sets.
CCI: Charlson Comorbidity Index; ECOG: Eastern Cooperative Oncology Group; IPT: Inverse probability of treatment; IPTW: Inverse probability of treatment weighted; ISS: International Staging System; LOT: Line of therapy; MICE: Multiple imputation by chained equations; SCT: Stem cell transplant; SMD: Standardized mean difference.

Table 2. Baseline demographics and disease characteristics of the elranatamab arm from MagnetisMM-3 Cohort A and the physician's choice arms from Flatiron Health before and after inverse probability of treatment weighting.

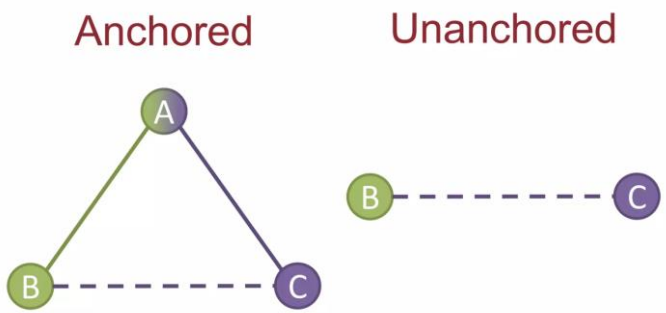
	Unweighted data before MICE				IPTW sample after MICE†		
	MagnetisMM-3 Cohort A (N = 123)	FH cohort (N = 152)	p-value	SMD	MagnetisMM-3 Cohort A (N = 108.9)	FH cohort (N = 150.6)	SMD
Age at index, mean (SD)	67.1 (9.4)	69.5 (10.0)	0.043	**0.246**	68.6	69.2	0.062
Male, n (%)	68 (55.3)	80 (52.6)	0.661	0.053	53.5	55.6	0.041
White, n (%)	72 (58.5)	102 (67.1)	0.143	0.178	69.0	63.2	0.122
ISS disease stage, n (%)							
I	35 (28.5)	11 (7.2)	0.000	**1.309**	19.9	16.6	0.135
II	45 (36.6)	19 (12.5)			25.8	29.1	
III	25 (20.3)	19 (12.5)			13.3	10.5	
Unknown or not assessed	18 (14.6)	103 (67.8)			41.0	43.9	
ECOG performance status, n (%)							
0	45 (36.6)	47 (30.9)	0.030	**0.367**	27.4	30.8	**0.261**
1	71 (57.7)	81 (53.3)			58.7	46.9	
2	7 (5.7)	24 (15.8)			13.9	22.3	
Time from diagnosis to index, mean (SD), years	6.6 (3.8)	4.1 (2.2)	0.000	**0.798**	5.8	4.8	**0.349**
Bone lesions during the baseline period or on the index date, n (%)	34 (27.6)	18 (11.8)	0.001	**0.405**	20.5	23.7	0.078
Extramedullary disease, n (%)	38 (30.9)	–	–	–	–	–	–
High-risk cytogenetics (t[4;14], t[14;16], or del[17p]), n (%)	31 (25.2)	38 (25.0)	0.969	0.005	24.4	22.8	0.037
CCI score, n (%)							
2	83 (67.5)	121 (79.6)	0.189	**0.261**	77.2	70.6	**0.250**
3	21 (17.1)	14 (9.2)			12.4	14.4	
4	11 (8.9)	10 (6.6)			6.1	6.7	
5	6 (4.9)	4 (2.6)			3.3	7.0	
≥6	2 (1.6)	3 (2.0)			1.0	1.2	
Number of LOTs used prior to index date, mean (SD)	5.2 (2.6)	4.0 (1.7)	0.000	**0.555**	5.4	4.4	**0.528**
Penta-drug refractory, n (%)	52 (42.3)	23 (15.1)	0.000	**0.629**	35.6	23.4	**0.270**
SCT during the baseline period, n (%)	87 (70.7)	55 (36.2)	0.000	**0.738**	65.2	44.9	**0.416**
Aspartate aminotransferase, mean (SD), microkat/l	0.4 (0.2)	0.4 (0.5)	0.866	0.020	0.4	0.4	0.061
Alanine aminotransferase, mean (SD), microkat/l	0.3 (0.3)	0.3 (0.3)	0.758	0.037	0.3	0.3	0.030
Hemoglobin, mean (SD), g/l	9.0 (7.0)	8.3 (4.7)	0.377	0.110	9.4	9.4	0.004
Creatinine clearance, mean (SD), ml/min	74.2 (30.8)	62.5 (34.2)	0.003	**0.361**	70.9	71.4	0.014
Calcium in serum or plasma, mean (SD), mmol/l	2.3 (0.2)	2.3 (0.2)	0.700	0.046	2.3	2.3	0.110
Bilirubin, mean (SD), μmol/l	36.1 (5.4)	34.1 (5.3)	0.002	**0.374**	34.7	34.0	0.114
Serum albumin, mean (SD), g/dl	104.0 (17.1)	103.5 (20.5)	0.846	0.023	103.5	104.0	0.025

Bolded SMD values indicate those over the a priori defined threshold of 0.20.
†The MICE procedure generated five unique data sets with alternative imputed values. The descriptive statistics post IPTW and post MICE reflect the average of the descriptive statistics across these five multiple imputation data sets.
CCI: Charlson Comorbidity Index; ECOG: Eastern Cooperative Oncology Group; FH: Flatiron Health; IPT: Inverse probability of treatment; IPTW: Inverse probability of treatment weighed; ISS: International Staging System; LOT: Line of therapy; MICE: Multiple imputation by chained equations; SCT: Stem cell transplant; SMD: Standardized mean difference.

Future Oncol. 2024 Feb 28. doi: 10.2217/fon-2023-0995. Epub ahead of print. PMID: 38415370

Population-adjusted Indirect Comparisons

- Population-adjusted Indirect Comparisons use patient-level data from a trial of a given treatment (referred to as the index trial) to derive a comparison of outcomes with competing treatments, based on published information from similarly designed studies, after adjusting for differences in the characteristics of the populations.
- *in other words*: individual patient data (IPD) in one or more trials are used to adjust for between-trial differences in the distribution of variables that influence outcome
- "anchored" indirect comparison (common comparator arm in each trial) Vs "unanchored" indirect comparison (disconnected treatment network or single-arm studies)
  - *an unanchored comparison assumes that all effect modifiers and prognostic factors are accounted for*
- unanchored methods for population adjustment are problematic and should not be used when anchored methods can be applied



https://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

Network Meta-Analysis (NMA)

(M. Cinquini)

Network meta-analysis

Combines direct and indirect evidence. Also known as:

1) Mixed treatment comparison

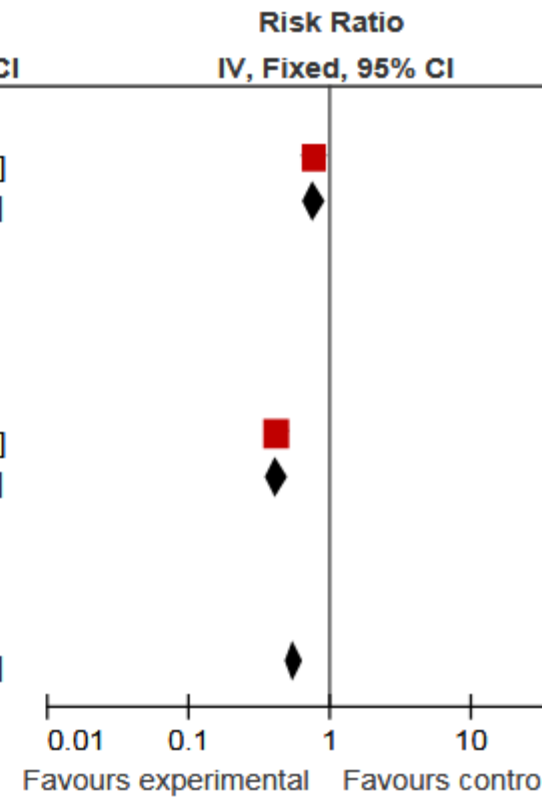
2) Multiple treatment meta-analysis

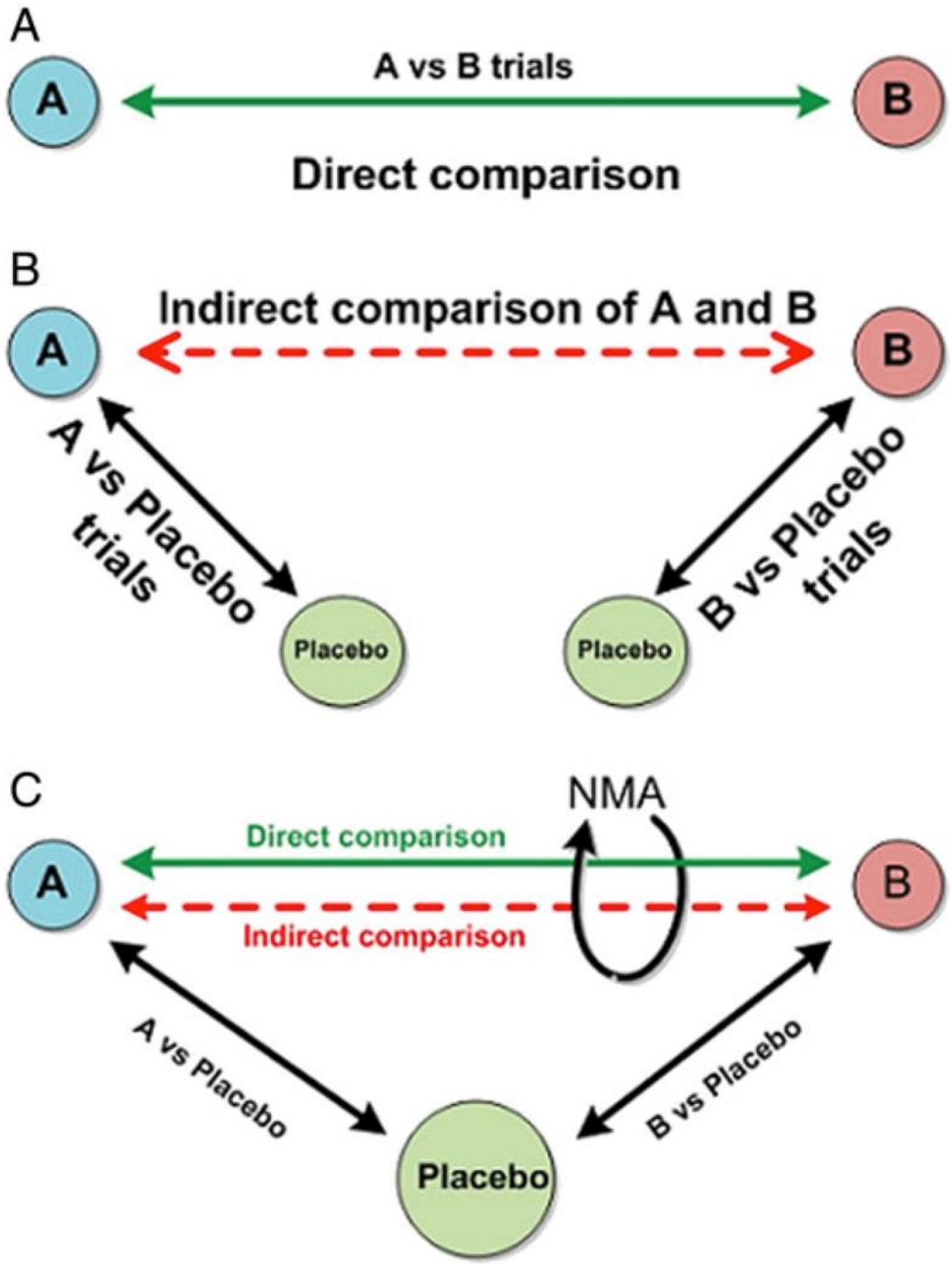
ALL 3 mean the same thing – **simultaneous** comparison of multiple competing treatments using direct & indirect evidence (usually from RCTs) in a single analysis.

SAME assumption as made for indirect comparison alone: the consistency assumption.

Using GIV to combine in RevMan

Study or Subgroup	log[Risk Ratio]	SE	Weight	Risk Ratio IV, Fixed, 95% CI
2.2.1 Direct				
Direct B vs C	-0.2571	0.095	47.9%	0.77 [0.64, 0.93]
Subtotal (95% CI)			**47.9%**	**0.77 [0.64, 0.93]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 2.71 (P = 0.007)				
2.2.2 Indirect				
Indirect B vs C	-0.87	0.091	52.1%	0.42 [0.35, 0.50]
Subtotal (95% CI)			**52.1%**	**0.42 [0.35, 0.50]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 9.56 (P < 0.00001)				
Total (95% CI)			**100.0%**	**0.56 [0.49, 0.64]**
Heterogeneity: Chi² = 21.71, df = 1 (P < 0.00001); I² = 95%				
Test for overall effect: Z = 8.78 (P < 0.00001)				





Consistency Assumption

there must be no relevant discrepancy between direct and indirect evidence

Indirect Comparisons

Basic assumptions underlying indirect comparisons include:

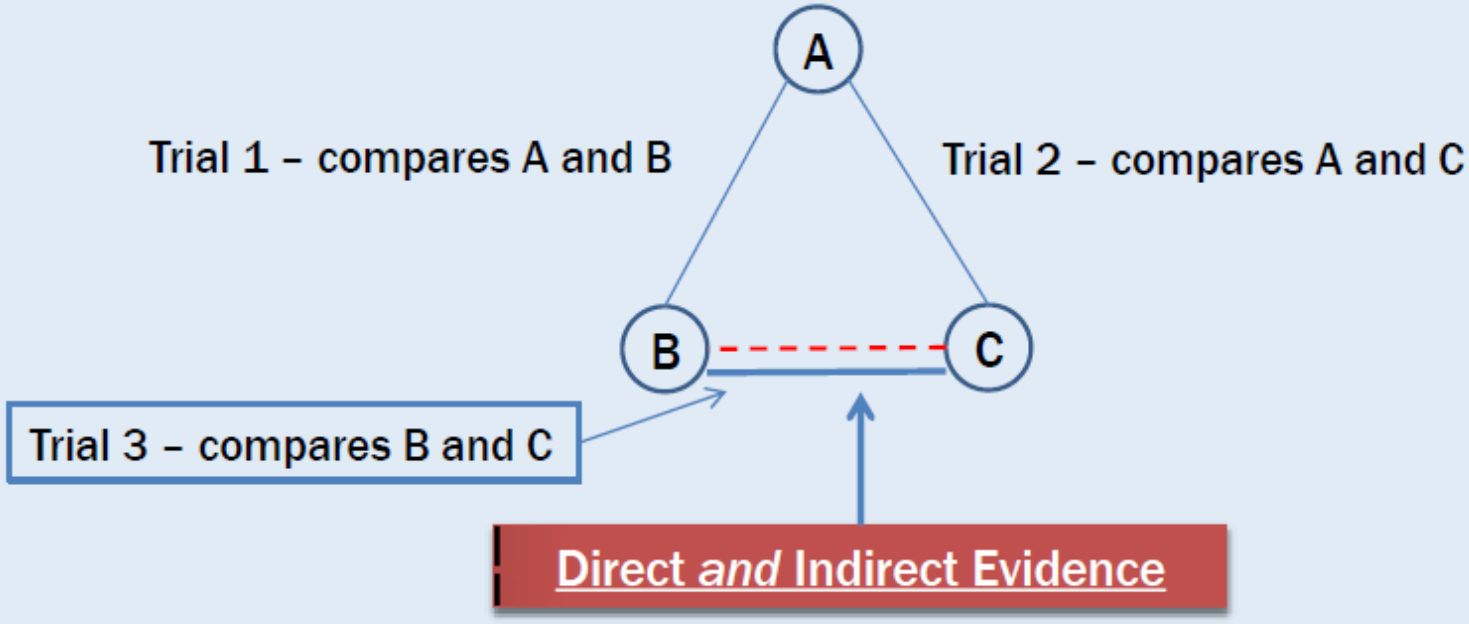
- ✓ **similarity** assumption for adjusted indirect comparison,
- ✓ **homogeneity** assumption for standard meta-analysis and
- ✓ **consistency** assumption for the combination of direct and indirect evidence. It is essential to fully understand and appreciate these basic assumptions in order to use adjusted indirect and mixed treatment comparisons appropriately.

CONSISTENCY ASSUMPTION

- When both direct and indirect evidence is available, an assumption of evidence consistency is required to quantitatively combine the direct and indirect estimates.
- It is important to investigate possible causes of discrepancy between the direct and indirect evidence, such as the play of chance, invalid indirect comparison, bias in head-to-head comparative trials, and clinically meaningful heterogeneity
- When the direct comparison differs from the adjusted indirect comparison, we should usually give more credibility to evidence from head-to-head comparative trials. However, evidence from direct comparative trials may not always be valid.

Song, What is …? 2009; Song et al, J Clin Epidemiol 2008

THERE ARE 2 TYPES OF TRIAL EVIDENCE



Consistency ⇒ Direct and indirect evidence agree

Inconsistency ⇒ Direct and indirect evidence disagree

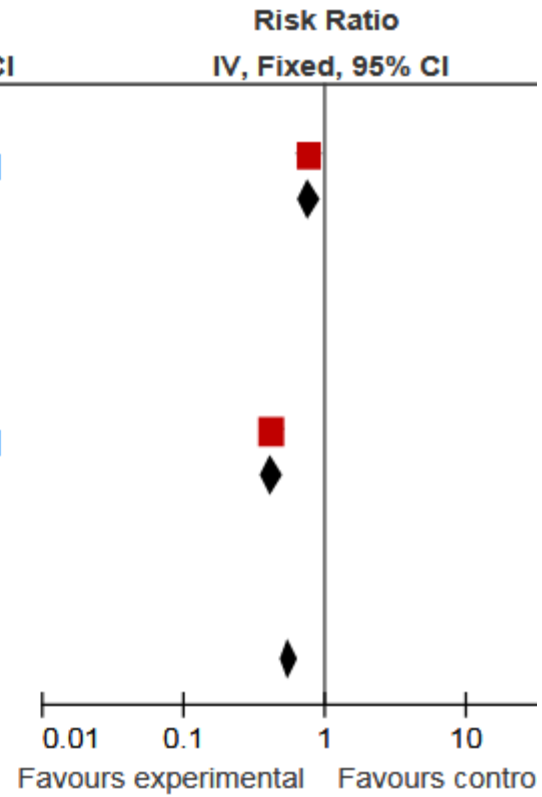
Differing effect modifiers among the trials can cause inconsistency

METHODS TO TEST FOR INCONSISTENCY

1. Bucher method
   - Can be used on triangle structures where three direct estimates are available
   - All such “triangles” should be evaluated one by one
2. Node-splitting
   - Direct and indirect studies are separated and a difference in estimates is calculated
   - Repeated for all treatment comparisons where inconsistency is possible
3. Inconsistency model
   - Could be considered “independence” model because all treatment comparisons are estimated independently
   - Treatment effects are not estimated relative to a reference treatment

DISCUSSION on INDIRECT and DIRECT ESTIMATES

Study or Subgroup	log[Risk Ratio]	SE	Weight	Risk Ratio IV, Fixed, 95% CI
2.2.1 Direct				
Direct B vs C	-0.2571	0.095	47.9%	0.77 [0.64, 0.93]
Subtotal (95% CI)			**47.9%**	**0.77 [0.64, 0.93]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 2.71 (P = 0.007)				
2.2.2 Indirect				
Indirect B vs C	-0.87	0.091	52.1%	0.42 [0.35, 0.50]
Subtotal (95% CI)			**52.1%**	**0.42 [0.35, 0.50]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 9.56 (P < 0.00001)				
Total (95% CI)			**100.0%**	**0.56 [0.49, 0.64]**
Heterogeneity: Chi² = 21.71, df = 1 (P < 0.00001); I² = 95%				
Test for overall effect: Z = 8.78 (P < 0.00001)				



Bucher approach to checking consistency

The difference ω between direct LRR_{BC} and indirect LRR_{BC}

$\hat{\omega}$ = -0.257 - -0.87 = 0.61

To calculate the standard error of the difference we sum the SE from the direct and indirect log risk ratios

$$SE(\Delta) = \sqrt{SE(LLR^{Direct})^2 + SE(LRR^{Indirect})^2}$$

$$= \sqrt{0.095^2 + 0.091^2} \quad = 0.13$$

Bucher approach to checking consistency

Calculate confidence intervals & p-values for : $\hat{\omega}$

95% CI = $\hat{\omega} \pm (1.96*SE)$ = exp [0.36] to exp [0.86]

= **1.43 to 2.37**

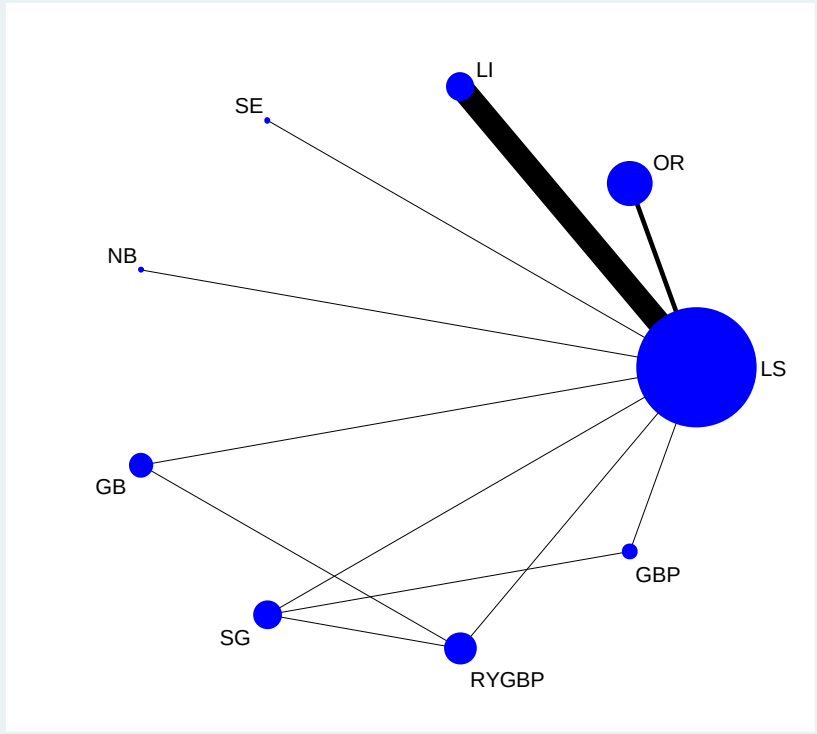
z-score = $\frac{\hat{\omega}}{SE(\hat{\omega})}$ = 4.64 p-value = <0.000002

PROs & CONS

Bucher analyses can be used only when there is a single study per treatment comparison. The Bucher method is suitable, or even ideal, in such situations. However, it can also be used when multiple studies are available for one or more comparisons. If so, estimates from multiple studies for a treatment contrast are pooled into one estimate using classical (pairwise) meta-analysis approach before computing Bucher indirect estimate for a different treatment contrast.

In reality, where the treatment comparisons involve simple networks with two pairwise comparisons or a star-shaped network with a single common comparator, Bucher ITC is likely to provide adequate results.
However, with more complex networks involving closed loops and multi-arm RCTs, the Bucher methodology cannot be applied, as it assumes independence between pairwise comparisons – something not found in multi-arm studies.

NODE SPLITTING



La rete di trattamenti è rappresentata da un network plot, dove i trattamenti sono i **nodi** e i **confronti diretti** tra trattamenti sono le **linee** del grafico.

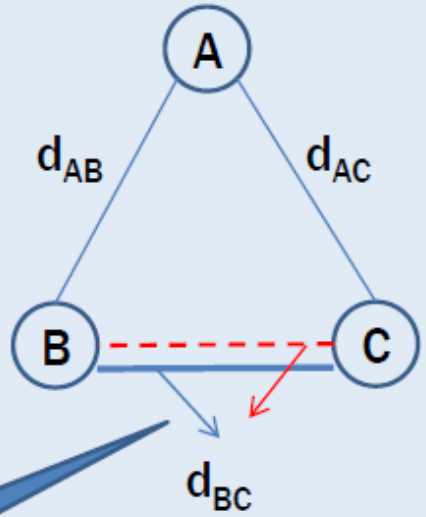
Nodo condiviso: Quando un trattamento appare in più confronti nella rete, può essere visto come un nodo "condiviso". Ad esempio, se il trattamento LS viene confrontato con SG e GBP in studi separati, il trattamento LS è un nodo condiviso tra due confronti.

Problema con il nodo condiviso: Quando si eseguono analisi, l'assunzione di transitività (che significa che i confronti indiretti tra trattamenti sono validi, come se fossero confronti diretti) potrebbe non essere soddisfatta se non trattiamo separatamente i vari utilizzi di un nodo condiviso.

Node Splitting: Il node splitting aiuta a gestire questo problema separando il nodo condiviso in più "versioni", corrispondenti ai diversi confronti. In pratica, si divide il trattamento LS in due nodi distinti: uno per il confronto LS vs SG e uno per LS vs GBP. In questo modo, ogni confronto ha il proprio nodo e viene trattato come un'entità separata.

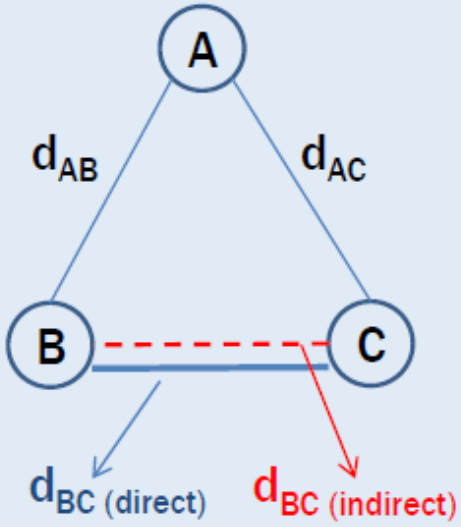
#2 NODE-SPLITTING

Full NMA estimates 3 parameters



Direct and indirect evidence inform this comparison

Node-splitting estimates separate parameters for direct and indirect evidence



Inconsistency is present if
$d_{BC\ (direct)} \neq d_{BC\ (indirect)}$

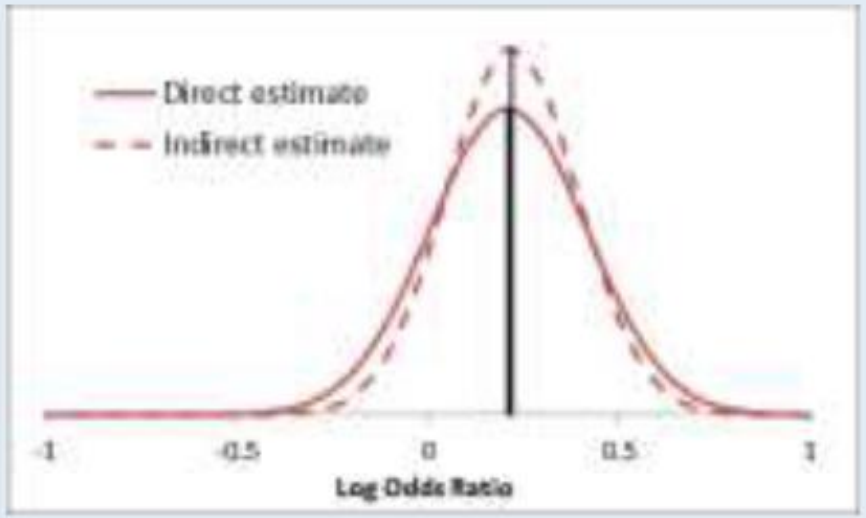
Dias, S, Welton, N, Caldwell, D & Ades, A 2010, 'Checking consistency in mixed treatment comparison meta-analysis'. Statistics in Medicine, vol 29., pp. 932 - 944

Modello statistico: Dopo aver separato i nodi, si utilizza un modello statistico che stima separatamente gli effetti dei trattamenti, prendendo in considerazione i confronti diretti e indiretti, mantenendo la coerenza della rete e riducendo il rischio di distorsioni dovute alla presenza di un nodo condiviso.

#2 NODE-SPLITTING

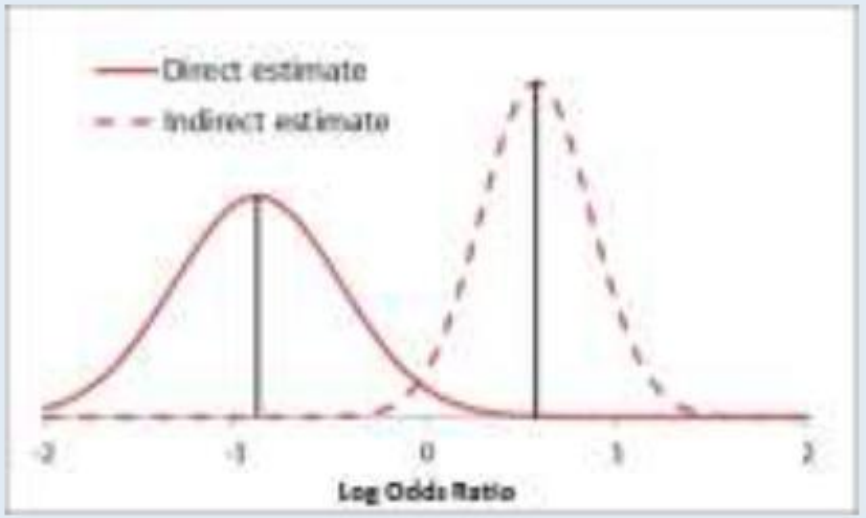
Example of posterior distributions with direct and indirect evidence

Consistent Evidence



Posterior densities overlap indicating absence of inconsistency

Inconsistent Evidence

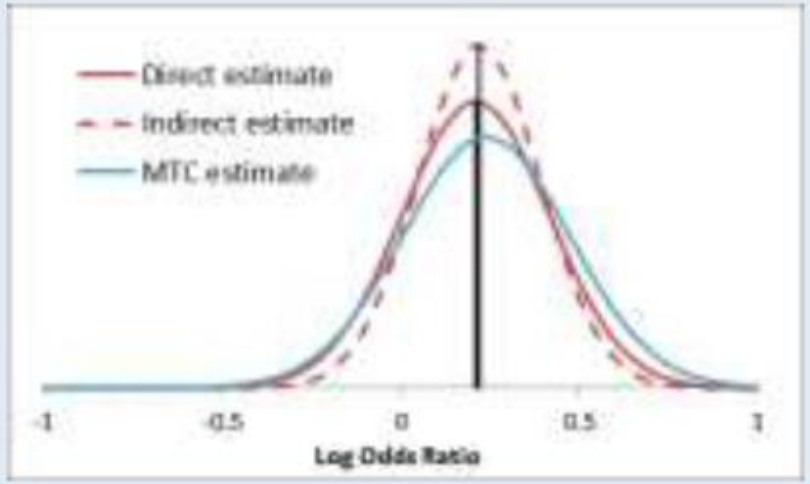


Posterior densities hardly overlap indicating presence of inconsistency

#2 NODE-SPLITTING

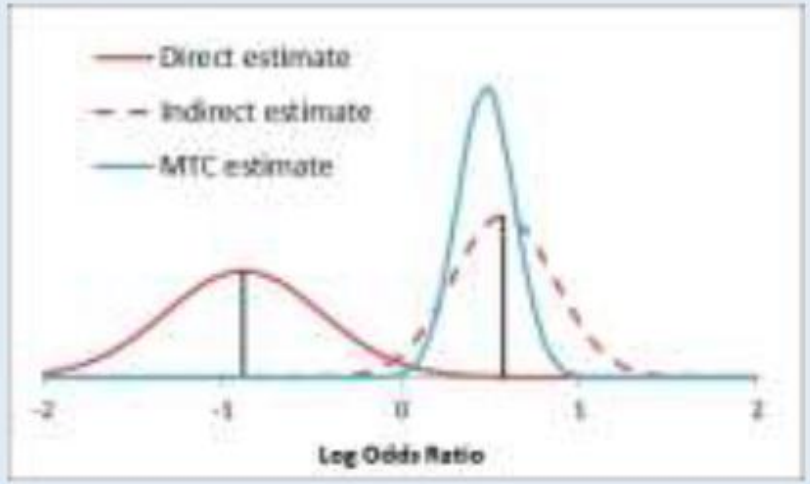
What do we do with this information?

Consistent Evidence



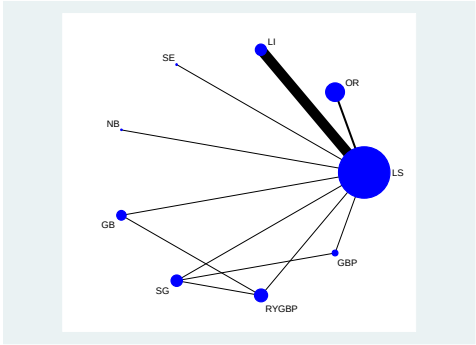
MTC estimate is similar to direct and indirect estimates

Inconsistent Evidence



MTC estimate is closer to indirect estimate, possibly because indirect trials are larger and more precise

Step 1: generating network geometry



Step 2: testing for inconsistency

Table 1. Inconsistency test between direct and indirect treatment comparisons in mixed treatment comparison

Side	Direct		Indirect		Difference		p>z
	Coefficient	SE	Coefficient	SE	Coefficient	SE	
A B	-1.083	0.174	-0.877	0.620	-0.206	0.636	0.746
A C	-1.388	0.247	-1.869	0.493	0.481	0.542	0.375
A D	-1.378	0.265	-0.738	0.413	-0.640	0.479	0.182
A E	-3.425	0.940	-3.221	1.005	-0.204	0.937	0.828
B C	-0.894	0.655	-0.312	0.297	-0.581	0.715	0.416
B D	0.099	0.462	-0.241	0.329	0.340	0.567	0.548
B E	-2.152	0.881	-2.615	1.087	0.463	0.896	0.605
C D	0.490	0.492	0.177	0.350	0.313	0.604	0.605
D E	-2.550	1.254	-1.956	0.958	-0.595	1.314	0.651

SE, standard error; A, placebo; B, IV_single; C, IV_double; D, topical; E, combination.

```
Multivariate meta-analysis
Variance-covariance matrix = proportional .5*I(4)+.5*J(4,4,1)
Method = reml                                 Number of dimensions    =      4
Restricted log likelihood = -30.939719        Number of observations  =     25

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_y_B         |
      des_ABC |   .2528377   .5704516     0.44   0.658    -.8652269    1.370902
      des_ABD |  -.7433714   .5269164    -1.41   0.158    -1.776108    .2893657
      des_ABE |  -.1959024   .5311986    -0.37   0.712    -1.237033    .8452278
        _cons |  -.9727775   .2201655    -4.42   0.000    -1.404294   -.5412611
-------------+----------------------------------------------------------------
_y_C         |
      des_AC |    .217719   .6845858     0.32   0.750    -1.124045    1.559483
        _cons |  -1.58294   .6293945    -2.52   0.012    -2.816531   -.3493498
-------------+----------------------------------------------------------------
_y_D         |
      des_AD |   .5489224   .5775957     0.95   0.342    -.5831443    1.680989
     des_BDE |   1.020097   .9029483     1.13   0.259    -.7496496    2.789843
      des_CD |    .633251   .9312281     0.68   0.496    -1.191923    2.458425
        _cons |  -1.72662   .4786004    -3.61   0.000     -2.66466   -.7885806
-------------+----------------------------------------------------------------
_y_E         |
     des_BDE |   .4401131   1.862385     0.24   0.813    -3.210095    4.090321
        _cons |  -3.402272   1.051331    -3.24   0.001    -5.462844     -1.3417
------------------------------------------------------------------------------

Estimated between-studies SDs and correlation matrix:
                  SD        _y_B        _y_C        _y_D        _y_E
_y_B    1.767e-09            1           .           .           .
_y_C    1.767e-09           .5           1           .           .
_y_D    1.767e-09           .5          .5           1           .
_y_E    1.767e-09           .5          .5          .5           1

Testing for inconsistency:
  1)  [_y_B]des_ABC = 0
  2)  [_y_B]des_ABD = 0
  3)  [_y_B]des_ABE = 0
  4)  [_y_C]des_AC = 0
  5)  [_y_D]des_AD = 0
  6)  [_y_D]des_BDE = 0
  7)  [_y_E]des_BDE = 0
  8)  [_y_D]des_CD = 0

           chi2(  8) =    4.00
         Prob > chi2 =    0.8567
```

Step 3: creating plots and league table of effect size by treatment

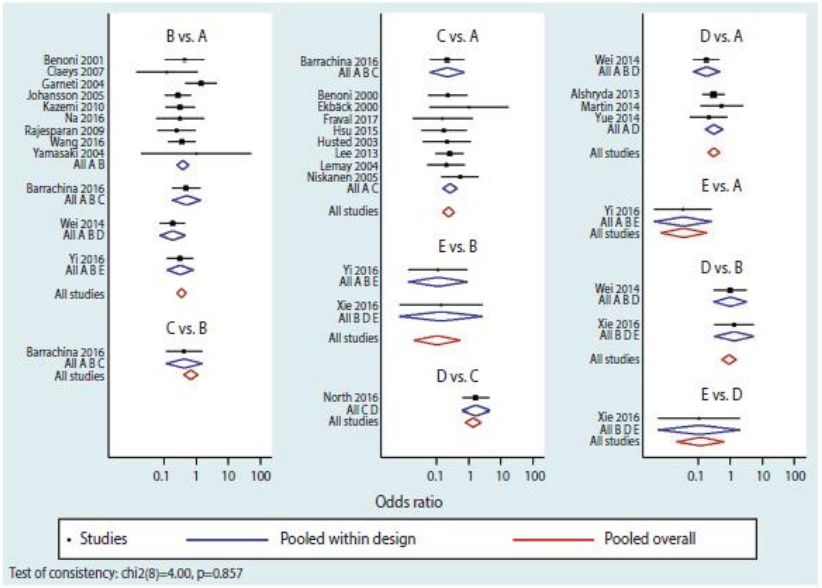


Figure 4. Network forest plot. A, placebo; B, IV_single; C, IV_double; D, topical; E, combination.

Step 4: determining relative rankings of treatment

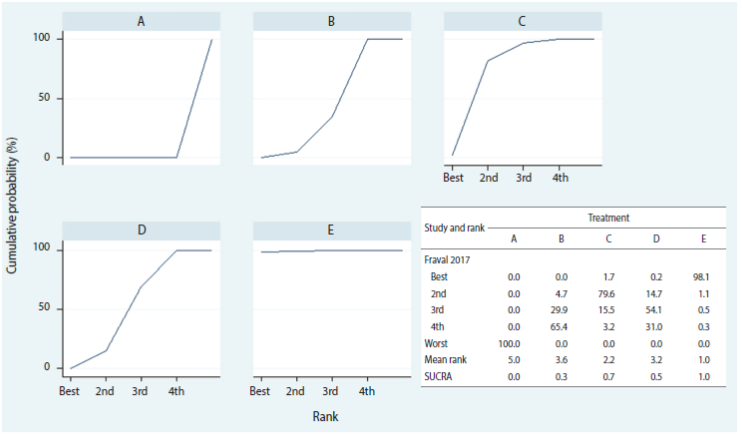


Figure 6. Results of network rank test. A, placebo; B, IV_single; C, IV_double; D, topical; E, combination; SCURA, surface under the cumulative ranking.

ref	-21.74 (-26.96,-16.52)	-24.71 (-27.96,-21.46)	-18.05 (-21.66,-14.44)	-12.26 (-15.55,-8.96)	-4.50 (-8.17,-0.84)	-9.83 (-14.12,-5.55)	-3.85 (-6.19,-1.51)	-2.93 (-4.34,-1.52)
21.74 (16.52,26.96)	_y_I	-2.97 (-8.43,2.49)	3.70 (-1.01,8.40)	9.49 (3.44,15.53)	17.24 (10.86,23.62)	11.91 (5.16,18.67)	17.89 (12.17,23.61)	18.81 (13.41,24.21)
24.71 (21.46,27.96)	2.97 (-2.49,8.43)	_y_H	6.66 (3.18,10.14)	12.45 (8.28,16.62)	20.21 (15.30,25.11)	14.88 (9.49,20.26)	20.86 (16.85,24.87)	21.78 (18.25,25.31)
18.05 (14.44,21.66)	-3.70 (-8.40,1.01)	-6.66 (-10.14,-3.18)	_y_G	5.79 (1.14,10.44)	13.54 (8.40,18.69)	8.21 (2.61,13.82)	14.20 (9.89,18.50)	15.11 (11.25,18.98)
12.26 (8.96,15.55)	-9.49 (-15.53,-3.44)	-12.45 (-16.62,-8.28)	-5.79 (-10.44,-1.14)	_y_F	7.76 (2.83,12.68)	2.43 (-2.98,7.83)	8.41 (4.37,12.45)	9.33 (5.84,12.81)
4.50 (0.84,8.17)	-17.24 (-23.62,-10.86)	-20.21 (-25.11,-15.30)	-13.54 (-18.69,-8.40)	-7.76 (-12.68,-2.83)	_y_E	-5.33 (-10.97,0.31)	0.65 (-3.70,5.00)	1.57 (-2.36,5.50)
9.83 (5.55,14.12)	-11.91 (-18.67,-5.16)	-14.88 (-20.26,-9.49)	-8.21 (-13.82,-2.61)	-2.43 (-7.83,2.98)	5.33 (-0.31,10.97)	_y_D	5.98 (1.10,10.87)	6.90 (2.39,11.41)
3.85 (1.51,6.19)	-17.89 (-23.61,-12.17)	-20.86 (-24.87,-16.85)	-14.20 (-18.50,-9.89)	-8.41 (-12.45,-4.37)	-0.65 (-5.00,3.70)	-5.98 (-10.87,-1.10)	_y_C	0.92 (-1.82,3.65)
2.93 (1.52,4.34)	-18.81 (-24.21,-13.41)	-21.78 (-25.31,-18.25)	-15.11 (-18.98,-11.25)	-9.33 (-12.81,-5.84)	-1.57 (-5.50,2.36)	-6.90 (-11.41,-2.39)	-0.92 (-3.65,1.82)	_y_B

Presenting the data

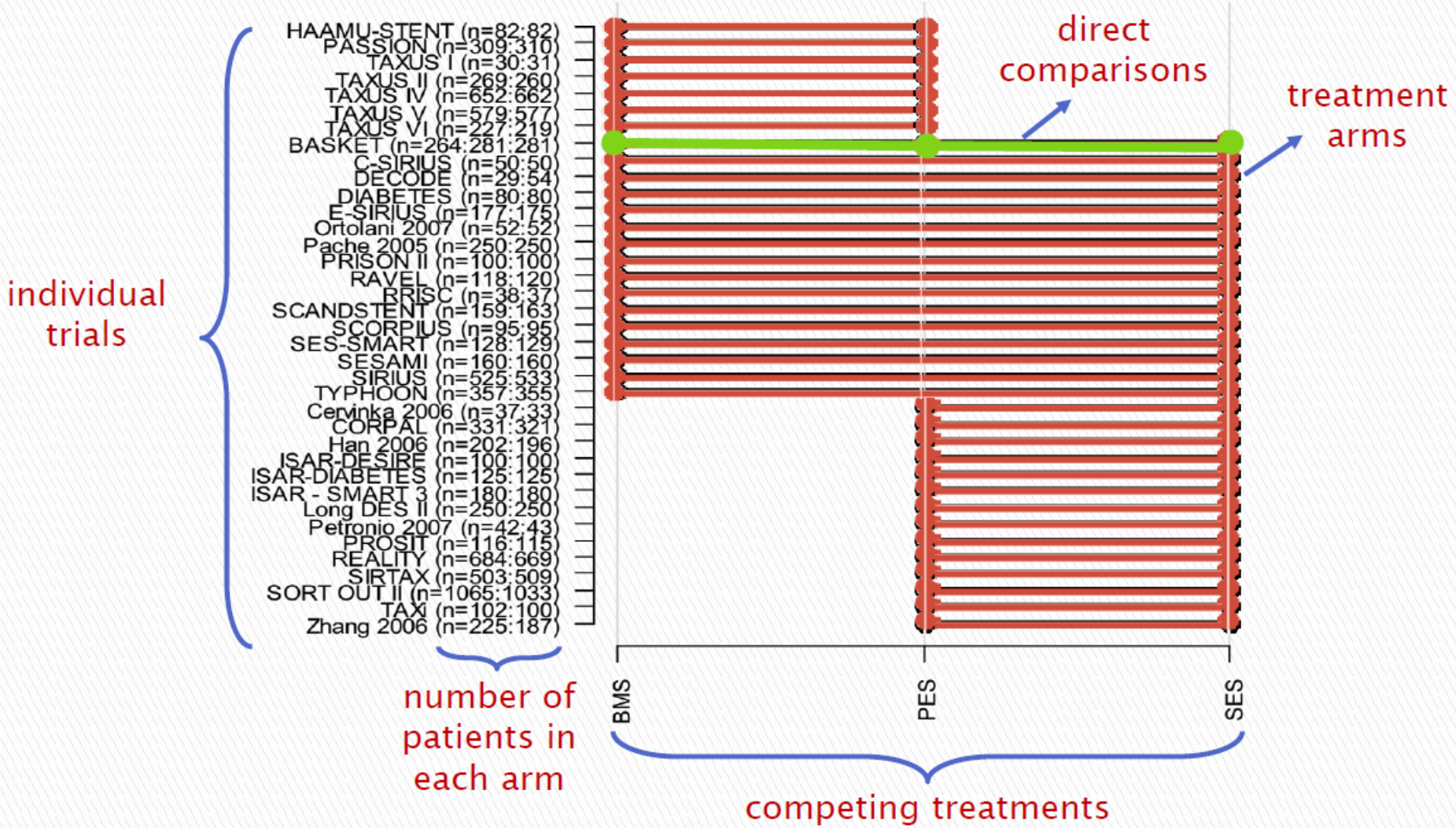
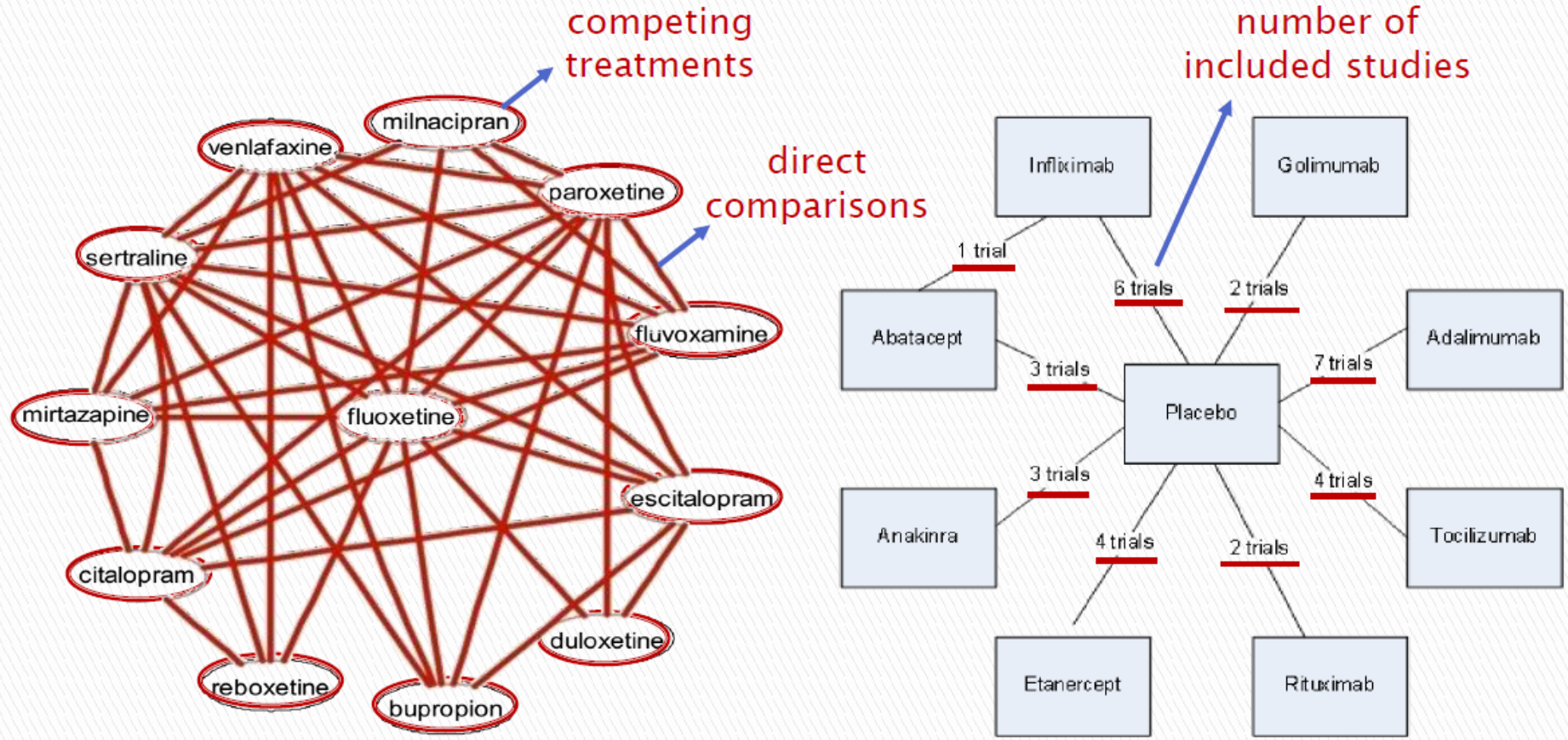


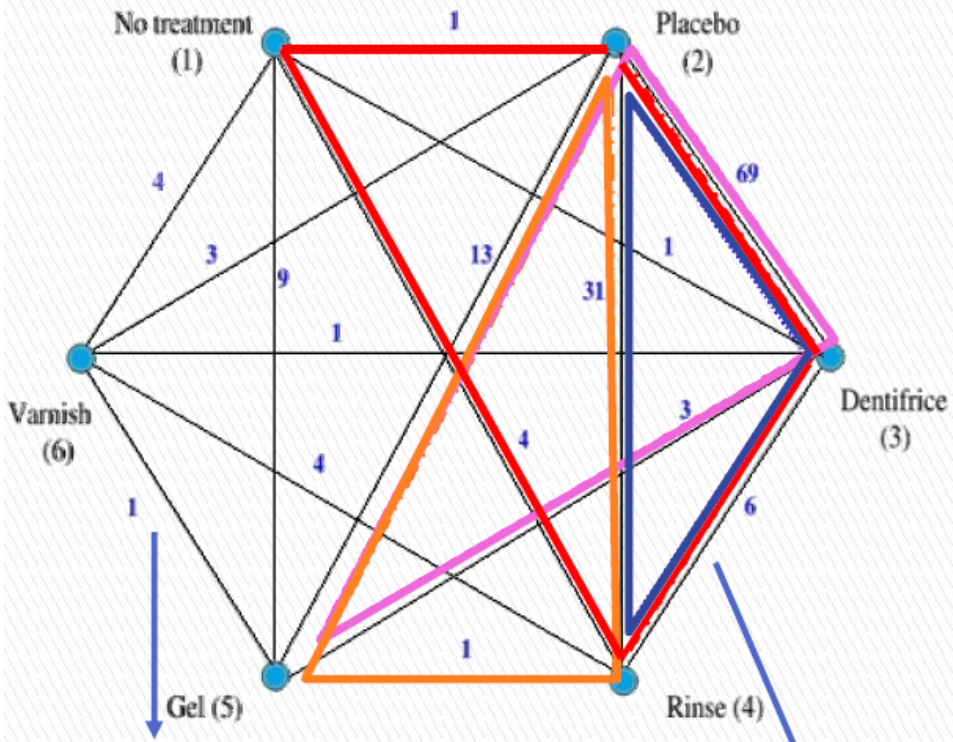
Diagram showing the comparisons involved in the individual studies of the network

[Example in Hoaglin et al. 2011]



Network graph showing the available direct comparisons in the network

[Examples in Hoaglin et al. 2011 & Jonas et al. 2013]



Type (g)	No. of studies	No. trt (1)	Placebo (2)	Dentifrice (3)	Rinse (4)	Gel (5)	Varnish (6)
1	9	X				X	
2	3	X			X		
3	4	X					X
4	61		X	X			
5	9		X			X	
6	25		X		X		
7	3		X				X
8	1			X	X		
9	1			X		X	
10	1					X	X
11	4				X		X
12	4		X	X	X		
13	3		X	X		X	
14	1		X		X	X	
15	1	X	X	X	X		

direct comparisons from two-arm trials

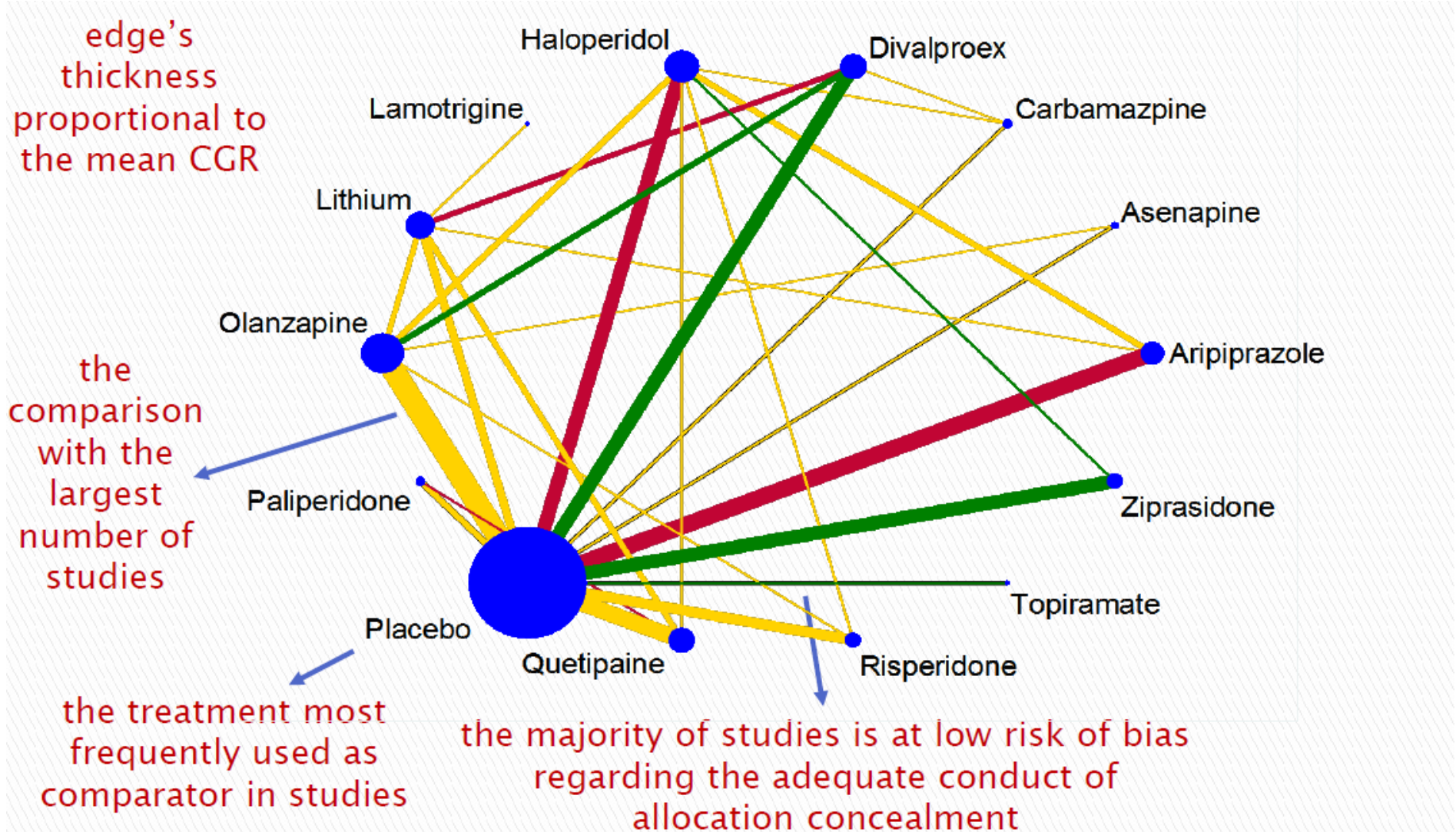
direct comparisons from multi-arm trials

15 different study designs

treatment arms in each design

Network graph showing the presence of multi-arm trials & table showing the network structure; the available study designs in the network

[Examples in Lu et al. 2011]

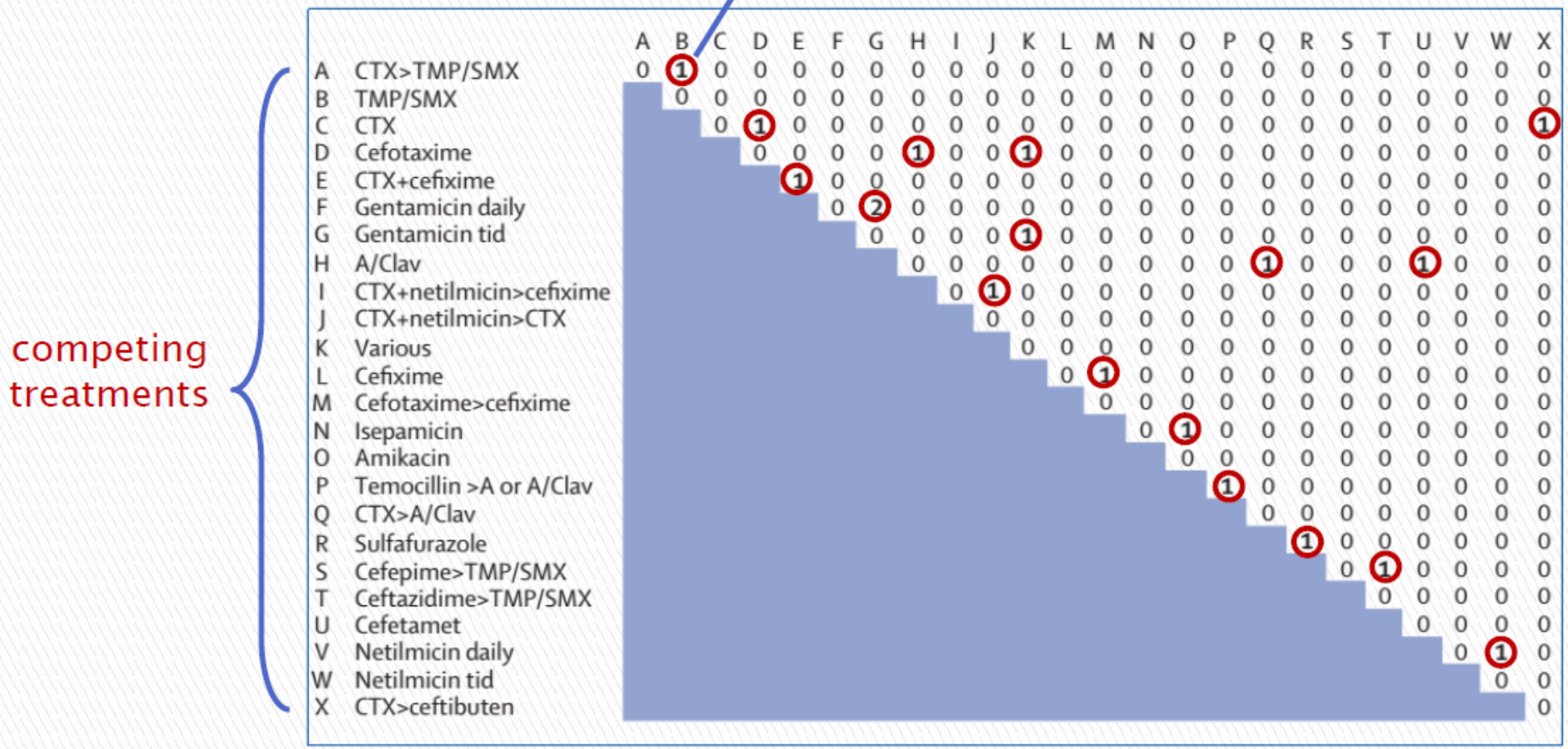


Network graph with weighted and/or colored nodes and edges

[Examples in Chaimani et al. 2013]

number of trials comparing the treatments in the respective row and column

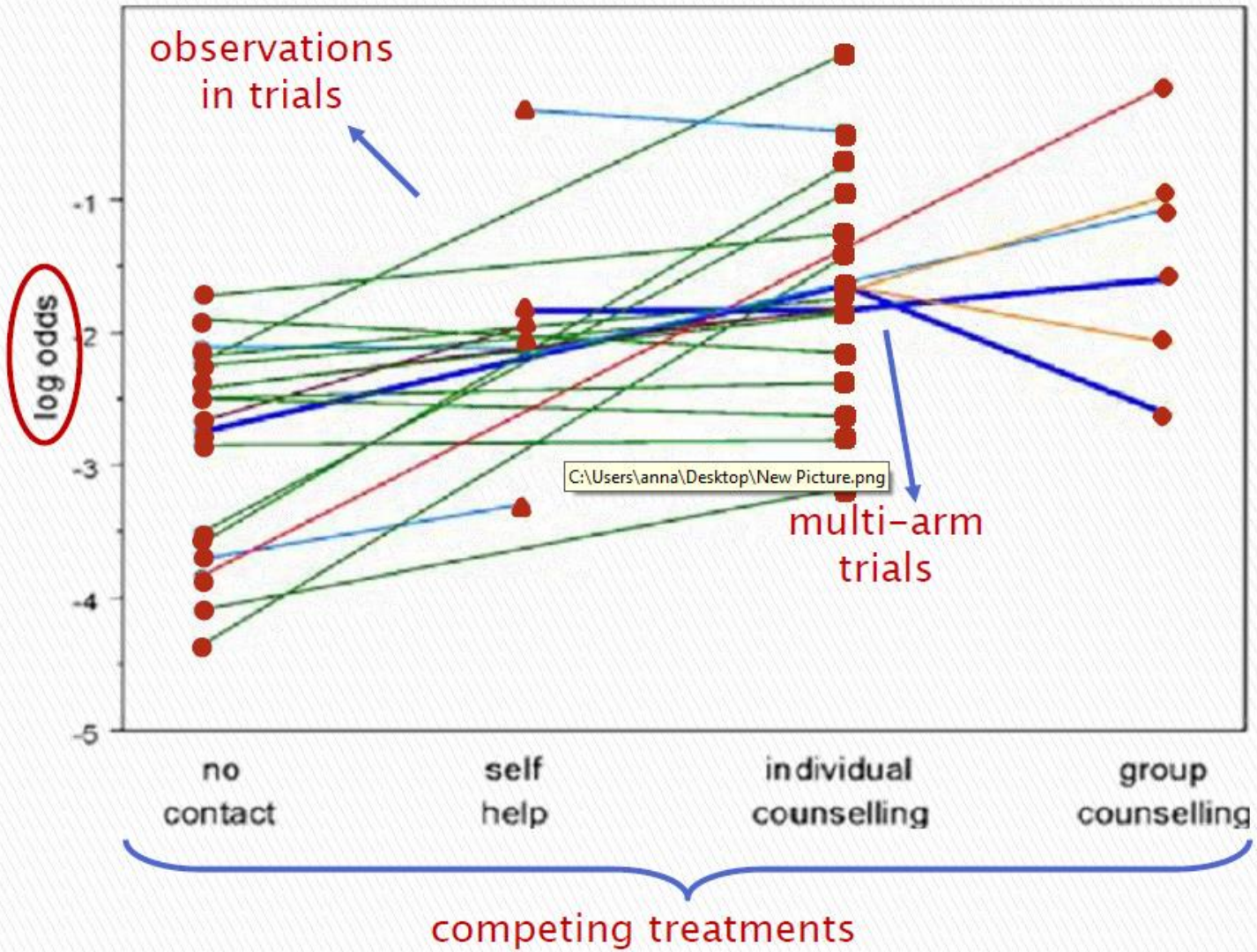
competing treatments



Matrix showing the available direct comparisons in the network

[Example in Ioannidis 2006]

the slope of the lines shows which treatments are favored in studies

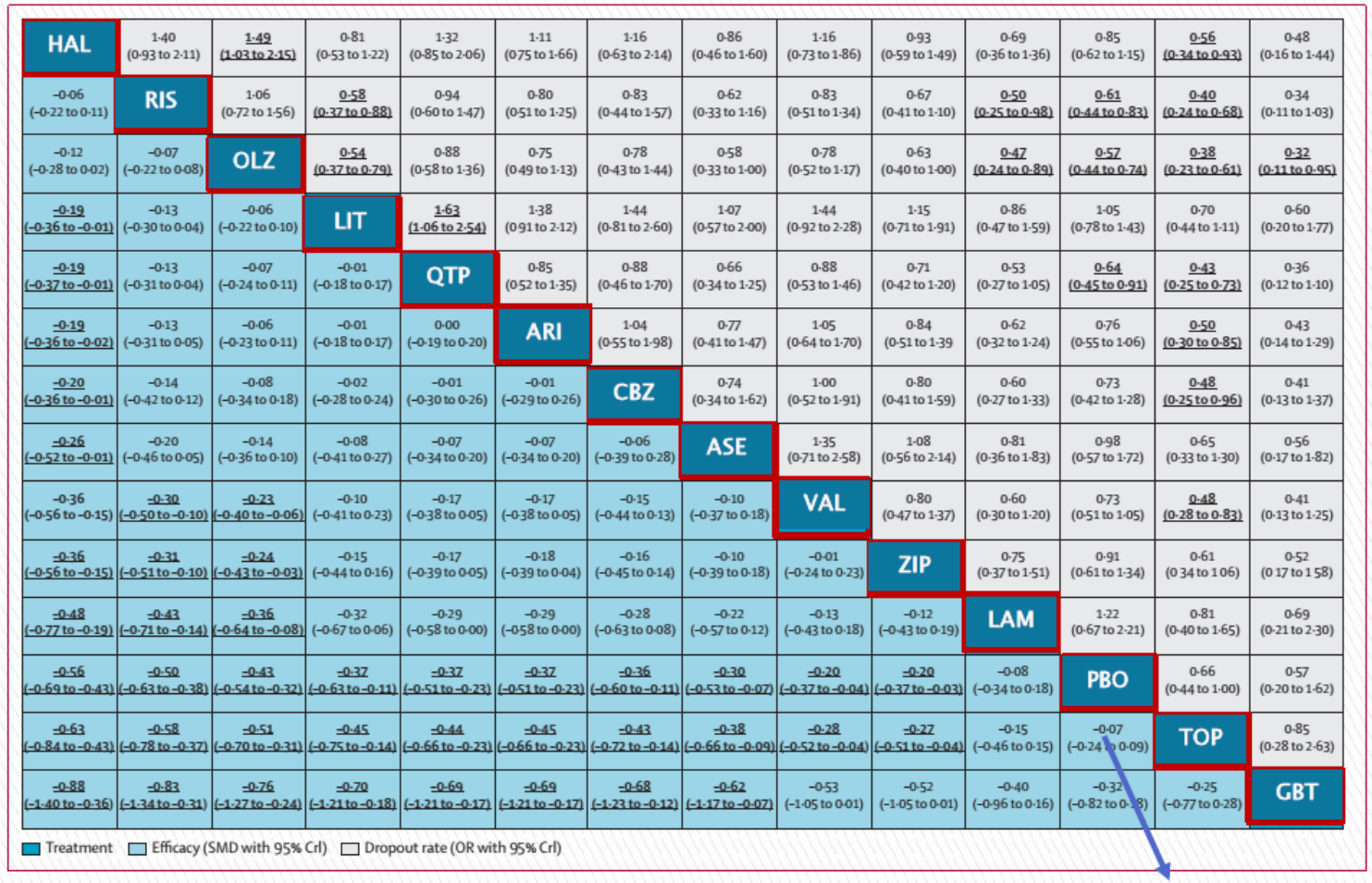


Graph showing the data provided by the individual studies of the network

[Example in Lu & Ades 2006]

Presenting the results measures of effect

relative treatment effects for efficacy
SMD<0 favor the treatment in column



HAL	1·40 (0·93 to 2·11)	**1·49 (1·03 to 2·15)**	0·81 (0·53 to 1·22)	1·32 (0·85 to 2·06)	1·11 (0·75 to 1·66)	1·16 (0·63 to 2·14)	0·86 (0·46 to 1·60)	1·16 (0·73 to 1·86)	0·93 (0·59 to 1·49)	0·69 (0·36 to 1·36)	0·85 (0·62 to 1·15)	**0·56 (0·34 to 0·93)**	0·48 (0·16 to 1·44)
−0·06 (−0·22 to 0·11)	RIS	1·06 (0·72 to 1·56)	**0·58 (0·37 to 0·88)**	0·94 (0·60 to 1·47)	0·80 (0·51 to 1·25)	0·83 (0·44 to 1·57)	0·62 (0·33 to 1·16)	0·83 (0·51 to 1·34)	0·67 (0·41 to 1·10)	**0·50 (0·25 to 0·98)**	**0·61 (0·44 to 0·83)**	**0·40 (0·24 to 0·68)**	0·34 (0·11 to 1·03)
−0·12 (−0·28 to 0·02)	−0·07 (−0·22 to 0·08)	OLZ	**0·54 (0·37 to 0·79)**	0·88 (0·58 to 1·36)	0·75 (0·49 to 1·13)	0·78 (0·43 to 1·44)	0·58 (0·33 to 1·00)	0·78 (0·52 to 1·17)	0·63 (0·40 to 1·00)	**0·47 (0·24 to 0·89)**	**0·57 (0·44 to 0·74)**	**0·38 (0·23 to 0·61)**	**0·32 (0·11 to 0·95)**
−0·19 (−0·36 to −0·01)	−0·13 (−0·30 to 0·04)	−0·06 (−0·22 to 0·10)	LIT	**1·63 (1·06 to 2·54)**	1·38 (0·91 to 2·12)	1·44 (0·81 to 2·60)	1·07 (0·57 to 2·00)	1·44 (0·92 to 2·28)	1·15 (0·71 to 1·91)	0·86 (0·47 to 1·59)	1·05 (0·78 to 1·43)	0·70 (0·44 to 1·11)	0·60 (0·20 to 1·77)
−0·19 (−0·37 to −0·01)	−0·13 (−0·31 to 0·04)	−0·07 (−0·24 to 0·11)	−0·01 (−0·18 to 0·17)	QTP	0·85 (0·52 to 1·35)	0·88 (0·46 to 1·70)	0·66 (0·34 to 1·25)	0·88 (0·53 to 1·46)	0·71 (0·42 to 1·20)	0·53 (0·27 to 1·05)	**0·64 (0·45 to 0·91)**	**0·43 (0·25 to 0·73)**	0·36 (0·12 to 1·10)
−0·19 (−0·36 to −0·02)	−0·13 (−0·31 to 0·05)	−0·06 (−0·23 to 0·11)	−0·01 (−0·18 to 0·17)	0·00 (−0·19 to 0·20)	ARI	1·04 (0·55 to 1·98)	0·77 (0·41 to 1·47)	1·05 (0·64 to 1·70)	0·84 (0·51 to 1·39	0·62 (0·32 to 1·24)	0·76 (0·55 to 1·06)	**0·50 (0·30 to 0·85)**	0·43 (0·14 to 1·29)
−0·20 (−0·36 to −0·01)	−0·14 (−0·42 to 0·12)	−0·08 (−0·34 to 0·18)	−0·02 (−0·28 to 0·24)	−0·01 (−0·30 to 0·26)	−0·01 (−0·29 to 0·26)	CBZ	0·74 (0·34 to 1·62)	1·00 (0·52 to 1·91)	0·80 (0·41 to 1·59)	0·60 (0·27 to 1·33)	0·73 (0·42 to 1·28)	**0·48 (0·25 to 0·96)**	0·41 (0·13 to 1·37)
−0·26 (−0·52 to −0·01)	−0·20 (−0·46 to 0·05)	−0·14 (−0·36 to 0·10)	−0·08 (−0·41 to 0·27)	−0·07 (−0·34 to 0·20)	−0·07 (−0·34 to 0·20)	−0·06 (−0·39 to 0·28)	ASE	1·35 (0·71 to 2·58)	1·08 (0·56 to 2·14)	0·81 (0·36 to 1·83)	0·98 (0·57 to 1·72)	0·65 (0·33 to 1·30)	0·56 (0·17 to 1·82)
−0·36 (−0·56 to −0·15)	**−0·30 (−0·50 to −0·10)**	**−0·23 (−0·40 to −0·06)**	−0·10 (−0·41 to 0·23)	−0·17 (−0·38 to 0·05)	−0·17 (−0·38 to 0·05)	−0·15 (−0·44 to 0·13)	−0·10 (−0·37 to 0·18)	VAL	0·80 (0·47 to 1·37)	0·60 (0·30 to 1·20)	0·73 (0·51 to 1·05)	**0·48 (0·28 to 0·83)**	0·41 (0·13 to 1·25)
−0·36 (−0·56 to −0·15)	**−0·31 (−0·51 to −0·10)**	**−0·24 (−0·43 to −0·03)**	−0·15 (−0·44 to 0·16)	−0·17 (−0·39 to 0·05)	−0·18 (−0·39 to 0·04)	−0·16 (−0·45 to 0·14)	−0·10 (−0·39 to 0·18)	−0·01 (−0·24 to 0·23)	ZIP	0·75 (0·37 to 1·51)	0·91 (0·61 to 1·34)	0·61 (0·34 to 1·06)	0·52 (0·17 to 1·58)
−0·48 (−0·77 to −0·19)	**−0·43 (−0·71 to −0·14)**	**−0·36 (−0·64 to −0·08)**	−0·32 (−0·67 to 0·06)	−0·29 (−0·58 to 0·00)	−0·29 (−0·58 to 0·00)	−0·28 (−0·63 to 0·08)	−0·22 (−0·57 to 0·12)	−0·13 (−0·43 to 0·18)	−0·12 (−0·43 to 0·19)	LAM	1·22 (0·67 to 2·21)	0·81 (0·40 to 1·65)	0·69 (0·21 to 2·30)
−0·56 (−0·69 to −0·43)	**−0·50 (−0·63 to −0·38)**	**−0·43 (−0·54 to −0·32)**	**−0·37 (−0·63 to −0·11)**	**−0·37 (−0·51 to −0·23)**	**−0·37 (−0·51 to −0·23)**	**−0·36 (−0·60 to −0·11)**	**−0·30 (−0·53 to −0·07)**	**−0·20 (−0·37 to −0·04)**	**−0·20 (−0·37 to −0·03)**	−0·08 (−0·34 to 0·18)	PBO	0·66 (0·44 to 1·00)	0·57 (0·20 to 1·62)
−0·63 (−0·84 to −0·43)	**−0·58 (−0·78 to −0·37)**	**−0·51 (−0·70 to −0·31)**	**−0·45 (−0·75 to −0·14)**	**−0·44 (−0·66 to −0·23)**	**−0·45 (−0·66 to −0·23)**	**−0·43 (−0·72 to −0·14)**	**−0·38 (−0·66 to −0·09)**	**−0·28 (−0·52 to −0·04)**	**−0·27 (−0·51 to −0·04)**	−0·15 (−0·46 to 0·15)	−0·07 (−0·24 to 0·09)	TOP	0·85 (0·28 to 2·63)
−0·88 (−1·40 to −0·36)	**−0·83 (−1·34 to −0·31)**	**−0·76 (−1·27 to −0·24)**	**−0·70 (−1·21 to −0·18)**	**−0·69 (−1·21 to −0·17)**	**−0·69 (−1·21 to −0·17)**	**−0·68 (−1·23 to −0·12)**	**−0·62 (−1·17 to −0·07)**	−0·53 (−1·05 to 0·01)	−0·52 (−1·05 to 0·01)	−0·40 (−0·96 to 0·16)	−0·32 (−0·82 to 0·18)	−0·25 (−0·77 to 0·28)	GBT

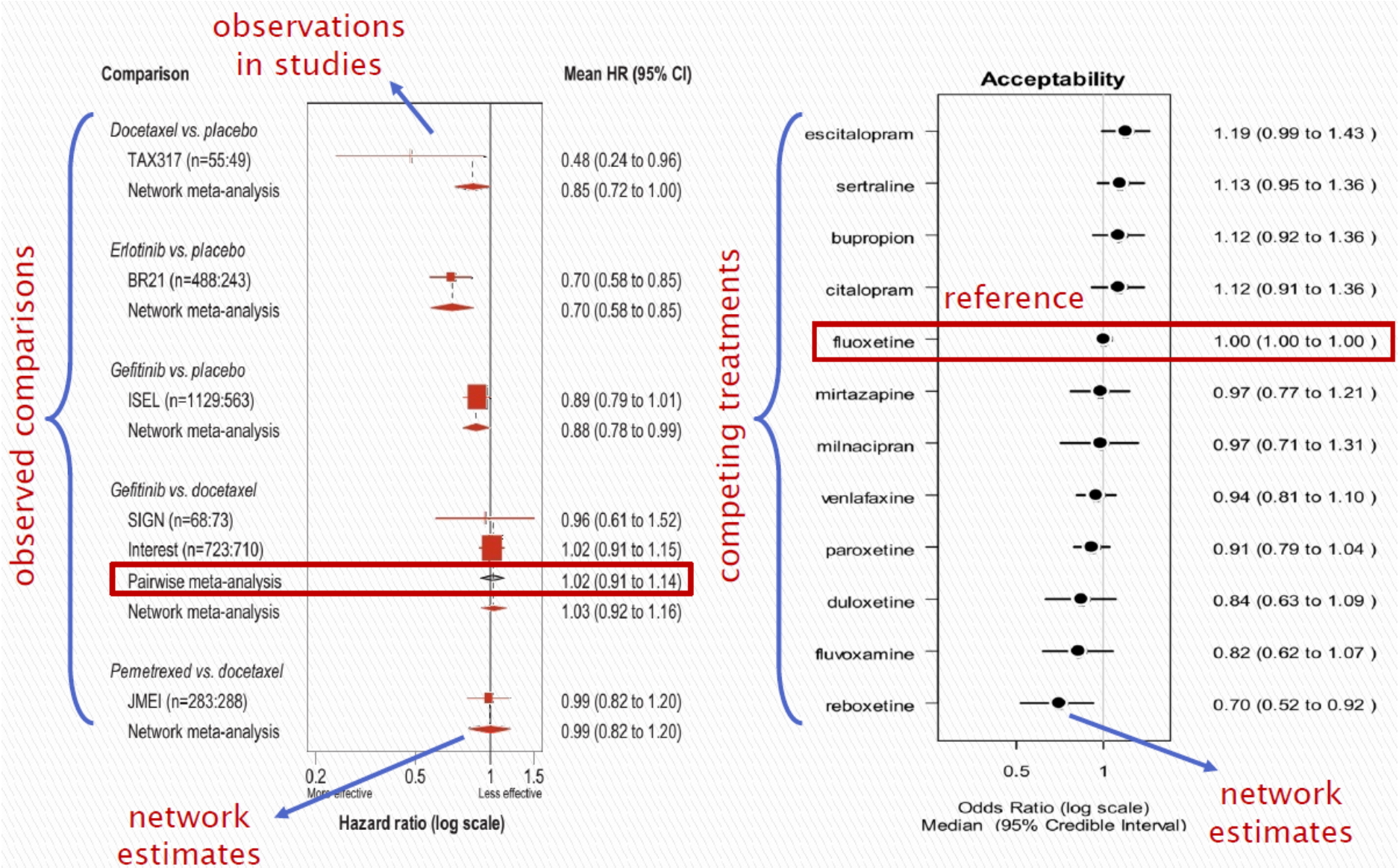
Treatment Efficacy (SMD with 95% CrI) Dropout rate (OR with 95% CrI)

relative treatment effects for dropout rate
OR>1 favor the treatment in column

significant effects are in bold and competing treatments *underscored font*

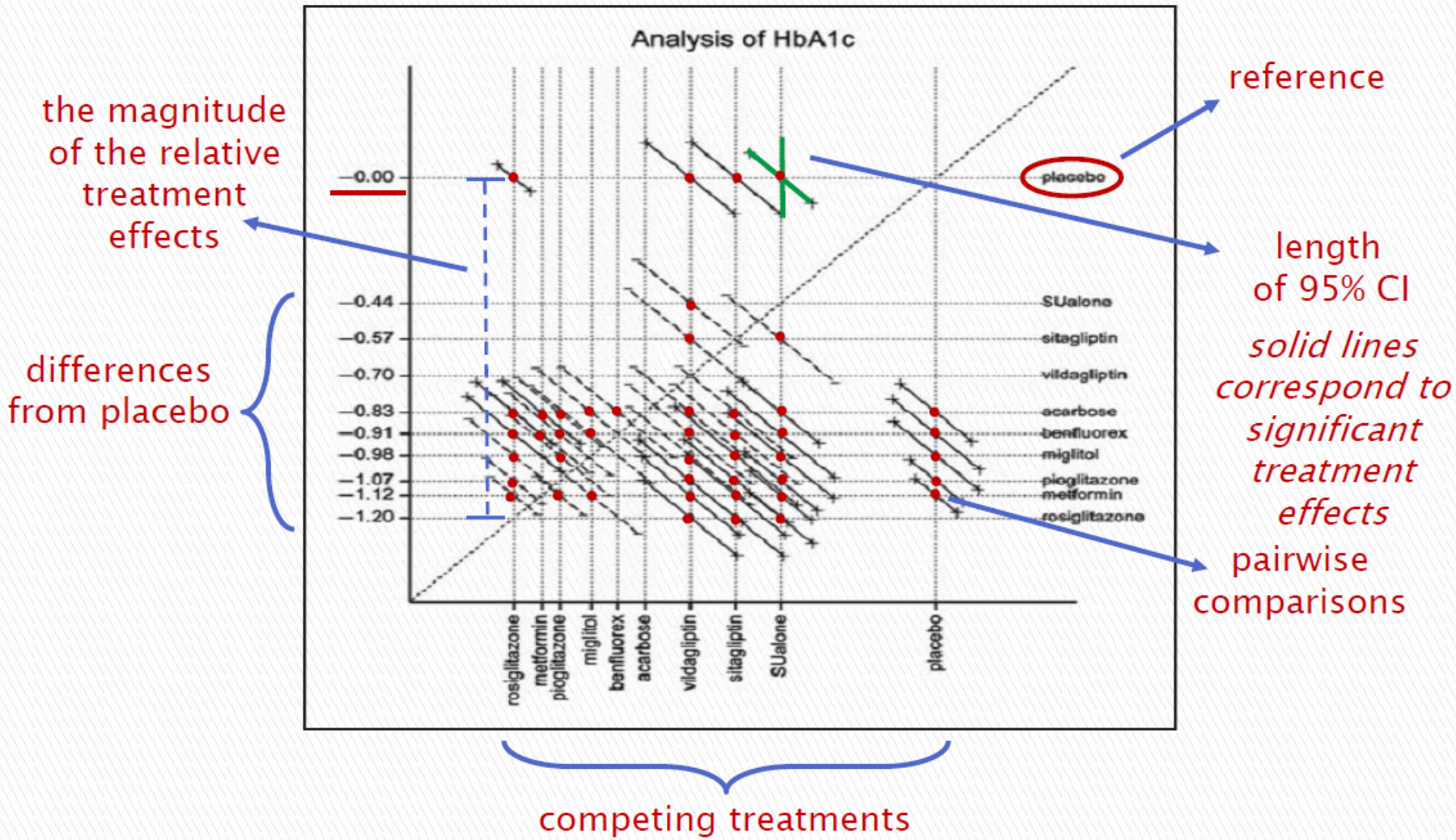
Table showing all the pairwise relative treatment effects with their 95% CI for one or two outcomes

[Example in Cipriani et al. 2011]



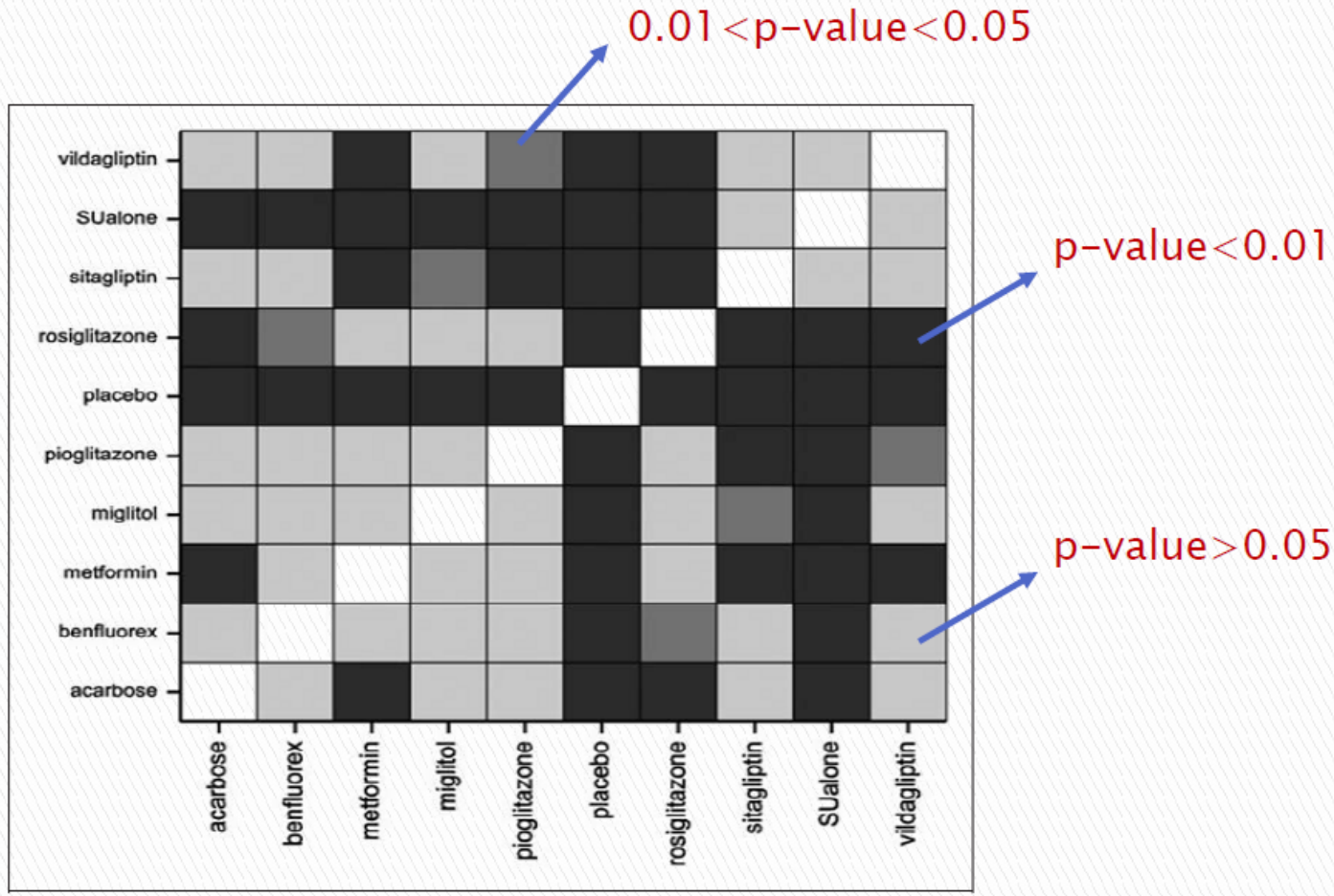
Forest plot with the treatment effects estimates for the pairwise comparisons

[Examples in Hawkins et al. 2009 & Hoaglin et al. 2011]



'Hsu mean-mean plot' showing the network estimates with the 95% CI for all pairwise comparisons

[Example in Senn et al. 2013]

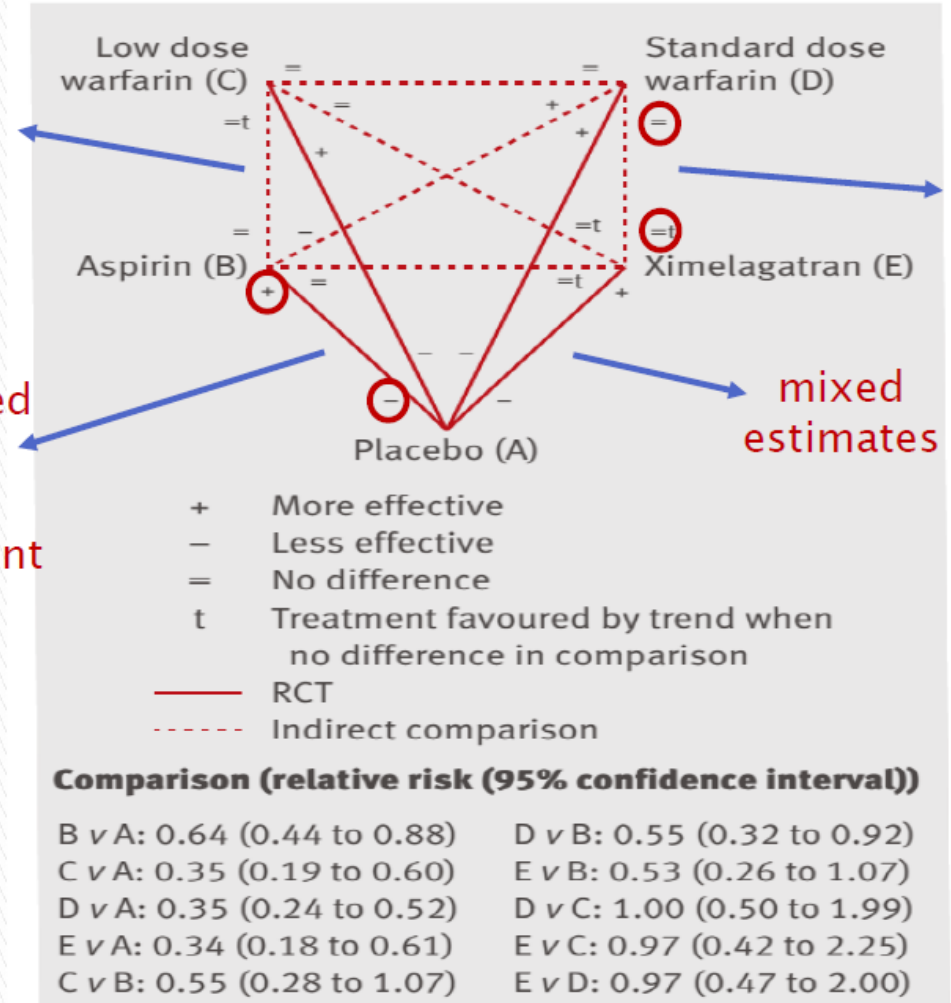


Shade plot showing the p-values of the treatment effects for all pairwise comparisons in the network

[Example in Senn et al. 2013]

indirect estimates

in comparisons with significant relative treatment effects (based on the network estimates):
'+' the favored treatment
'–' the non-favored treatment



mixed estimates

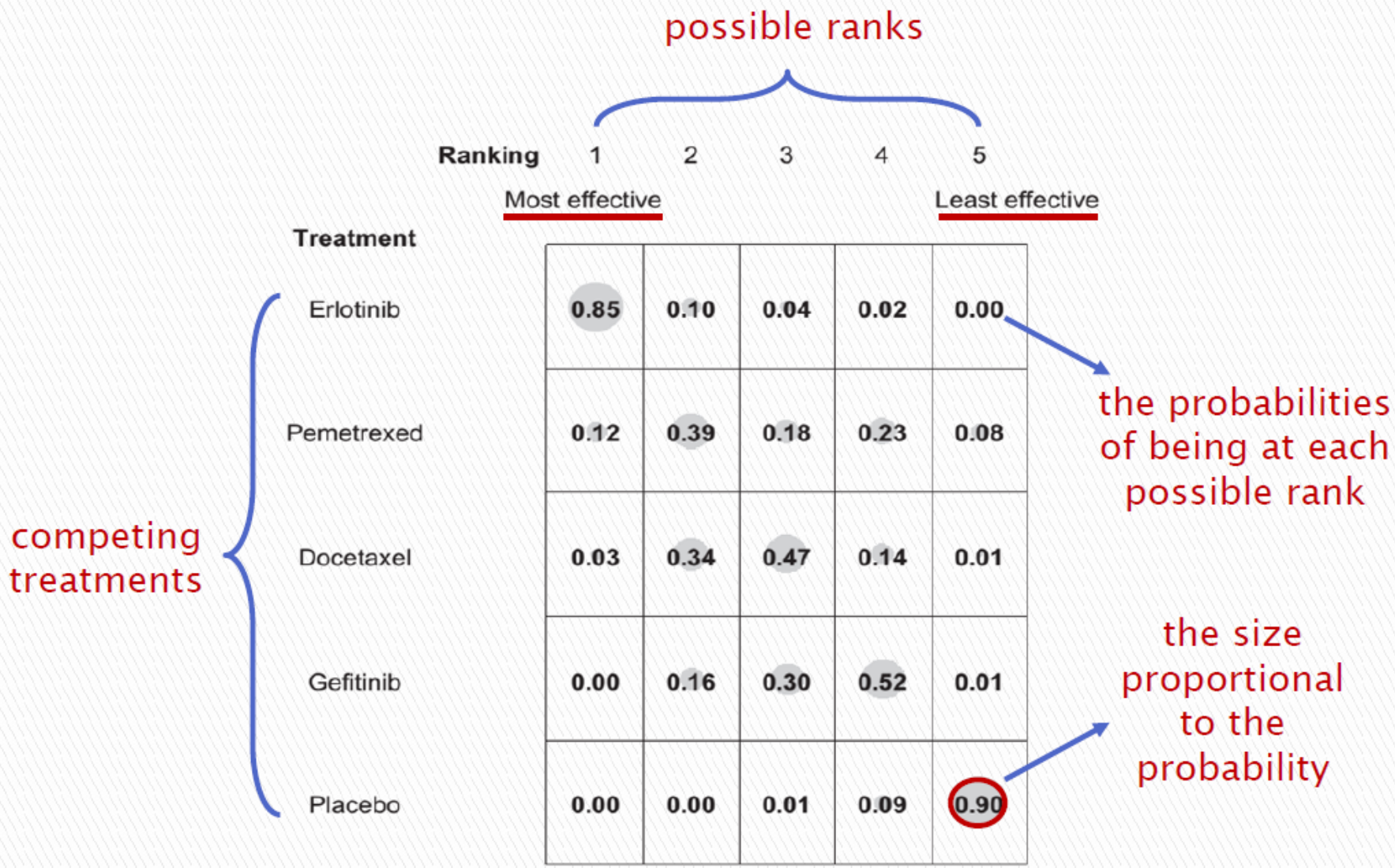
in comparisons with non-significant relative treatment effects (based on the network estimates):
'=t' the favored treatment
'=' the non-favored treatment

Network graph presenting the relative treatment effects for each pairwise comparison

[Example in Fadda et al. 2011]

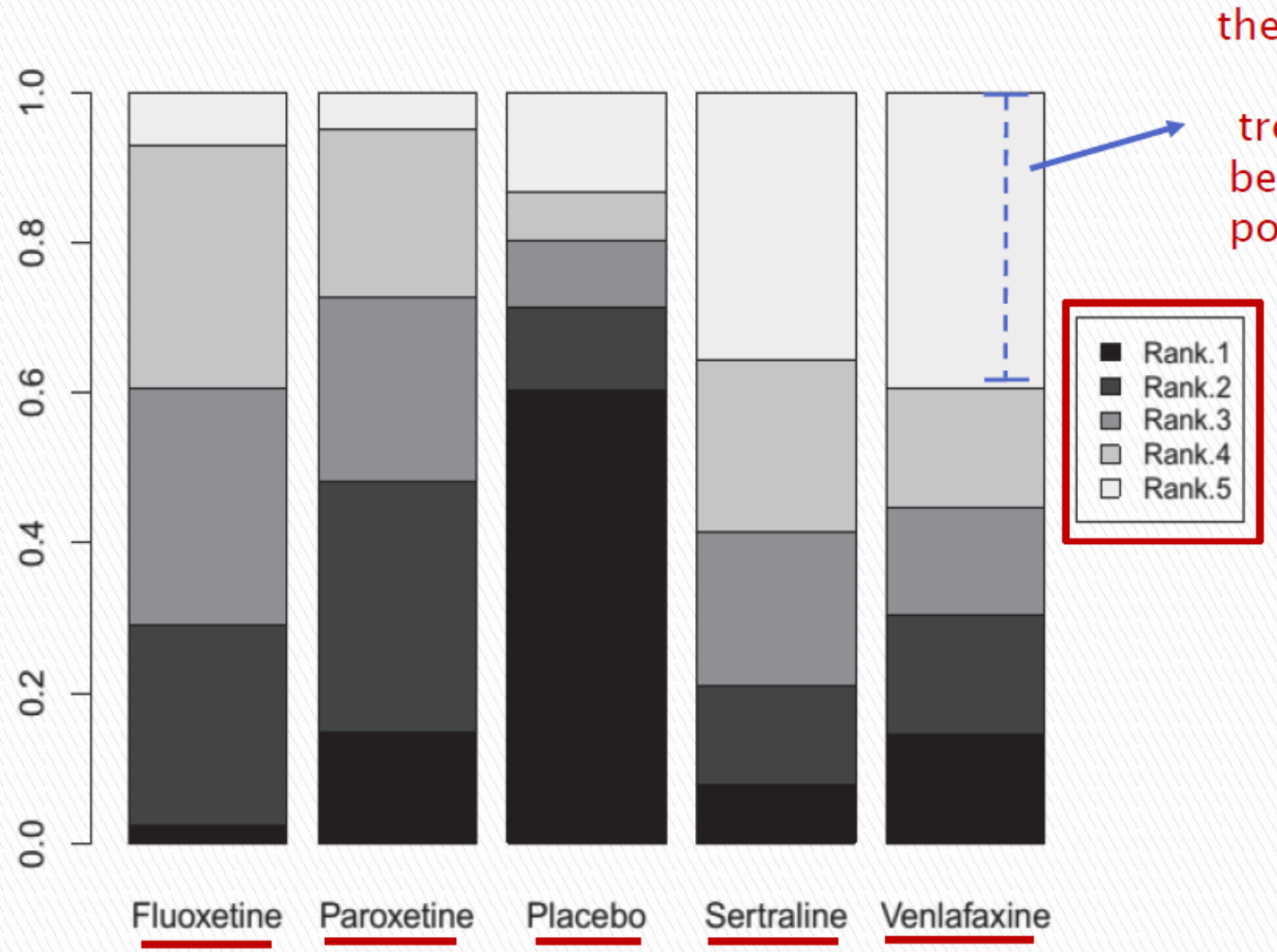
Presenting the results ranking

- Using probability of being the best
- Using probabilities of being at each possible rank
- Using SUCRAS



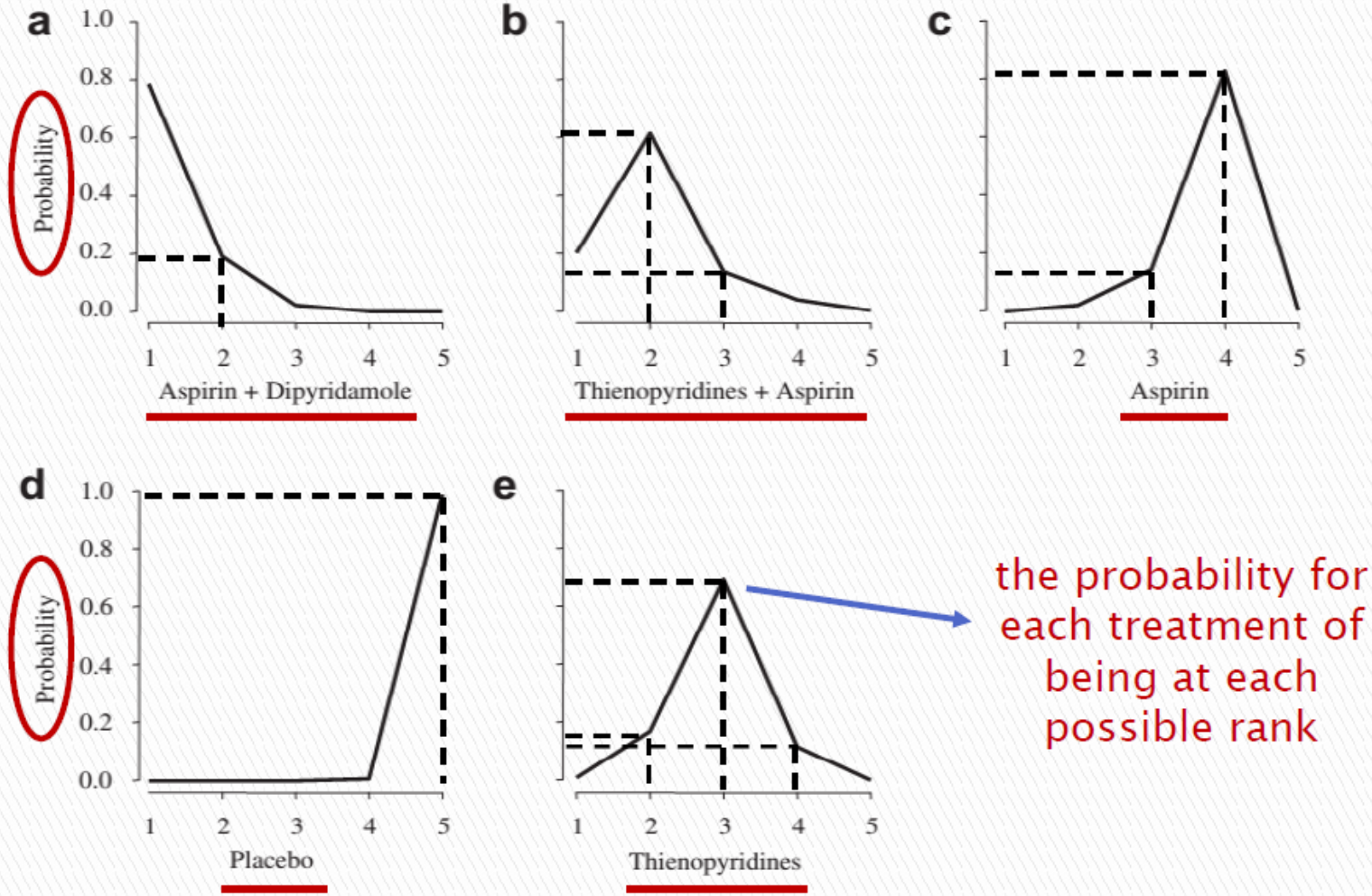
'Bubble-plot' including the ranking probabilities for all treatments

[Example in Hawkins et al. 2009]



Bar plots showing the probability for each treatment of being at a specific rank

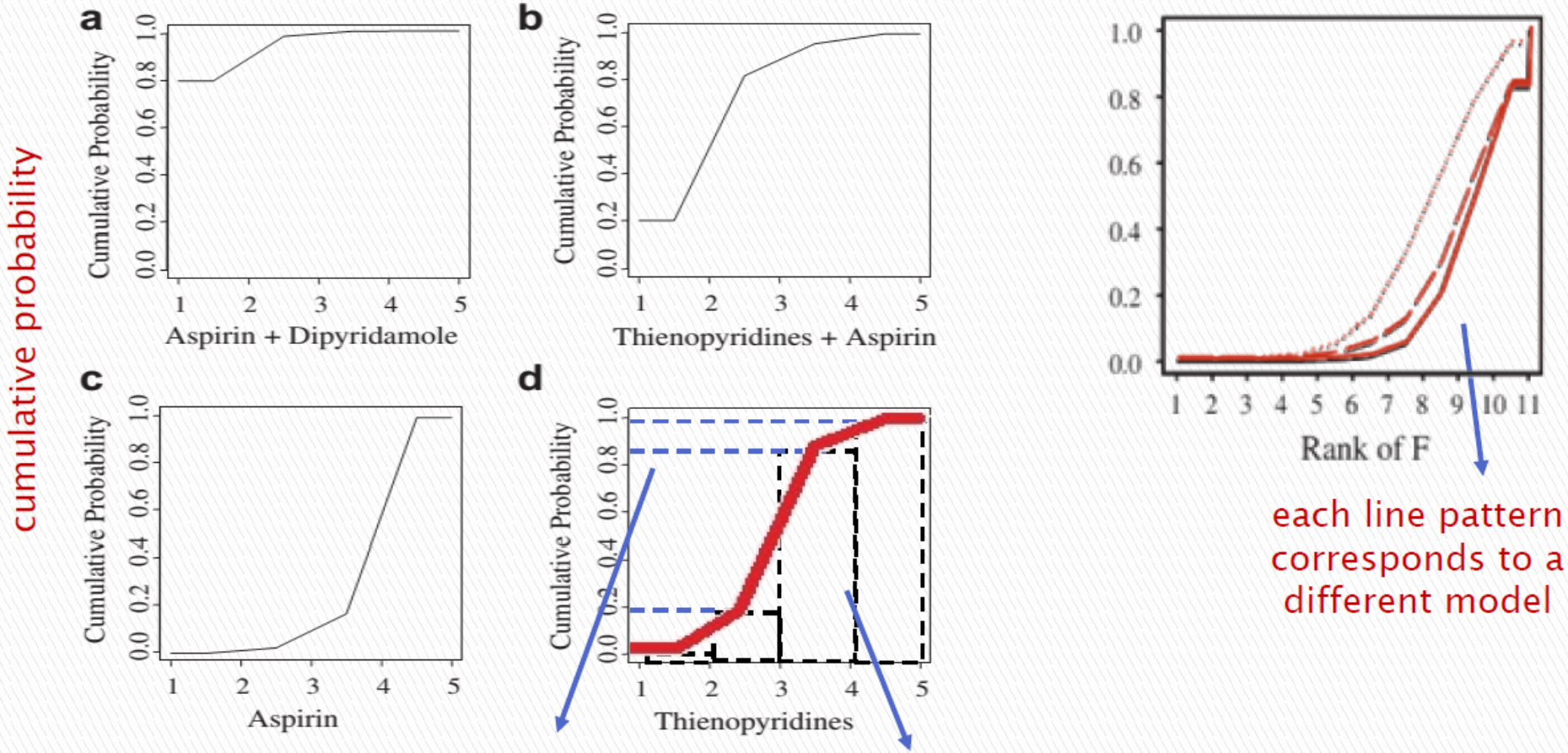
[Example in van Valkenhoef et al. 2012]



the probability for each treatment of being at each possible rank

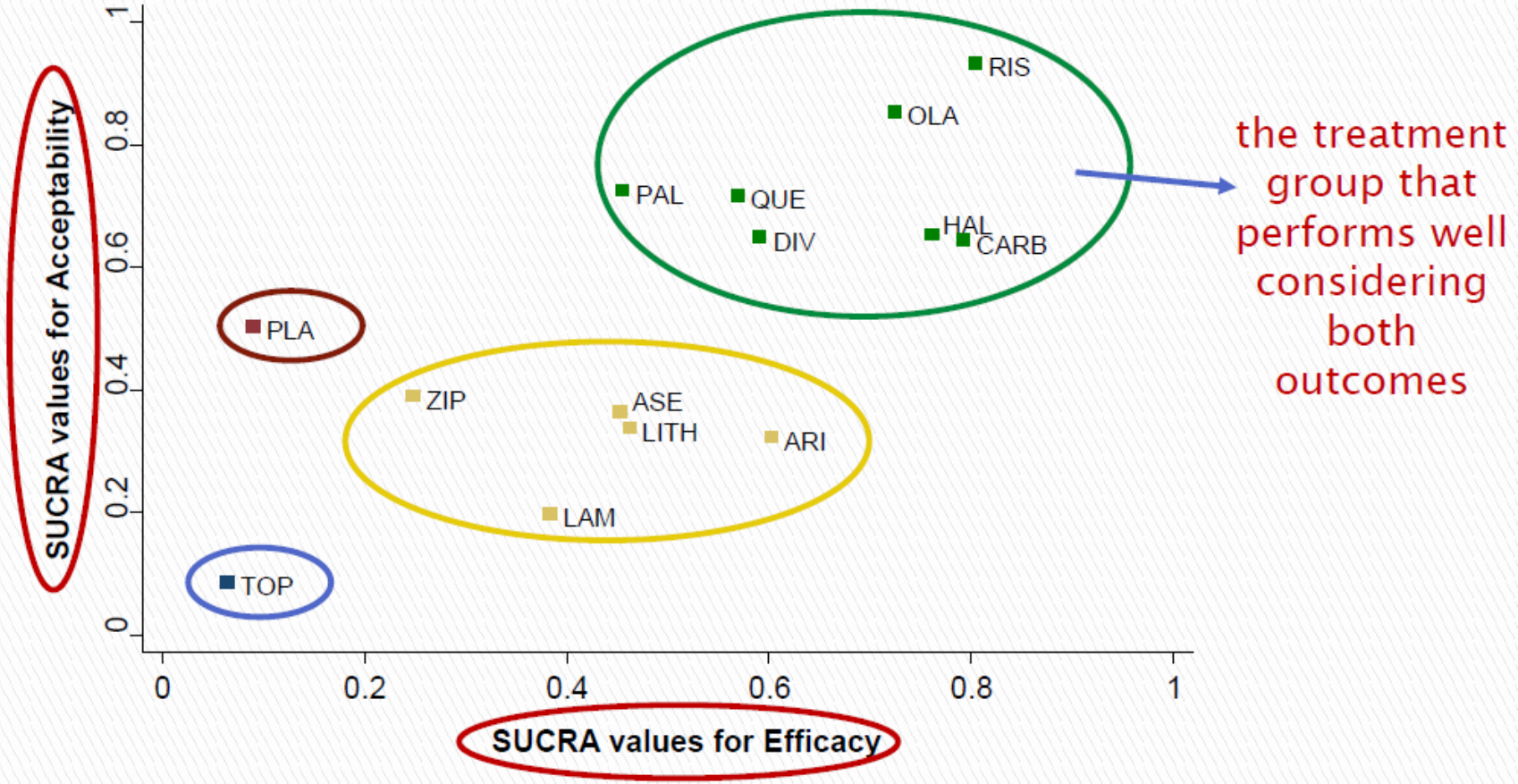
'Rankograms' showing the probability for each treatment of being at a specific rank

[Example in Salanti et al. 2011]



'SUCRA plots' showing the cumulative probability for each treatment of being up to a specific rank

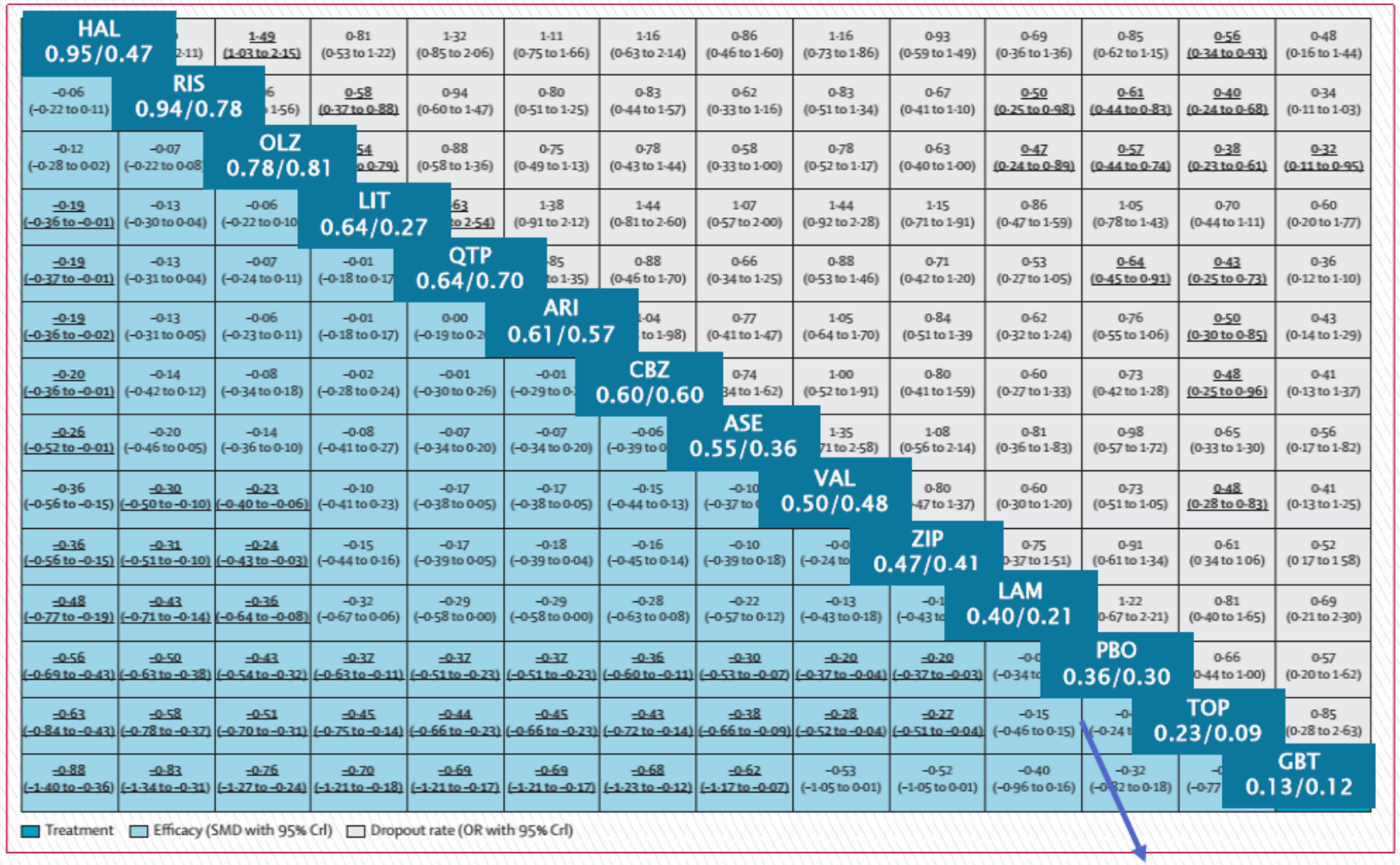
[Examples in Salanti et al. 2011 & Salanti et al. 2010]



the treatment group that performs well considering both outcomes

Scatterplot showing jointly the ranking results for two different outcomes

[Example in Chaimani et al. 2013]



competing treatments ordered according to their relative ranking for efficacy

Table showing all the pairwise relative treatment effects with their 95% CI for one or two outcomes along with the SUCRA values

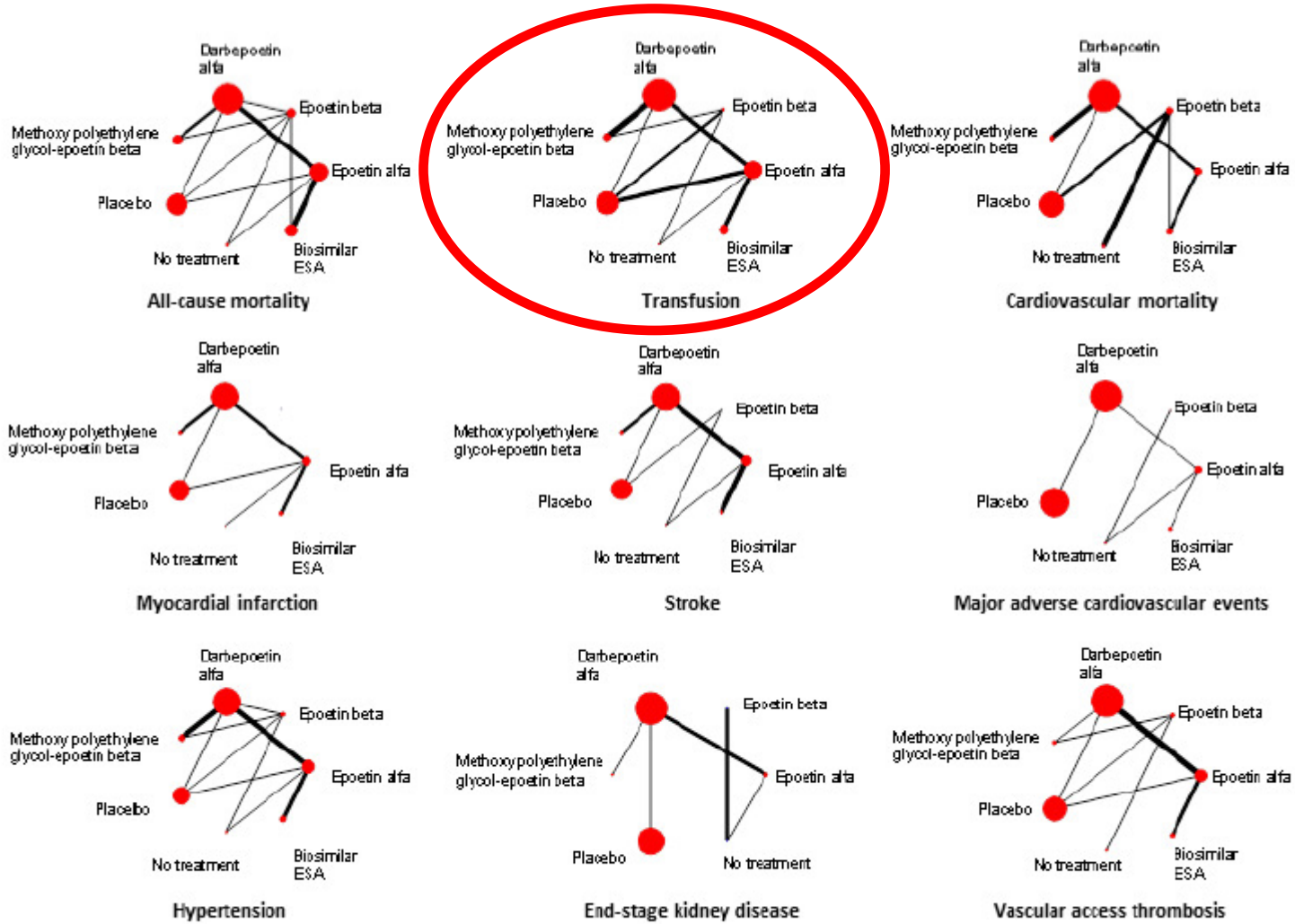
Erythropoiesis-stimulating agents for anaemia in adults with chronic kidney disease: a network meta-analysis (Review)

Palmer SC, Saglimbene V, Mavridis D, Salanti G, Craig JC, Tonelli M, Wiebe N, Strippoli GFM

OBJECTIVES

To compare the efficacy and safety of ESAs (epoetin alfa, epoetin beta, darbepoetin alfa, or methoxy polyethylene glycol-epoetin beta, and biosimilar ESAs, against each other, placebo, or no treatment) to treat anaemia in adults with CKD.

Figure 5. Networks of the treatment efficacy and safety of ESA drugs in the treatment of anaemia in chronic kidney disease. Values lower than 1 favour the active treatment in the comparison



Assessment of clinical and methodological

nerated

Analysis 1.1. Comparison 1 ESA versus ESA or placebo/no treatment, Outcome 1 Blood transfusion.

Review: Erythropoiesis-stimulating agents for anaemia in adults with chronic kidney disease: a network meta-analysis

Comparison: 1 ESA versus ESA or placebo/no treatment

Outcome: 1 Blood transfusion

Study or subgroup	Intervention n/N	Comparator n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds H,Ran
1 Epoetin alfa versus placebo					
Kleinman 1989	0/7	3/7		24.9 %	0.09 [0.00,
Canadian EPO Study 1990	2/67	23/32		36.4 %	0.01 [0.00,
Roth 1994	4/43	9/40		38.7 %	0.35 [0.10,
Subtotal (95% CI)	**117**	**79**		**100.0 %**	**0.07 [0.01, 0.**
Total events: 6 (Intervention), 35 (Comparator)					
Heterogeneity: Tau² = 3.61; Chi² = 10.53, df = 2 (P = 0.01); I² =81%					
Test for overall effect: Z = 2.10 (P = 0.036)					
2 Epoetin beta versus placebo					
Bennett 1991	0/90	1/41		10.4 %	0.15 [0.01,
Bahlmann 1991	5/53	28/46		89.6 %	0.07 [0.02,
Subtotal (95% CI)	**143**	**87**		**100.0 %**	**0.07 [0.03, 0.**
Total events: 5 (Intervention), 29 (Comparator)					
Heterogeneity: Tau² = 0.0; Chi² = 0.21, df = 1 (P = 0.64); I² =0.0%					
Test for overall effect: Z = 4.95 (P < 0.00001)					
3 Darbepoetin alfa versus placebo					
TREAT Study 2005	297/2012	496/2026		100.0 %	0.53 [0.46,

0.001 0.01 0.1 1 10 100 1000

Favours intervention Favours comparator

(Continue

(. . . Continued)

Study or subgroup	Intervention n/N	Comparator n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
ARCTOS Study 2008	11/162	4/162		23.7 %	2.88 [0.90, 9.23]
CORDATUS Study 2011	5/154	10/153		25.4 %	0.48 [0.16, 1.44]
PATRONUS Study 2010	32/244	39/245		46.2 %	0.80 [0.48, 1.32]
Subtotal (95% CI)	**585**	**606**		**100.0 %**	**0.94 [0.45, 1.95]**
Total events: 48 (Intervention), 54 (Comparator)					
Heterogeneity: Tau² = 0.23; Chi² = 5.35, df = 3 (P = 0.15); I² =44%					
Test for overall effect: Z = 0.18 (P = 0.86)					

0.001 0.01 0.1 1 10 100 1000

Favours intervention Favours comparator

(. . . Continued)

Study or subgroup	Intervention n/N	Comparator n/N	Odds Ratio M-H,Random,95% CI	Weight	Odds Ratio M-H,Random,95% CI
Subtotal (95% CI)	**2012**	**2026**		**100.0 %**	**0.53 [0.46, 0.63]**
Total events: 297 (Intervention), 496 (Comparator)					
Heterogeneity: not applicable					
Test for overall effect: Z = 7.71 (P < 0.00001)					
4 Epoetin alfa versus control					
Patel 2012	4/118	0/39		100.0 %	3.10 [0.16, 58.97]
Subtotal (95% CI)	**118**	**39**		**100.0 %**	**3.10 [0.16, 58.97]**
Total events: 4 (Intervention), 0 (Comparator)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.75 (P = 0.45)					
5 Epoetin beta versus no control					
Van Biesen 2005	2/22	4/18		100.0 %	0.35 [0.06, 2.18]
Subtotal (95% CI)	**22**	**18**		**100.0 %**	**0.35 [0.06, 2.18]**
Total events: 2 (Intervention), 4 (Comparator)					
Heterogeneity: not applicable					
Test for overall effect: Z = 1.12 (P = 0.26)					
6 Epoetin alfa versus darbepoetin alfa					
Akizawa 2011	0/160	1/161		2.8 %	0.33 [0.01, 8.24]
Locatelli 2001	3/37	6/129		14.2 %	1.81 [0.43, 7.61]
Nissenson 2002	37/335	17/369		82.9 %	2.57 [1.42, 4.66]
Subtotal (95% CI)	**532**	**659**		**100.0 %**	**2.31 [1.34, 3.97]**
Total events: 40 (Intervention), 24 (Comparator)					
Heterogeneity: Tau² = 0.0; Chi² = 1.64, df = 2 (P = 0.44); I² =0.0%					
Test for overall effect: Z = 3.03 (P = 0.0025)					
7 Epoetin alfa versus biosimilar ESA					
Krivoshiev 2010	1/230	3/232		5.5 %	0.33 [0.03, 3.23]
Krivoshiev 2008	9/304	10/305		33.6 %	0.90 [0.36, 2.25]
rtin 2007	11/192	46/560		60.9 %	0.68 [0.34, 1.34]
otal (95% CI)	726	1097		100.0 %	0.72 [0.42, 1.22]
ents: 21 (Intervention), 59 (Comparator)					
ogeneity: Tau² = 0.0; Chi² = 0.70, df = 2 (P = 0.70); I² =0.0%					
r overall effect: Z = 1.22 (P = 0.22)					
etin beta versus methoxy polyethylene glycol-epoetin beta					
ICUS Study 2007	2/46	7/135		100.0 %	0.83 [0.17, 4.15]
otal (95% CI)	46	135		100.0 %	0.83 [0.17, 4.15]
ents: 2 (Intervention), 7 (Comparator)					
ogeneity: not applicable					
r overall effect: Z = 0.23 (P = 0.82)					
epoetin alfa versus methoxy polyethylene glycol-epoetin beta					
OLI Study 2013	0/25	1/46		4.7 %	0.59 [0.02, 15.14]

0.001 0.01 0.1 1 10 100 1000

Favours intervention Favours comparator

Assessment of similarity (transitivity) across treatment comparisons

Evaluation of the assumption is important and its plausibility determines the validity of the network meta-analysis results.

We inferred about the assumption of transitivity:

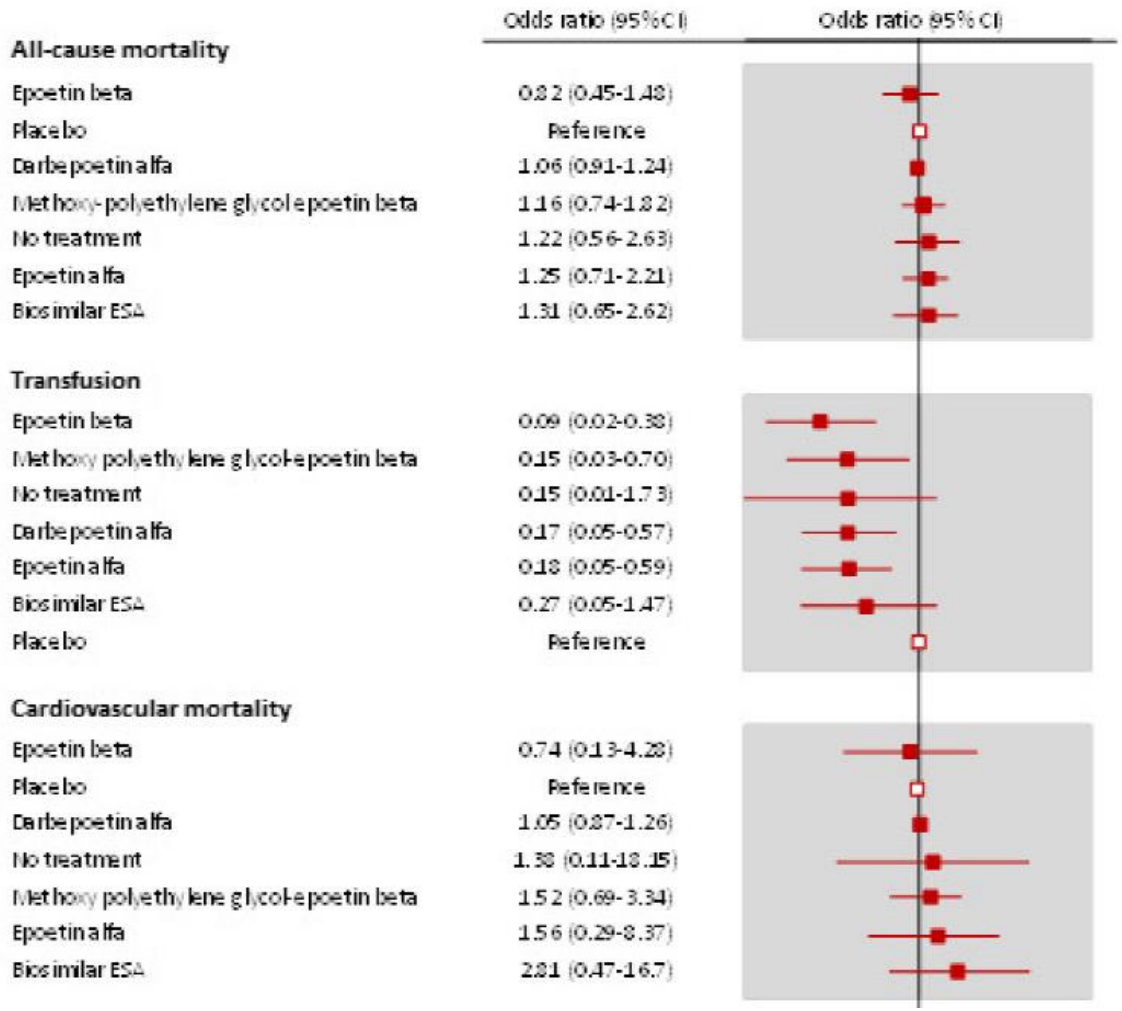
1. We assessed whether the included interventions were similar when they were evaluated in studies with different designs, for example, whether ESAs are administered the same way in studies comparing ESAs to placebo and in those comparing ESAs to other ESAs

2. We compared the distribution of the potential effect modifiers (age, stage of CKD, duration of treatment) across the different pairwise comparisons.

The inconsistency factor is the absolute difference in the log odds ratio estimated from indirect and direct treatment comparisons and is reported together with the 95% confidence interval. A 95% confidence interval that includes zero indicates that the result is compatible with zero inconsistency between effect estimates using indirect (networkmeta-analysis) and direct (conventional pairwise meta-analysis) treatment comparisons.

Transfusion		
Epoetin alfa - epoetin beta - placebo – no treatment	2.09	0.00-6.91
Epoetin alfa - darbepoetin alfa - placebo	1.97	0.00-4.20
Epoetin beta - darbepoetin alfa – methoxy polyethylene glycol-epoetin beta - placebo	1.26	0.00-3.39

Figure 6. Forest plots for results from network meta-analyses comparing ESAs versus placebo



Annals of Internal Medicine RESEARCH AND REPORTING METHODS

The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations

Brian Hutton, PhD, MSc; Georgia Salanti, PhD; Deborah M. Caldwell, PhD, MA, BA; Anna Chaimani, PhD; Christopher H. Schmid, PhD; Chris Cameron, MSc; John P.A. Ioannidis, MD, DSc; Sharon Straus, MD, MSc; Kristian Thorlund, PhD; Jeroen P. Jansen, PhD; Cynthia Mulrow, MD, MSc; Ferrán Catalá-López, PhD, MPH, PharmD; Peter C. Gøtzsche, MD, MSc; Kay Dickersin, PhD, MA; Isabelle Boutron, MD, PhD; Douglas G. Altman, DSc; and David Moher, PhD

The PRISMA statement is a reporting guideline designed to improve the completeness of reporting of systematic reviews and meta-analyses. Authors have used this guideline worldwide to prepare their reviews for publication. In the past, these reports typically compared 2 treatment alternatives. With the evolution of systematic reviews that compare multiple treatments, some of them only indirectly, authors face novel challenges for conducting and reporting their reviews. This extension of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) statement was developed specifically to improve the reporting of systematic reviews incorporating network meta-analyses.

A group of experts participated in a systematic review, Delphi survey, and face-to-face discussion and consensus meeting to establish new checklist items for this extension statement. Current PRISMA items were also clarified. A modified, 32-item PRISMA extension checklist was developed to address what the group considered to be immediately relevant to the reporting of network meta-analyses.

This document presents the extension and provides examples of good reporting, as well as elaborations regarding the rationale for new checklist items and the modification of previously existing items from the PRISMA statement. It also highlights educational information related to key considerations in the practice of network meta-analysis. The target audience includes authors and readers of network meta-analyses, as well as journal editors and peer reviewers.

Ann Intern Med. 2015;162:777-784. doi:10.7326/M14-2385 **www.annals.org**
For author affiliations, see end of text.

Table. Checklist of Items to Include When Reporting a Systen

Section/Topic	Item # *	Checklist Item†
TITLE		
Title	1	Identify the report as a systemati *meta-analysis).*
ABSTRACT		
Structured summary	2	Provide a structured summary in Background: main objectives Methods: data sources; study and *synthesis methods, such* Results: number of studies anc confidence/credible interval *to summarize pairwise comp brevity.* Discussion/Conclusions: limita Other: primary source of fundi
INTRODUCTION		
Rationale	3	Describe the rationale for the re *why a network meta-analysis*
Objectives	4	Provide an explicit statement of interventions, comparisons,
METHODS		
Protocol and registration	5	Indicate whether a review protoc and, if available, provide reg
Eligibility criteria	6	Specify study characteristics (e.g years considered, language, *Clearly describe eligible trea have been clustered or merg*
Information sources	7	Describe all information sources authors to identify additiona
Search	8	Present full electronic search str it could be repeated.
Study selection	9	State the process for selecting st and, if applicable, included
Data collection process	10	Describe method of data extract and any processes for obtai
Data items	11	List and define all variables for w assumptions and simplificati
Geometry of the network	S1	Describe methods used to explo potential biases related to it summarized for presentatio the evidence base to reader
Risk of bias within individual studies	12	Describe methods used for asse whether this was done at the in any data synthesis.
Summary measures	13	State the principal summary mea *of additional summary meas cumulative ranking curve (SU summary findings from meta*
Planned methods of analysis	14	Describe the methods of handlin meta-analysis. This should in *Handling of multigroup trials; Selection of variance structure; Selection of prior distributions Assessment of model fit.*
Assessment of inconsistency	S2	Describe the statistical methods the treatment network(s) stu
Risk of bias across studies	15	Specify any assessment of risk of bias, selective reporting with
Additional analyses	16	Describe methods of additional include, but not be limited t Sensitivity or subgroup analys Meta-regression analyses; *Alternative formulations of the Use of alternative prior distributions for Bayesian analyses (if applicable).*

RESEARCH AND REPORTING METHODS

PRISMA Extension for Network Meta-analysis

Table–Continued

Section/Topic	Item # *	Checklist Item†	Reported on Page #
RESULTS‡			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Presentation of network structure	S3	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	
Summary of network geometry	S4	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g., placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	
Exploration for inconsistency	S5	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses, *alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses,* and so forth).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, researchers, and policymakers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network geometry (e.g., avoidance of certain comparisons).*	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	

(Continued on following page)

Rilevanza e affidabilità dei confronti indiretti

(G.L. Pappagallo & M. Cinquini)

B.4 Identificazione dei medicinali comparatori

Nella presente sezione si richiede di indicare, se esistenti, le alternative terapeutiche utilizzate nel contesto assistenziale italiano per la popolazione *target,* su esiti riconosciuti come clinicamente rilevanti e validati per la patologia in oggetto, a partire dallo Standard of Care (SoC) raccomandato, nel momento di presentazione del Dossier, da linee guida nazionali, con particolare riferimento a quelle pubblicate nel Sistema Nazionale delle Linee guida. In caso di assenza di Linee guida nazionali, si suggerisce di far riferimento alle Linee guida europee e internazionali aggiornate, indicando eventuali differenze rispetto ai comparatori utilizzati nella pratica clinica nazionale.

Le alternative terapeutiche utilizzate nella pratica clinica rappresentano il/i comparatore/i ("Comparatore/i") con cui il Prodotto va confrontato ai fini della presente negoziazione: nello specifico, si richiede di individuare il/i Comparatore/i tenuto conto di indicazioni terapeutiche, medesima popolazione *target* ed eventuali sottopopolazioni e profili di efficacia, tollerabilità e sicurezza, anche alla luce dei principi sin qui adottati dall'AIFA in materia di valutazione dell'equivalenza terapeutica[5] e/o di sovrapponibilità terapeutica.

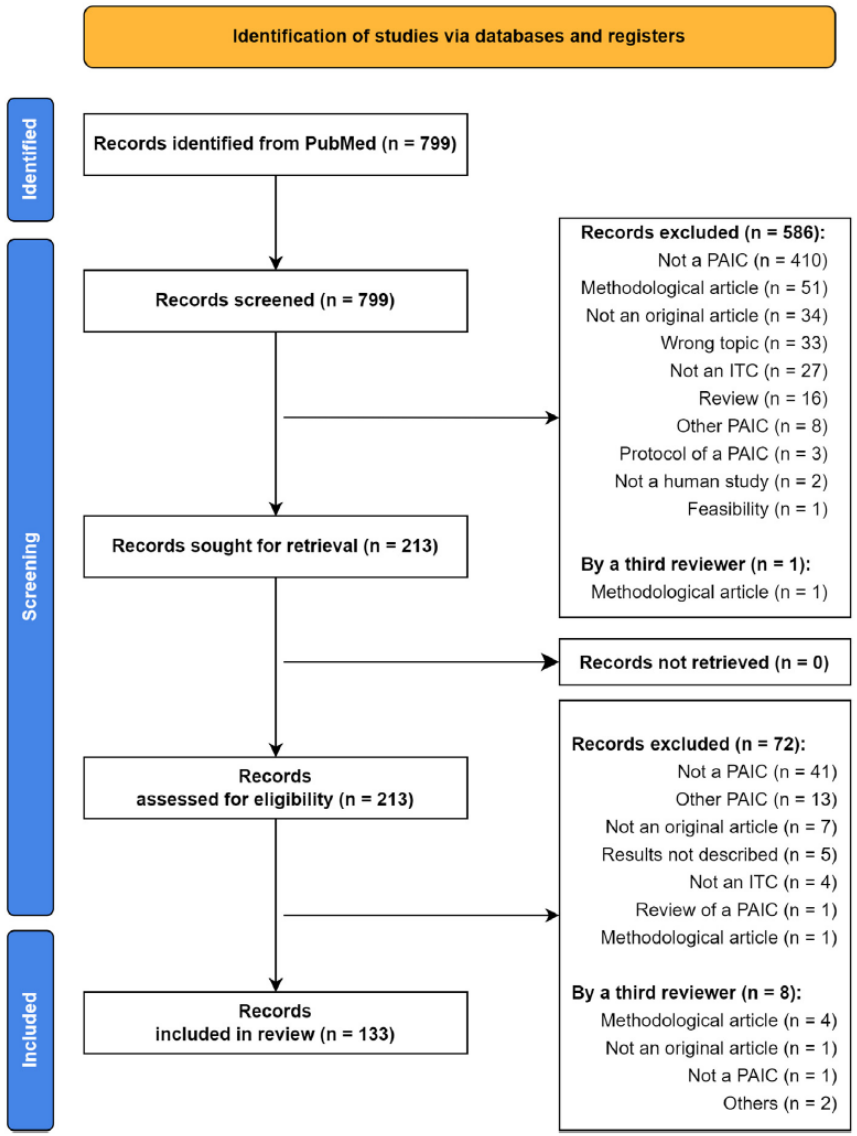
AGENZIA ITALIANA DEL FARMACO

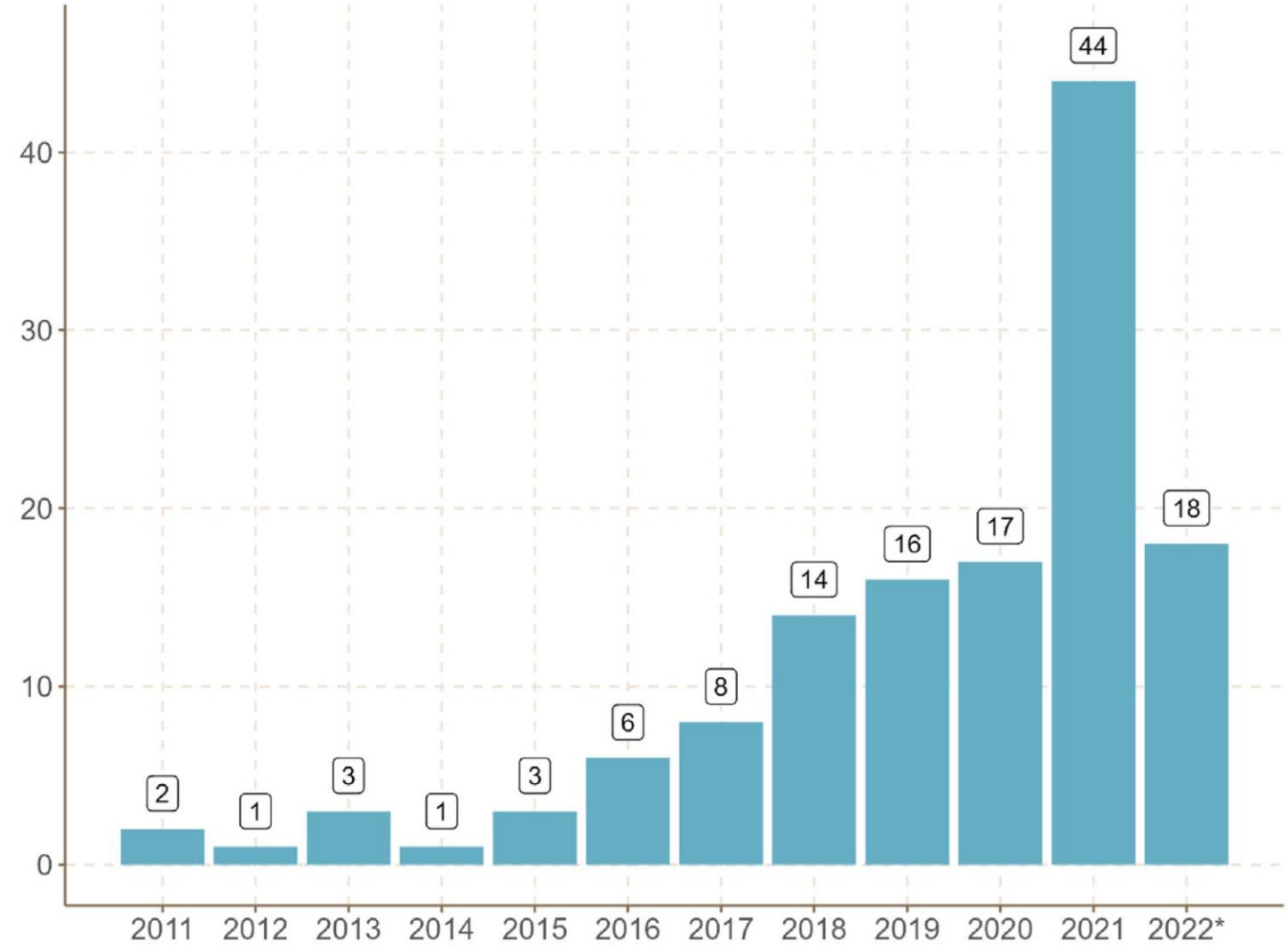
CRITERI DI VALUTAZIONE PER L'ATTRIBUZIONE DELL'INNOVATIVITÀ TERAPEUTICA E SULLA GESTIONE DEGLI AGENTI ANTINFETTIVI PER INFEZIONI DA GERMI MULTIRESISTENTI

Tenuto conto dell'implementazione del Regolamento (UE) 2021/2282, nella valutazione GRADE, il disegno dello studio (randomizzato/osservazionale) viene valutato considerando il livello randomizzato per i confronti indiretti ancorati o le metodologie di confronto valutate come rigorose dalla CSE, solo nel caso in cui non sia possibile sviluppare un disegno sperimentale con confronto diretto. Di contro, le metanalisi a rete o le metodologie di confronto indiretto non ancorato sono considerate sul medesimo livello di uno studio osservazionale.

Overview of NMA and MAIC posters at ASCO-GU

1. Network Meta-analysis (NMA) to Assess Comparative Efficacy of Lenvatinib plus Pembrolizumab Compared with other First-line Treatments for Management of Advanced Renal Cell Carcinoma (aRCC)
   - Grünwald *et al.* 2024 (Sponsored by Eisai Inc.)
2. Pembrolizumab plus lenvatinib versus alternate therapies in first line (1L) for advanced renal cell carcinoma (aRCC): a network meta-analysis (NMA)
   - Yan *et al.* 2024 (Sponsored by Merck & Co., Inc.)
3. Pembrolizumab plus Lenvatinib vs. alternative therapies in first-line (1L) advanced renal cell carcinoma (aRCC) by IMDC risk factor: a network meta-analysis (NMA)
   - Rane *et al.* 2024 (Sponsored by Merck & Co., Inc.)
4. Pembrolizumab plus Lenvatinib vs. nivolumab plus cabozantinib in patients with metastatic renal cell carcinoma: A matching-adjusted indirect comparison (MAIC)
   - Rane *et al.* 2024 (Sponsored by Merck & Co., Inc.)
5. Health-related quality of life (HRQoL) of first-line treatments in metastatic renal cell carcinoma (mRCC): A network meta-analysis
   - Abidoye *et al.* 2024

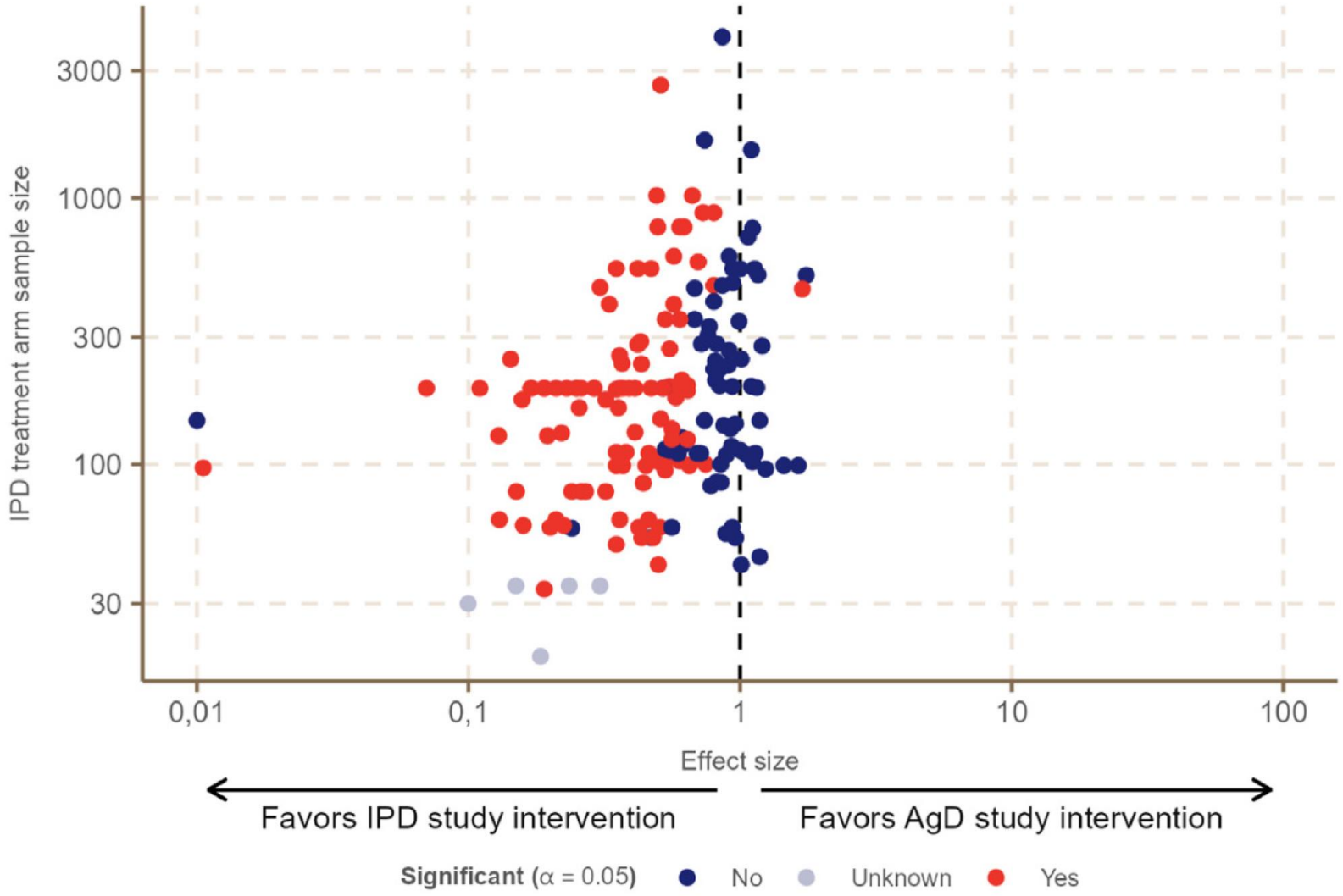




* For the year 2022, data collected up to April 2 only

The treatment effect was beneficial for the IPD treatment arm for all but one comparison (161/162, 99.6%).

The only PAIC in favor of the aggregated data treatment arm was one of the three articles without any involvement of the pharmaceutical industry.



Journal of Clinical Epidemiology 163 (2023) 1−10

A methodological review of population-adjusted indirect comparisons reveals inconsistent reporting and suggests publication bias

Arnaud Serret-Larmande[a,b,*], Belkacem Zenati[a], Agnès Dechartres[a], Jérôme Lambert[b], David Hajage[a]

Methodology and Reporting

Conclusion

The methodology and reporting of these studies were heterogeneous and overall insufficient regarding key criteria such as primary outcome definition and covariates selection for adjustment.

Published results suggest a major publication and reporting bias.

Table 1 Advantages/disadvantages of MAIC and NMA

Matching-adjusted indirect comparison (MAIC)	Network meta-analysis (NMA)
Advantages	
• Reduces heterogeneity between trials by matching the patient population • Treatment effects have clear clinical context for interpretation • Possible with and without placebo adjustment • Long-term analyses feasible	• Compares multiple treatments using published aggregate data • Can connect head-to-head RCTs and other RCTs via a common comparator (usually placebo) • Multiple simultaneous indirect paths • Based on relative effects, so randomisation is preserved • Established methodology
Disadvantages	
• Evolving method—NICE Technical Support Document published in December 2016 [2] • Interferes with/breaks randomisation • Reduced patient sample size • Only a single indirect path • Can only match observed characteristics, so heterogeneity may remain	• Assumes trials are comparable in terms of design and population (low heterogeneity) • Requires a common comparator (connected evidence network) • Often only short-term comparison due to lack of a long-term connected network (placebo switching)

Adapted from Ishak et al.
MAIC matching-adjusted indirect comparison, *NMA* network meta-analysis, *RCT* randomised controlled trials

Choy *et al. Arthritis Research & Therapy* (2019) 21:32
https://doi.org/10.1186/s13075-019-1812-3

Table 1 Advantages/disadvantages of MAIC and NMA

Matching-adjusted indirect comparison (MAIC)	Network meta-analysis (NMA)
Advantages	
• Reduces heterogeneity between trials by matching the patient population • Treatment effects have clear clinical context for interpretation • Possible with and without placebo adjustment • Long-term analyses feasible	• Compares multiple treatments using published aggregate data • Can connect head-to-head RCTs and other RCTs via a common comparator (usually placebo) • Multiple simultaneous indirect paths • Based on relative effects, so randomisation is preserved • Established methodology
Disadvantages	
• Evolving method—NICE Technical Support Document published in December 2016 [2] • Interferes with/breaks randomisation • Reduced patient sample size • Only a single indirect path • Can only match observed characteristics, so heterogeneity may remain	• Assumes trials are comparable in terms of design and population (low heterogeneity) • Requires a common comparator (connected evidence network) • Often only short-term comparison due to lack of a long-term connected network (placebo switching)

Adapted from Ishak et al.
MAIC matching-adjusted indirect comparison, *NMA* network meta-analysis, *RCT* randomised controlled trials

Choy *et al. Arthritis Research & Therapy* (2019) 21:32
https://doi.org/10.1186/s13075-019-1812-3


CAUTION