Editorial considerations for reviews that compare multiple interventions

Said Business School, Oxford, UK

March 22, 2013

Handout E3-L

Basic ideas of indirect comparisons and network meta-analysis

Deborah Caldwell

School of Social and Community Medicine

University of Bristol

Acknowledgements

- Georgia Salanti
- Julian Higgins
- Tianjing Li
- Nicky Welton
- Sofia Dias
- Tony Ades

Multiple treatment decision-making

- For many clinical indications there will often be several possible interventions.
- The Cochrane Database of Systematic Reviews
  - 22 interventions for adult smoking cessation
  - >12 interventions for chronic asthma in adults
  - 10 treatments for childhood nocturnal enuresis
  - 14 pharmacological treatments inducing labour
- Health care decisions should be based on 'best available' evidence from systematic reviews & meta-analysis of RCTs

Problem…

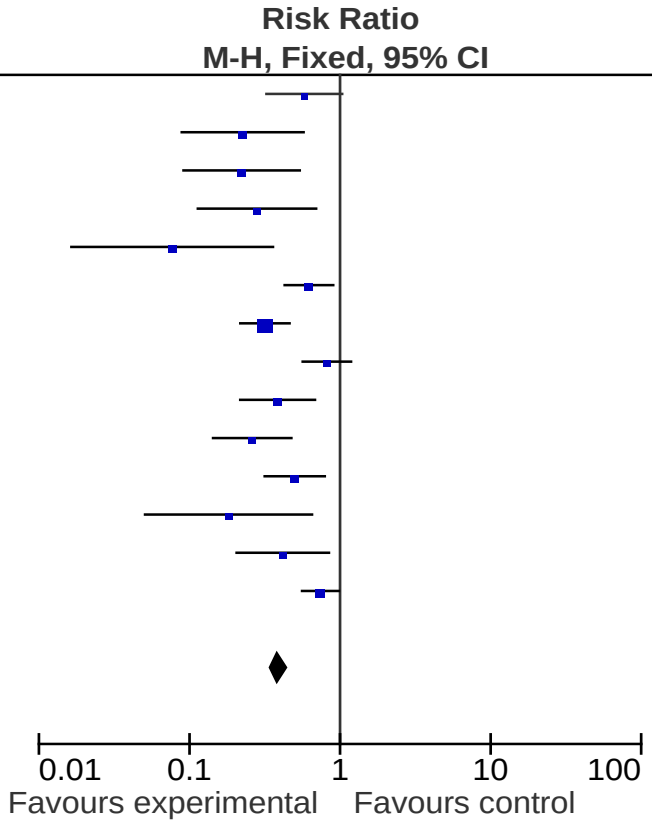
Systematic reviews typically focus on direct, head-to-head comparisons of interventions.

Study or Subgroup	Weight	Risk Ratio M-H, Fixed, 95% CI
Bennett 1985	3.7%	0.58 [0.32, 1.03]
Bollard 1981a	6.0%	0.23 [0.09, 0.57]
Bollard 1981b	7.0%	0.22 [0.09, 0.54]
Houts 1986	4.6%	0.28 [0.11, 0.69]
Jehu 1977	7.7%	0.08 [0.02, 0.36]
Lynch 1984	7.2%	0.62 [0.43, 0.90]
Moffat 1987	22.0%	0.32 [0.22, 0.46]
Nawaz 2002	4.3%	0.82 [0.57, 1.18]
Ronen 1992	7.3%	0.39 [0.22, 0.68]
Sacks 1974	4.8%	0.26 [0.14, 0.47]
Sloop 1973	7.7%	0.50 [0.32, 0.79]
Wagner 1982	4.3%	0.18 [0.05, 0.65]
Wagner 1985	4.6%	0.42 [0.21, 0.84]
Werry 1965	8.8%	0.74 [0.56, 0.98]
Total (95% CI)	**100.0%**	**0.39 [0.33, 0.45]**

Total events

Heterogeneity: $Chi^2 = 56.57$, $df = 13$ ($P < 0.00001$); $I^2 = 77\%$

Test for overall effect: $Z = 12.04$ ($P < 0.00001$)



Problem... (2)

Consequently, the evidence base consists of a set of pair-wise comparisons of interventions

- Placebo comparisons of limited use to the practitioner or policy-maker who wants to know the 'best' treatment to recommend/ prescribe.

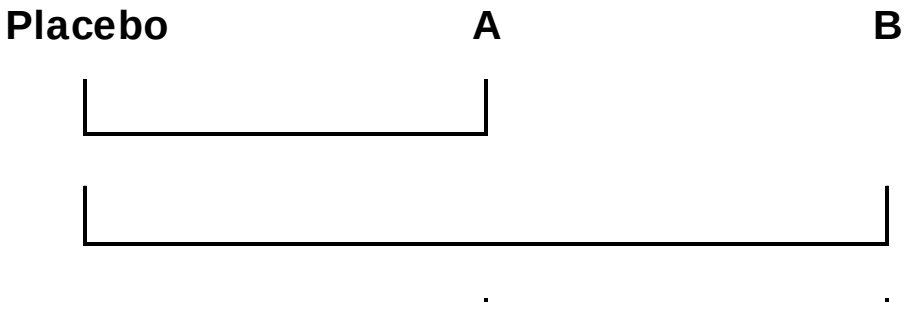
'Best available' evidence is not always available or sufficient

- Placebo controlled trials sufficient for regulatory approval of new drugs
- Even when active comparisons have been made such direct evidence is often limited.

Therefore, evidence base may not contain treatment comparisons of relevance for clinician or policy maker.

Example evidence structure

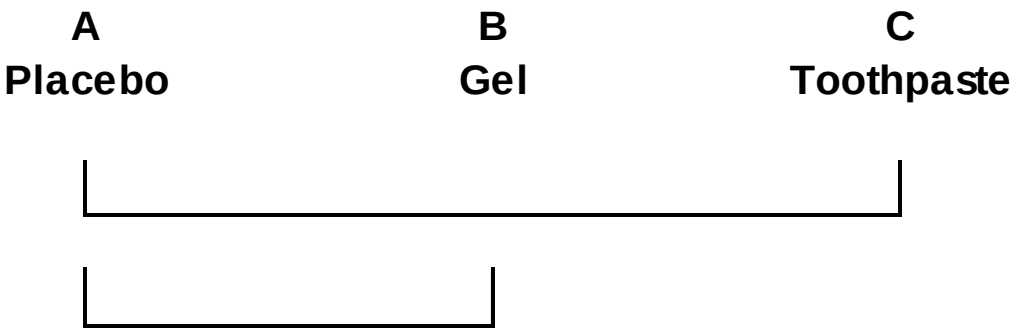
Common situation is to have multiple competing treatments (often within class) each studied in placebo-controlled RCTs but none compared directly to each other.



How do we know which treatment to use?

Case study: fluoride to prevent dental caries

Evidence base: 3 treatment options; 2 comparisons

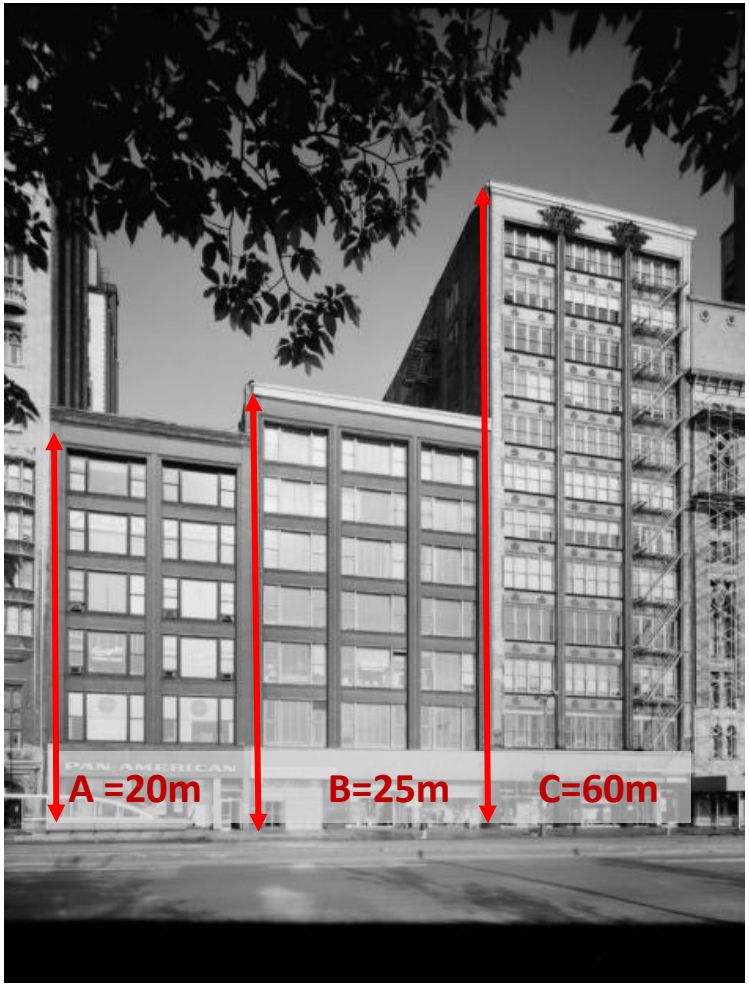


Summary of results: from 2 separate meta-analyses

Comparison	SMD	95% CI
Toothpaste vs placebo	-0.34	(-0.41, -0.28)
Gel vs placebo	-0.19	(-0.30, -0.10)

Indirect comparison

- If we know how much taller is B to A and how much taller is C to A we know how much taller is B compared to C



How much taller is building C compared to building B?

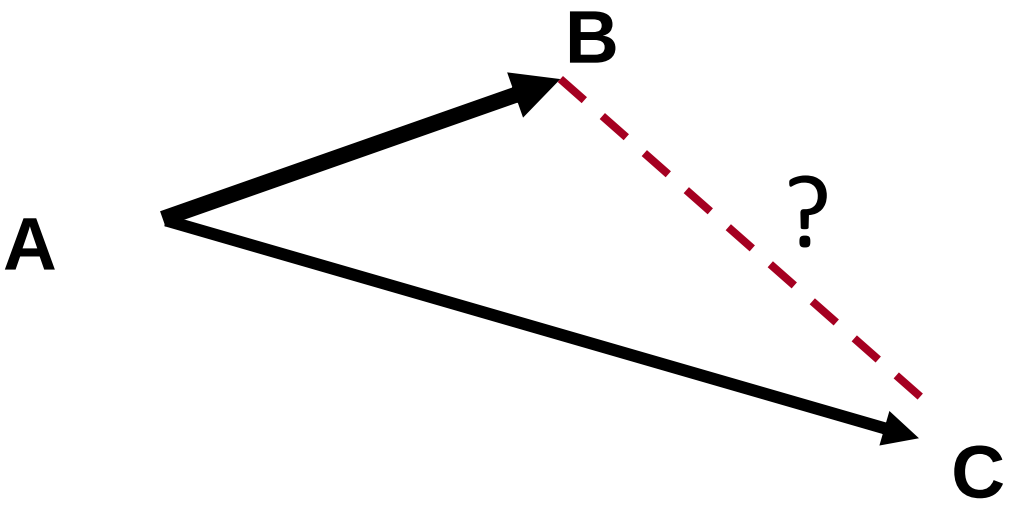
AB difference: B minus A = 5m

AC difference: C minus A = 40m

BC difference = 40m—5m = 35m

Indirect comparison

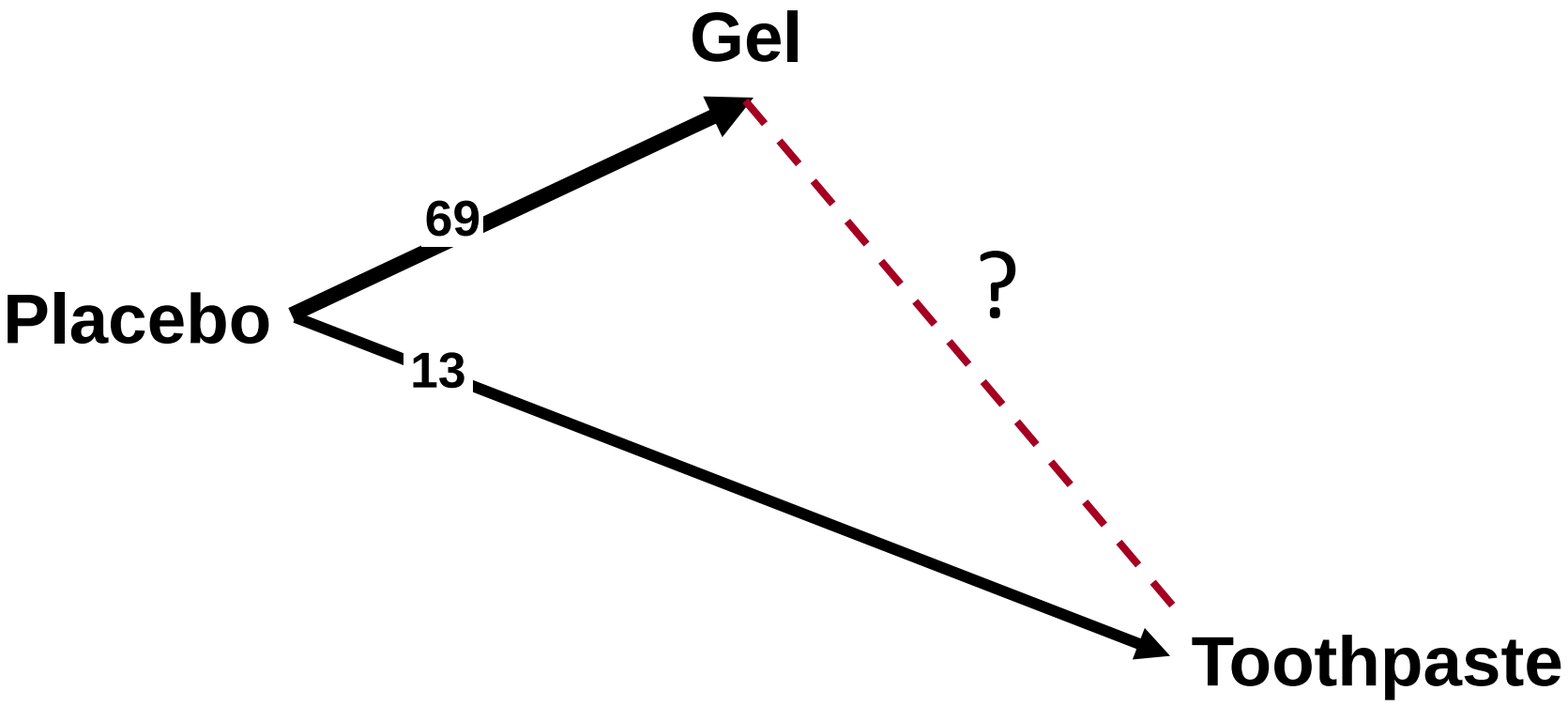
- We can obtain an indirect estimate for B vs C from RCTs comparing A vs B and A vs C:



$$SMD_{BC} = SMD_{AC} - SMD_{AB}$$

$$LRR_{BC} = LRR_{AC} - LRR_{AB}$$

Worked example: Toothpaste versus Gel



Comparison	SMD	CIs
Placebo vs Toothpaste	-0.34	(-0.41, -0.28)
Placebo vs Gel	-0.19	(-0.30, -0.10)

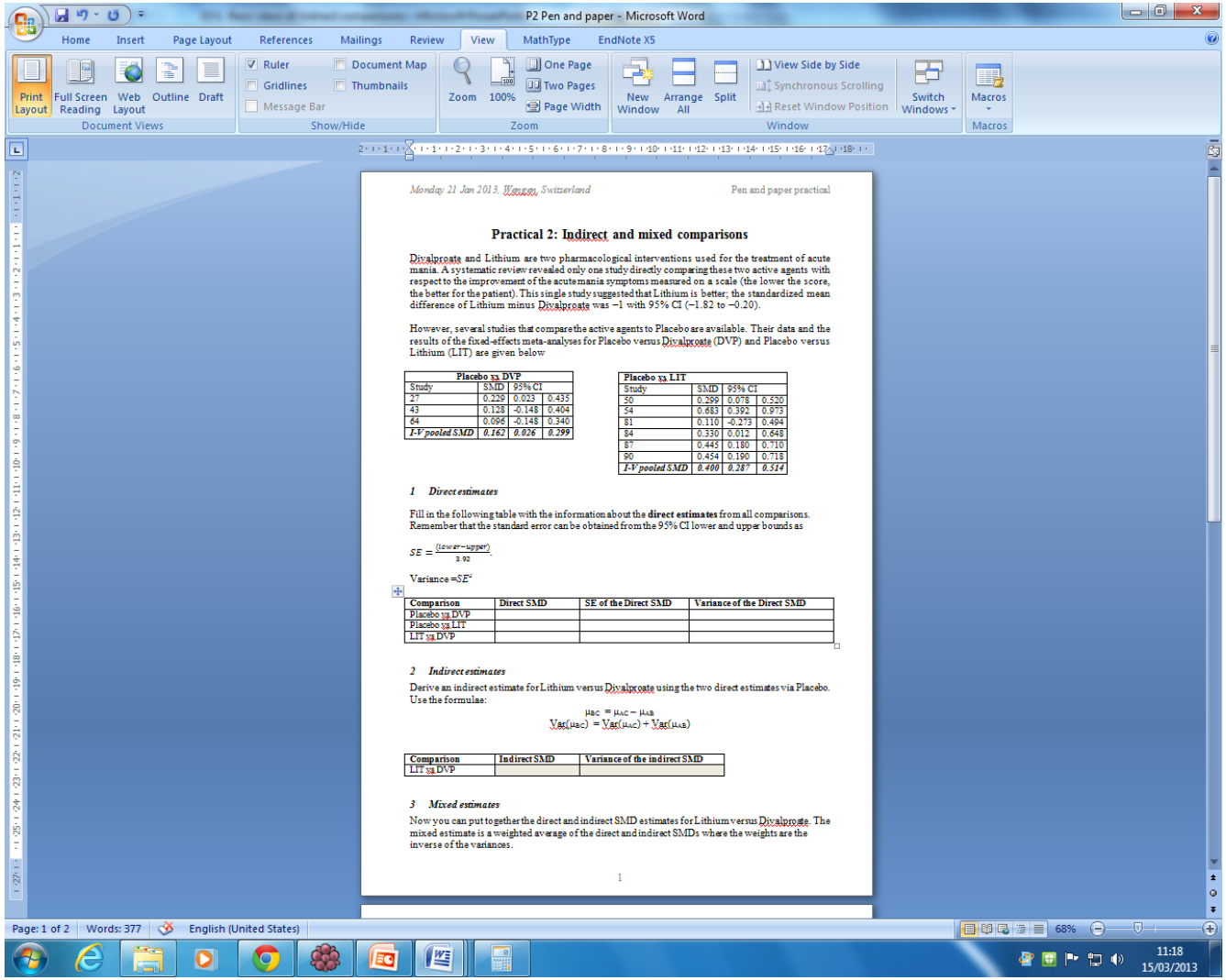
Example: Toothpaste versus Gel

- **Indirect SMD_{GvsT}** = $SMD_{PvsT} - SMD_{PvsG}$
- **Indirect SMD_{GvsT}** = –0.34 – (–0.19)= –0.15

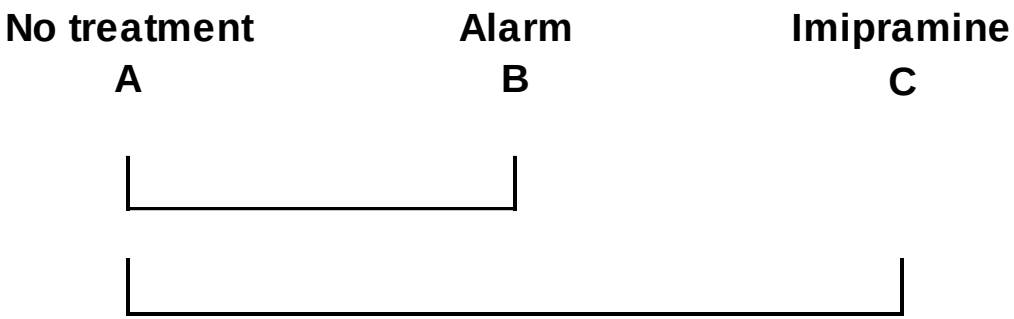
- **Variance Indirect SMD_{GvsT} = Variance SMD_{PvsT} + Variance SMD_{PvsG}**
- **Variance Indirect SMD_{GvsT}** = 0.0011 + 0.0026 = 0.0037

- **SE Indirect SMD_{GvsT}** = sqrt(0.0037) = 0.061
- **95% CI for Indirect SMD_{GvsT}** = (–0.15 – 1.96×0.061, -0.15 + 1.96×0.061)
- **95% CI for Indirect SMD_{GvsT}** = (–0.27, –0.03)

Pen and paper exercise



Exercise 2: treatments for nocturnal enuresis



Comparison	RR	CIs
No treatment vs Imipramine	0.95	(0.87 to 0.99)
No treatment vs Alarm	0.39	(0.33 to 0.46)

Outcome: failure to achieve 14 days dry nights

Group pen and paper exercise: Imipramine vs Alarm.

$$LRR_{BC} = LRR_{AC} - LRR_{AB}$$

lrr_{AB} =

lrr_{AC} =

$lrr_{BC} = lrr_{AC} - lrr_{AB}$ =

Indirect $RR_{BC} = \exp(lrr_{BC})$ =

Group pen and paper exercise: Imipramine vs Alarm.

$$LRR_{BC} = LRR_{AC} - LRR_{AB}$$

lrr_{AB} = -0.05

lrr_{AC} = -0.94

$lrr_{BC} = lrr_{AC} - lrr_{AB}$ =

Indirect $RR_{BC} = \exp(lrr_{BC})$ =

Pen and paper exercise: Imipramine vs Alarm.

$$LRR_{BC} = LRR_{AC} - LRR_{AB}$$

$lrr_{AB} = -0.05$

$lrr_{AC} = -0.94$

$lrr_{BC} = lrr_{AC} - lrr_{AB} = -0.94 - (-0.05) = -0.89$

Indirect $RR_{BC} = \exp(lrr_{BC}) =$ **0.41**

What NOT to do.

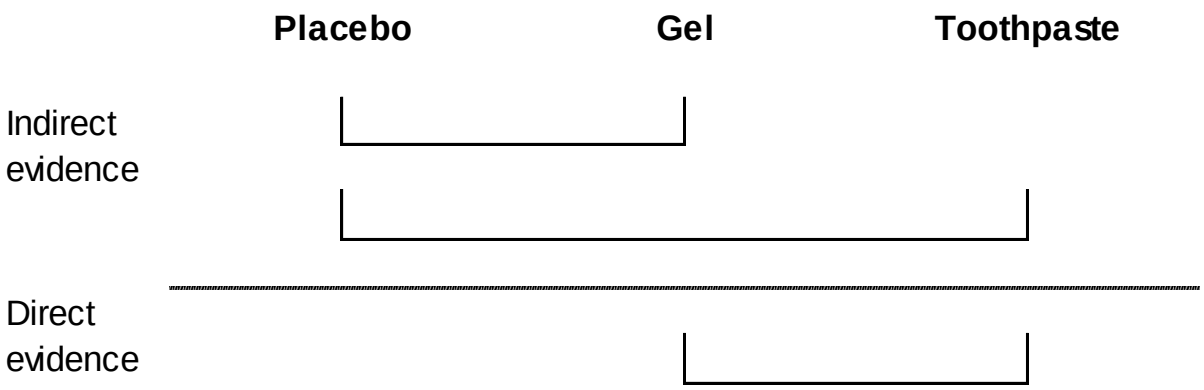
PLEASE DO NOT do a meta-analysis on all the A arms, and another on all the B arms, and another on all the C arms.

This breaks the randomised comparisons and Glenny (2005) calls this “unadjusted”

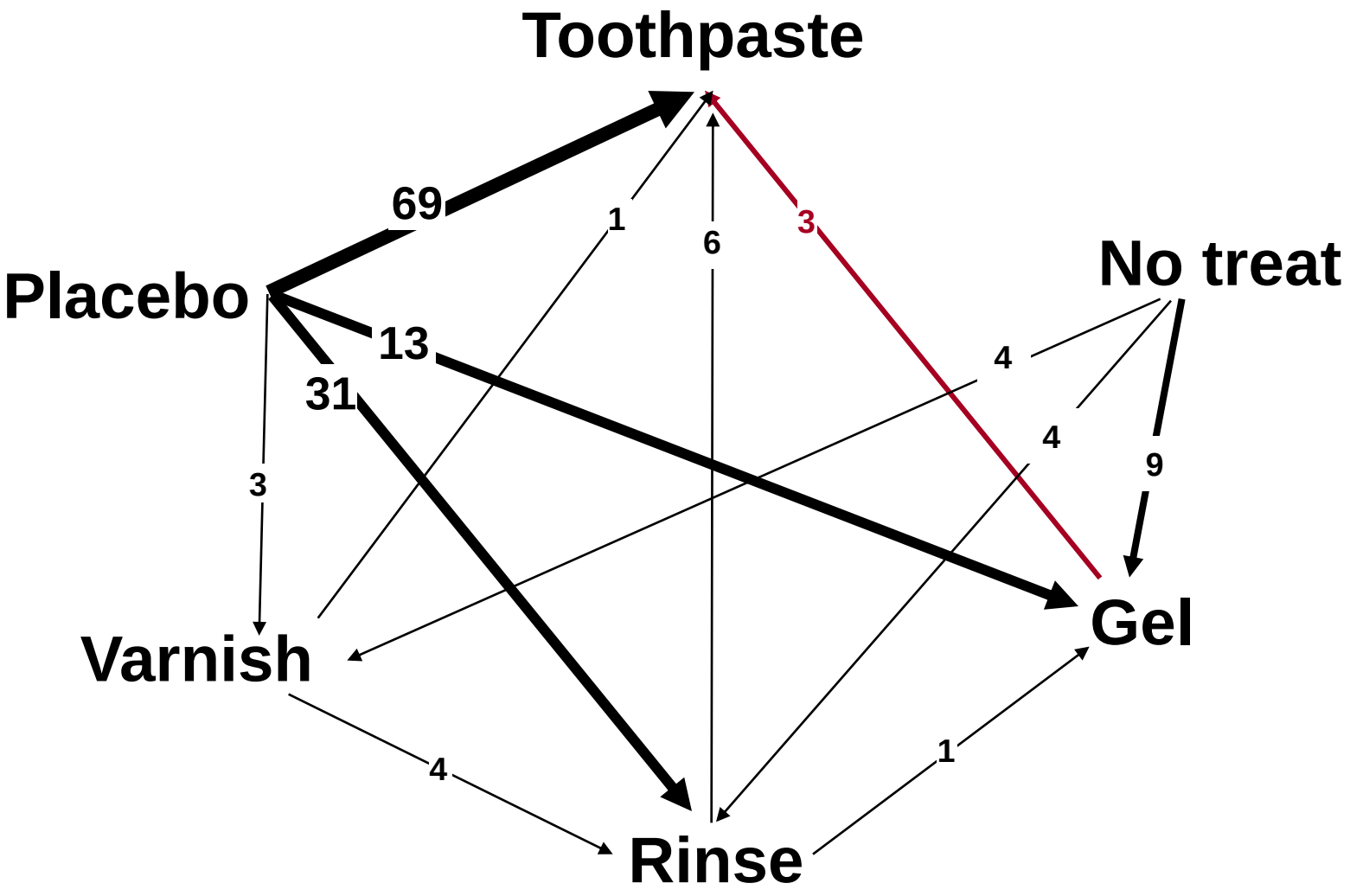
A correct analysis must be based on the relative treatment effects in each RCT

Example evidence structure #2

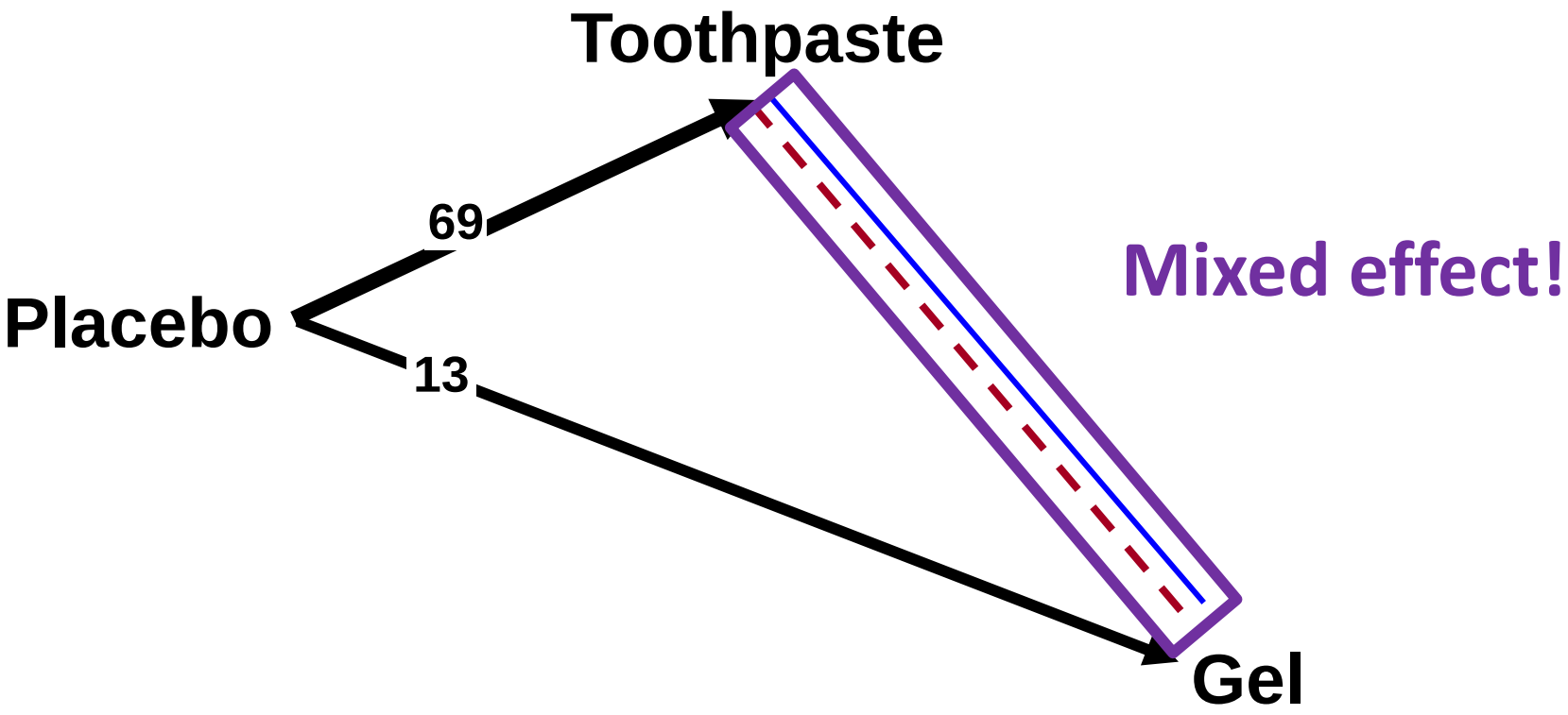
- Another common evidence structure is where we have **some** direct evidence on the relevant treatment comparisons (active vs active) but on its own its insufficient.



Example: Toothpaste versus Gel



Example: Toothpaste versus Gel



Mixed evidence: Combining direct and indirect evidence

Inverse variance approach to pooling direct and indirect evidence on SMD_{BC} (Toothpaste vs Gel)

1. SMD_{BC}^{direct} 2. $SMD_{BC}^{indirect}$

3. $$SMD_{BC}^{Mixed} = \frac{(w^{direct}\ SMD_{BC}^{direct}) + (w^{indirect}\ SMD_{BC}^{indirect})}{(w^{direct} + w^{indirect})}$$

Using the inverse variance method each estimate is 'weighted' by the inverse of the variance $w = 1/se(BC_i)^2$

Indirect evidence given less weight than direct evidence

Example: Toothpaste versus Gel

Indirect SMD_{GvsT} = - 0.15

Variance Indirect SMD_{GvsT} = 0.0037

Direct SMD_{GvsT} = 0.04

Variance Direct SMD_{GvsT} = 0.011

Mixed SMD_{GvsT} ?
Variance of Mixed SMD_{GvsT} ?
95% CI ?

Example: Toothpaste versus Gel

Indirect $SMD_{GvsT} = -0.15$

Variance Indirect $SMD_{GvsT} = 0.0037$

Direct $SMD_{GvsT} = 0.04$

Variance Direct $SMD_{GvsT} = 0.011$

Mixed SMD_{GvsT} ?
Variance of Mixed SMD_{GvsT} ?
95% CI ?

Mixed $SMD_{GvsT} = -0.102$

Var(Mixed SMD_{GvsT}) = 0.0028

95%CI: (-0.205, 0.001)

Mixed estimate: more precise!

Indirect SMD_{GvsT} = - 0.15

Variance Indirect SMD_{GvsT} = 0.0037

Direct SMD_{GvsT} = 0.04

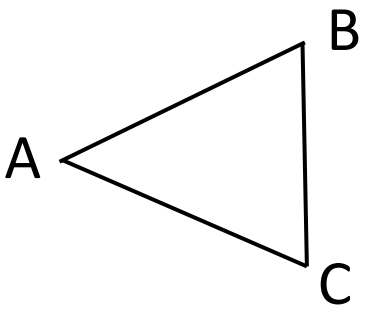
Variance Direct SMD_{GvsT} = 0.011

Mixed SMD_{GvsT} = -0.102

Var(Direct SMD_{GvsT}) = 0.0028

- Mixed estimates are more precise than the direct or the indirect estimate as they use both sources of information
- This might not be the case if
  - Direct and indirect estimates disagree (inconsistency)
  - If there is a lot of heterogeneity in the studies involved in the indirect evidence

Importance of “loops” of evidence



- Loops of evidence: e.g. AB, AC, BC

(1) Combines the “Indirect” and “direct” evidence

(2) Also, we can assess “inconsistency” between direct and indirect evidence (where inconsistency is defined as the discrepancy/ disagreement between the direct and indirect estimate of treatment effect).

Limitations of mixed approach

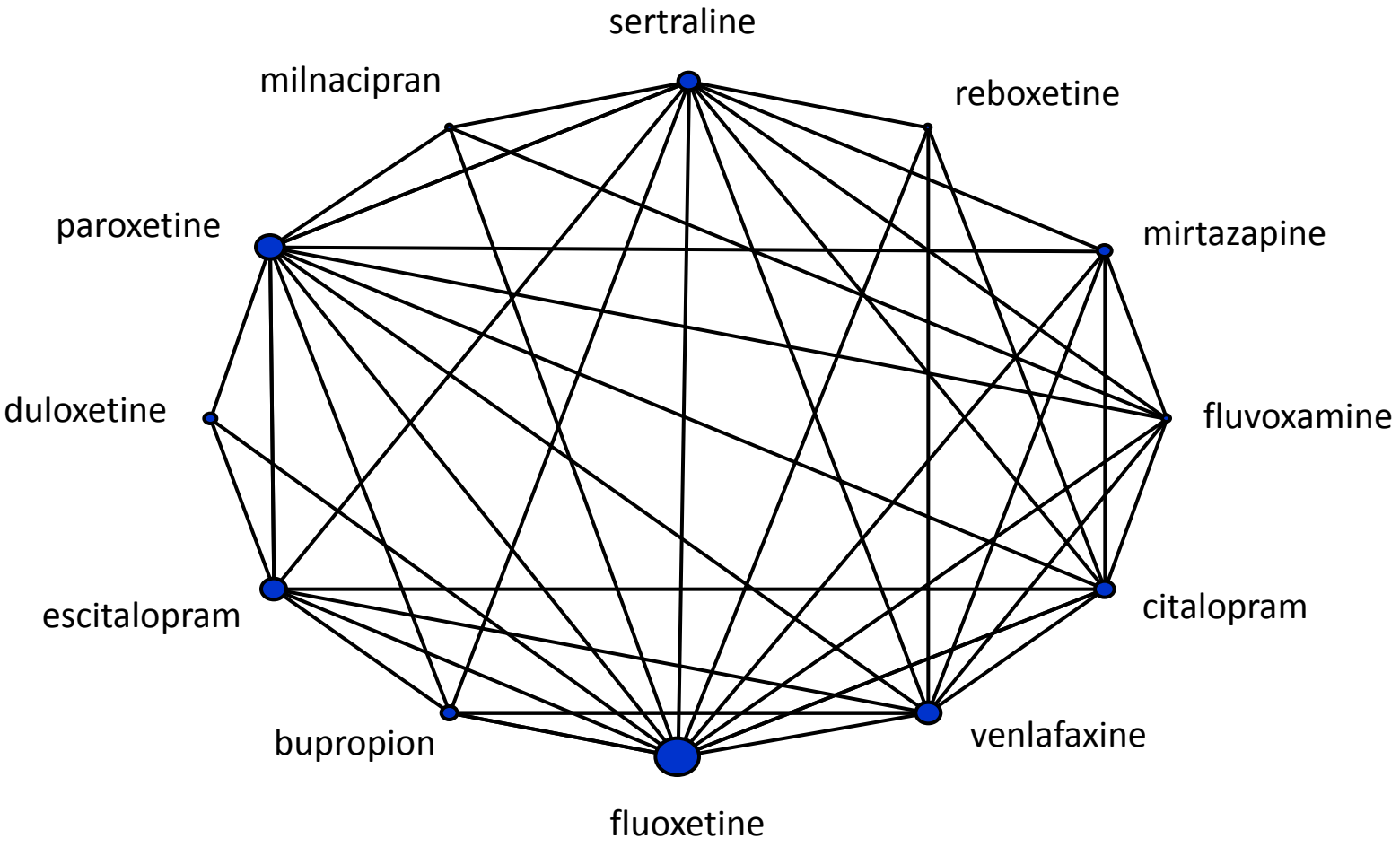
Straightforward & conceptually intuitive BUT it is very LIMITED:

- Pool separately for each treatment comparison (separate meta-analyses).
- Conduct indirect comparison (if appropriate).
- Combine with direct comparison (if appropriate).

What happens when:

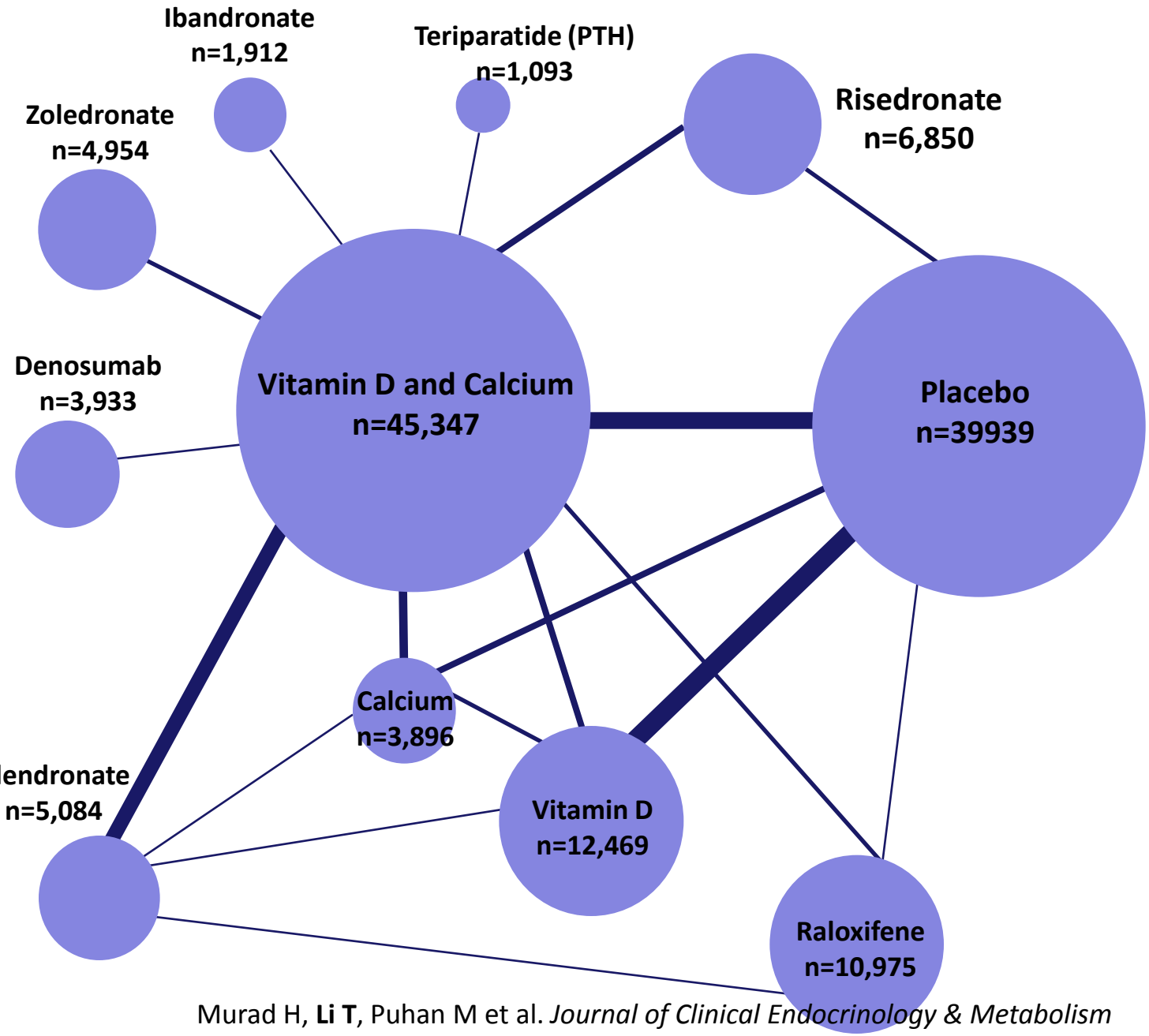
Treatments	4	5	6	7	8	9	10	11
Pairwise	6	10	15	21	28	36	45	55
Indirect	12	30	60	105	168	252	360	495

New-generation anti-depressants
12 treatments



Cipriani et al Lancet 2009

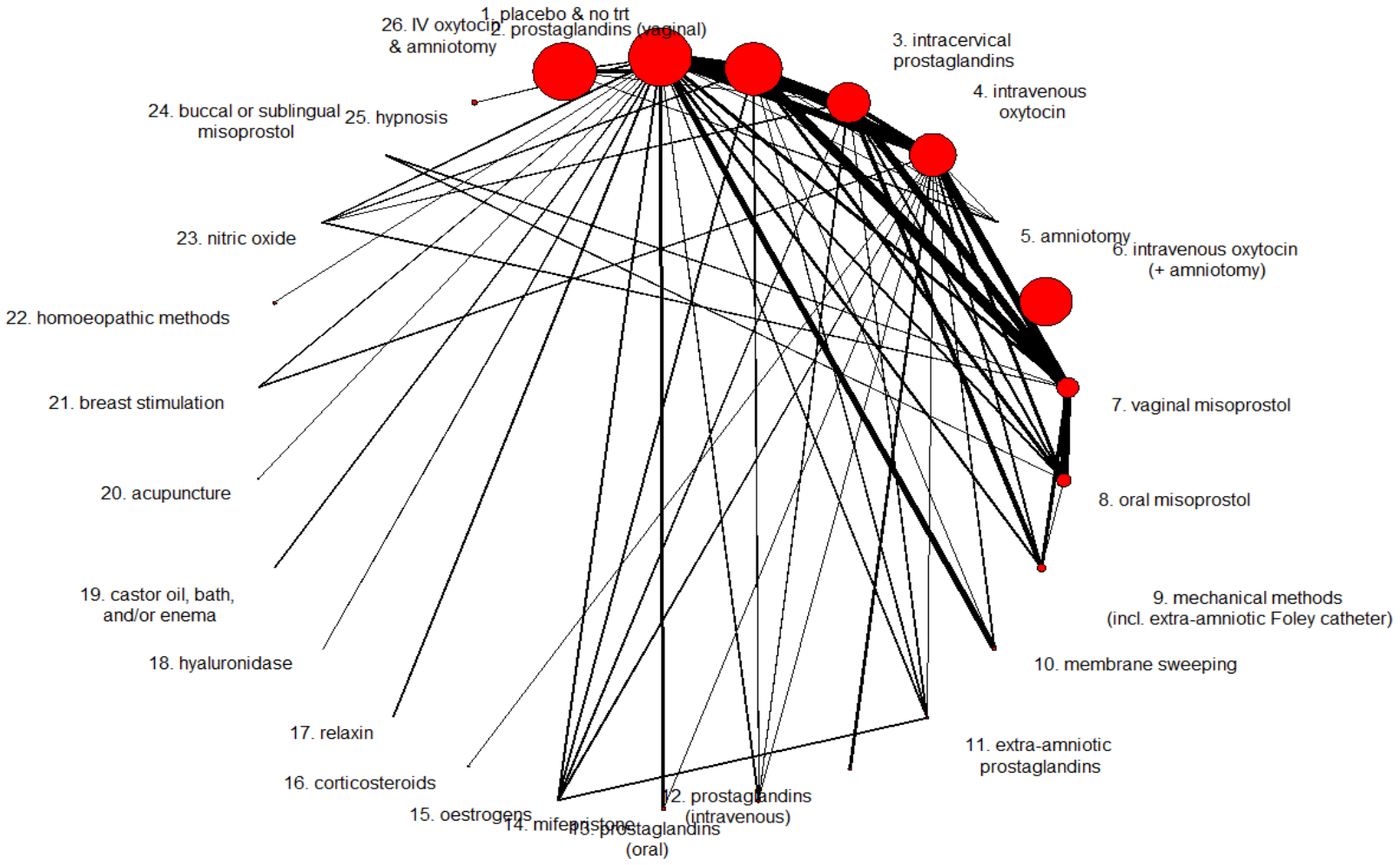
Treatment of Osteoporosis and Risk of Hip Fracture – 11 treatments



Murad H, **Li T**, Puhan M et al. *Journal of Clinical Endocrinology & Metabolism*

KER UNIT
Mayo Knowledge & Encounter Research Unit ◆ Mayo Clinic College of Medicine

Methods of induction for labour: 26 treatments

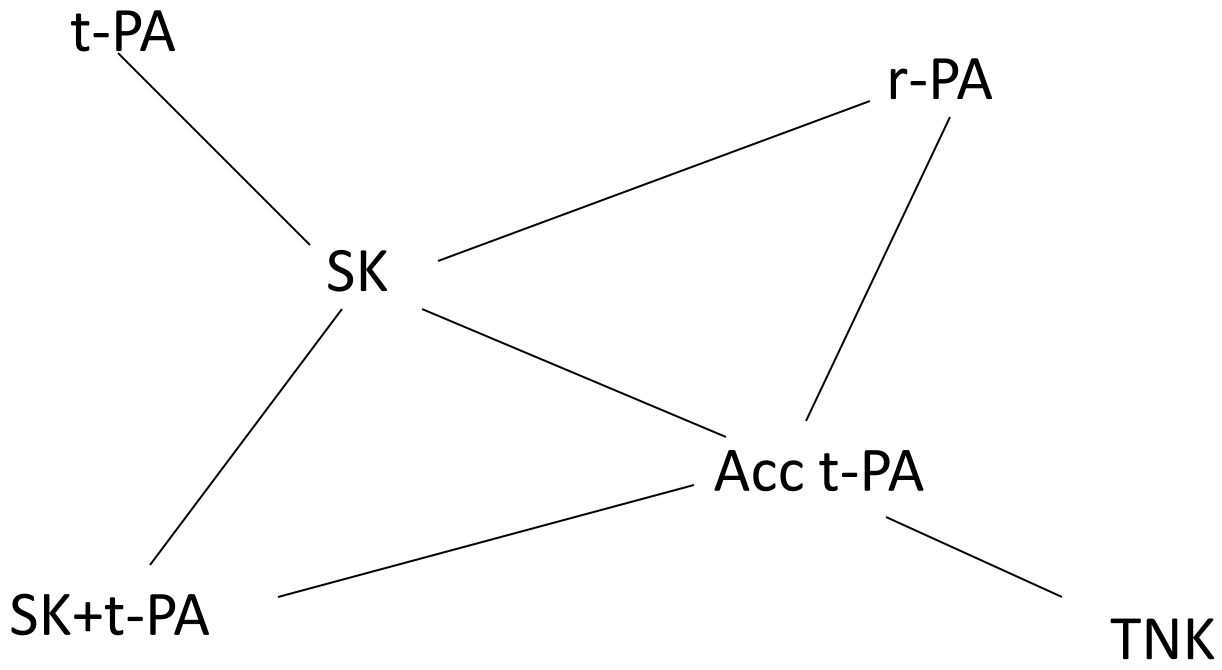


With thanks to: Leanne Jones, Therese Dowswell and Zarko Alfirevic (Pregnancy & Childbirth group)

Methods for larger networks

- Multiple treatment/ mixed treatment/ Network meta-analysis
- Simultaneous comparison of multiple treatments can only be done in a SINGLE ANALYSIS
  - Using frequentist (Stata command `mvmeta`)
  - or Bayesian approach (WinBUGS – code available from Bristol and Ioannina websites)
- Desire to determine which competing treatment is BEST?
  - Ranking of treatments using simulation approach
  - Estimates probability each treatment is the best.

Example: Thrombolysis network



Trial level data: 35-day mortality

Trial name	SK	t-PA	at-PA	SK+t-PA	r-PA	TNK
CI	9/130	6/123				
Cherng	5/63	2/59				
ECSG	3/65	3/64				
GISSI-2	887/10,396	929/ 10,372				
ISIS-3	1455/ 13,780	1418/ 13,746				
PAIMS	7/ 85	4/86				
TIMI-1	12/ 159	7/157				
White	10/ 135	5/135				
GUSTO-1	1472/ 20,251		652/ 10,396	723/ 10,374		
KAMIT	4/107				6/109	
INJECT	285/3004				270/3006	
GUSTO-3			356/ 4921		757/ 10,138	
RAPID-2			13/155		7/169	
ASSENT-2			522/8488			523/8461

Pairwise, 7 fixed effect meta-analyses (OR)

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00**	**0.86**	**0.96**	**0.95**	
t-PA		X				
Acc t-PA			X	**1.12**	**1.02**	**1.01**
t-PA+SK				X		
r-PA					X	
TNK						X

Conclusions from 7 pairwise meta-analyses

None achieves conventional statistical significance:

1. Streptokinase is as effective as non-accelerated alteplase.
2. Tenecteplase is as effective as accelerated alteplase
3. Reteplase is at least as effective as streptokinase.
4. Reteplase is possibly as effective as accelerated alteplase
5. No conclusion drawn for treatments forming three-arm trial

Fixed effect, pairwise meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B$$

Each trial compares treatments *A* and *B*

Study	SK (A)		t-PA (B)	
	r1	n2	r2	n2
CI	9	30	6	123
Cherng	5	63	2	59
ECSG	3	65	3	64
GISSI-2	887	10396	929	10372
ISIS-3	1455	13780	1418	13746
PAIMS	7	85	4	86
TIMI-1	12	159	7	157
White	10	135	5	135

Fixed effect, pairwise meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B$$

Each trial compares treatments *A* and *B*

$$\text{log-odds}(p_{jA}) = \mu_{jA} \quad \text{for arm A (SK)}$$

$$\text{log-odds}(p_{jB}) = \mu_{jB} + d_{AB} \quad \text{for arm B (t-PA)}$$

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

$$\text{log-odds}(p_{jb}) = \mu_{jb} \quad \text{for arm } b$$

$$\text{log-odds}(p_{jk}) = \mu_{jk} + d_{bk} \quad \text{for arm } k$$

Network meta-analysis is a generalisation of pairwise meta-analysis.

Network meta-analysis is an extension of pairwise meta-analysis

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

Study	Treatment indicator	SK (A)		t-PA (B)		A-tPA (C)		SK+t-PA (D)		r-PA (E)		TNK (F)	
		r1	n2	r2	n2	r3	n3	r4	n4	r5	n5	r6	n6
CI	B vs A	9	30	6	123								
Cherng	B vs A	5	63	2	59								
ECSG	B vs A	3	65	3	64								

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

Study	Treatment indicator	SK (A)		t-PA (B)		A-tPA (C)		SK+t-PA (D)		r-PA (E)		TNK (F)	
		r1	n2	r2	n2	r3	n3	r4	n4	r5	n5	r6	n6
CI	B vs A	9	30	6	123								
Cherng	B vs A	5	63	2	59								
ECSG	B vs A	3	65	3	64								
KAMIT	E vs A	4	107							6	109		
INJECT	E vs A	285	3004							270	3006		

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

Study	Treatment indicator	SK (A)		t-PA (B)		A-tPA (C)		SK+t-PA (D)		r-PA (E)		TNK (F)	
		r1	n2	r2	n2	r3	n3	r4	n4	r5	n5	r6	n6
CI	B vs A	9	30	6	123								
Cherng	B vs A	5	63	2	59								
ECSG	B vs A	3	65	3	64								
KAMIT	E vs A	4	107							6	109		
INJECT	E vs A	285	3004							270	3006		
GUSTO-3	E vs C					356	4921			757	10138		
ASSENT2	F vs C					522	8488					523	8461

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

Study	Treatment indicator	SK (A)		t-PA (B)		A-tPA (C)		SK+t-PA (D)		r-PA (E)		TNK (F)	
		r1	n2	r2	n2	r3	n3	r4	n4	r5	n5	r6	n6
CI	B vs A	9	30	6	123								
Cherng	B vs A	5	63	2	59								
ECSG	B vs A	3	65	3	64								
GUSTO-1	D vsC vs A	1472	20251			652	10396	723	10374				
KAMIT	E vs A	4	107							6	109		
INJECT	E vs A	285	3004							270	3006		
GUSTO-3	E vs C					356	4921			757	10138		
ASSENT2	F vs C					522	8488					523	8461

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

$$\text{log-odds}(p_{jb}) = \mu_{jb} \quad \text{for arm } b$$

$$\text{log-odds}(p_{jk}) = \mu_{jk} + d_{bk} \quad \text{for arm } k$$

Fixed effect, network meta-analysis

Number of events, $r_{j,k}$, out of total, $n_{j,k}$, on treatment *k* in study *j*

$$j = 1, \ldots, NS \qquad k = A, B, C, D \ldots, NT$$

Each trial compares treatments *b* and *k*

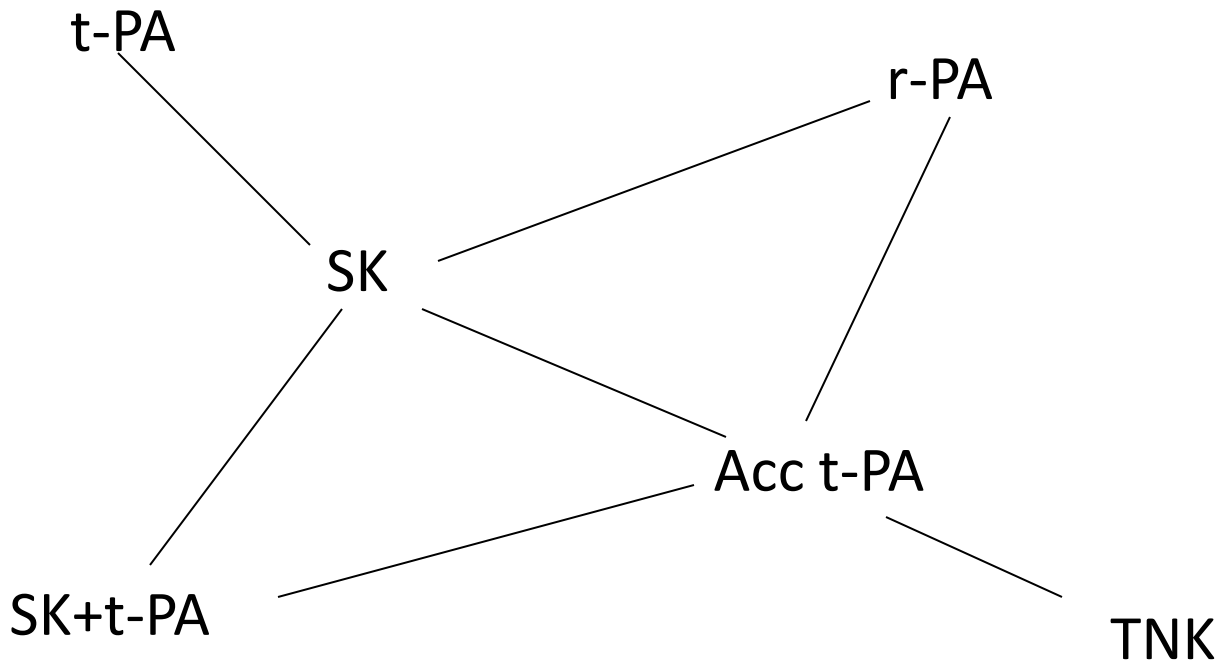
$$\text{log-odds}(p_{jb}) = \mu_{jb} \quad \text{for arm } b$$

$$\text{log-odds}(p_{jk}) = \mu_{jk} + (d_{Ak} - d_{Ab}) \quad \text{for arm } k$$

This should look familiar??

Higgins & Whitehead, *Statistics in Medicine* (1996).

Example: Thrombolysis network



Upper right: pair-wise ORs
Lower left: MTC ORs

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00**	**0.86**	**0.96**	**0.95**	
t-PA	**1.00**	X				
Acc t-PA	**0.87**	**0.87**	X	**1.12**	**1.02**	**1.01**
t-PA+SK	**0.96**	**0.97**	**1.11**	X		
r-PA	**0.90**	**0.91**	**1.04**	**0.94**	X	
TNK	**0.87**	**0.88**	**1.01**	**0.91**	**0.97**	X

95% Credible Intervals: Availability of direct evidence

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00** **0.94-1.06**	**0.86**	**0.96**	**0.95**	
t-PA	**1.00** **0.94-1.06**	X				
Acc t-PA	**0.87**	**0.87**	X	**1.12**	**1.02**	**1.01**
t-PA+SK	**0.96**	**0.97**	**1.11**	X		
r-PA	**0.90**	**0.91**	**1.04**	**0.94**	X	
TNK	**0.87**	**0.88**	**1.01**	**0.91**	**0.97**	X

95% Credible Intervals: Missing evidence for SK vs TNK

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00**	**0.86**	**0.96**	**0.95**	
t-PA	**1.00**	X				
Acc t-PA	**0.87**	**0.87**	X	**1.12**	**1.02**	**1.01**
t-PA+SK	**0.96**	**0.97**	**1.11**	X		
r-PA	**0.90**	**0.91**	**1.04**	**0.94**	X	
TNK	**0.87** **0.74-1.00**	**0.88**	**1.01**	**0.91**	**0.97**	X

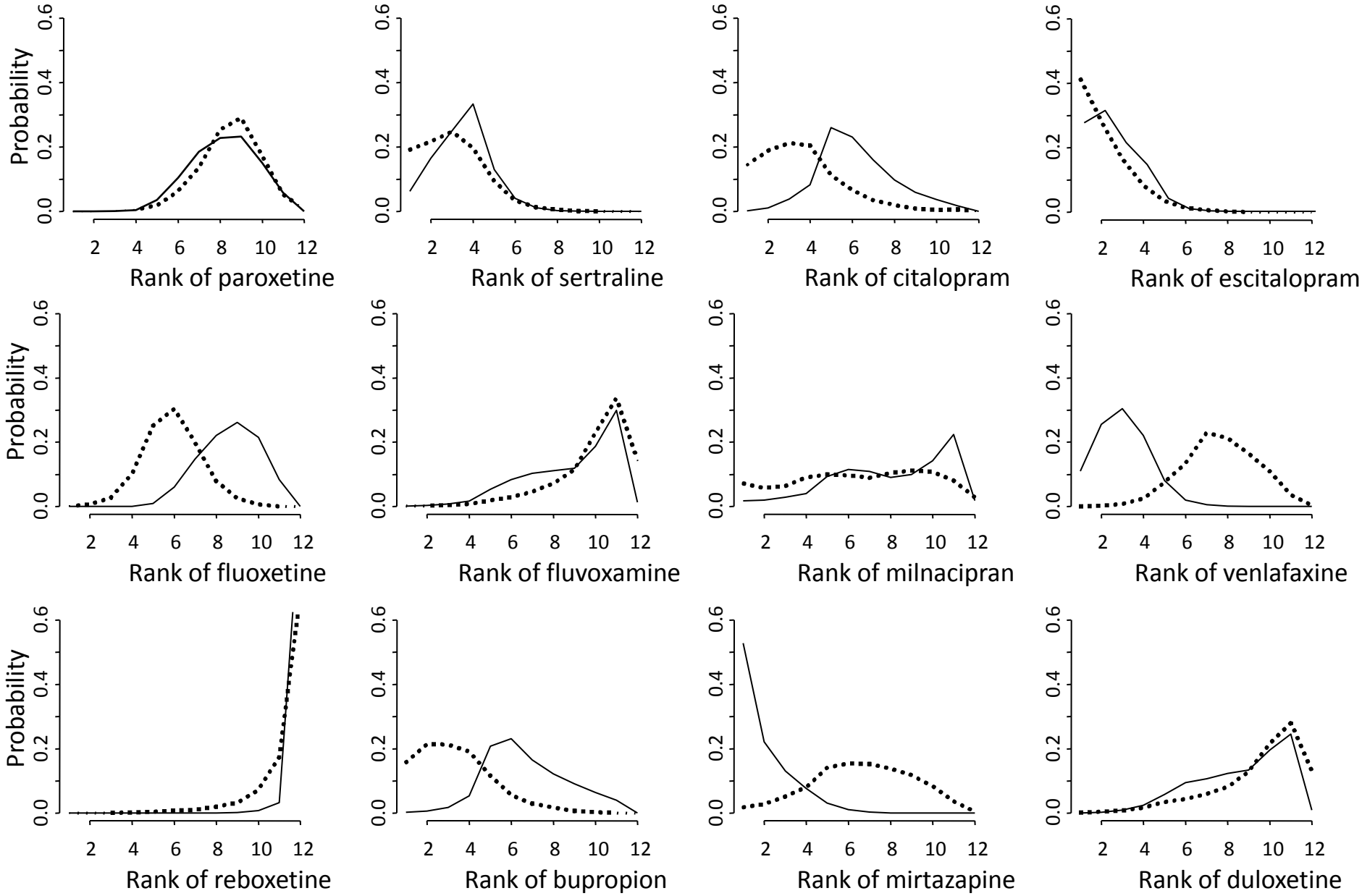
95% Credible Intervals: Increase in precision for At-PA vs r-PA

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00**	**0.86**	**0.96**	**0.95**	
t-PA	**1.00**	X				
Acc t-PA	**0.87**	**0.87**	X	**1.12**	**1.02** (0.90-1.16)	**1.01**
t-PA+SK	**0.96**	**0.97**	**1.11**	X		
r-PA	**0.90**	**0.91**	**1.04** (0.94-1.16)	**0.94**	X	
TNK	**0.87**	**0.88**	**1.01**	**0.91**	**0.97**	X

Probability each treatment is 'best'

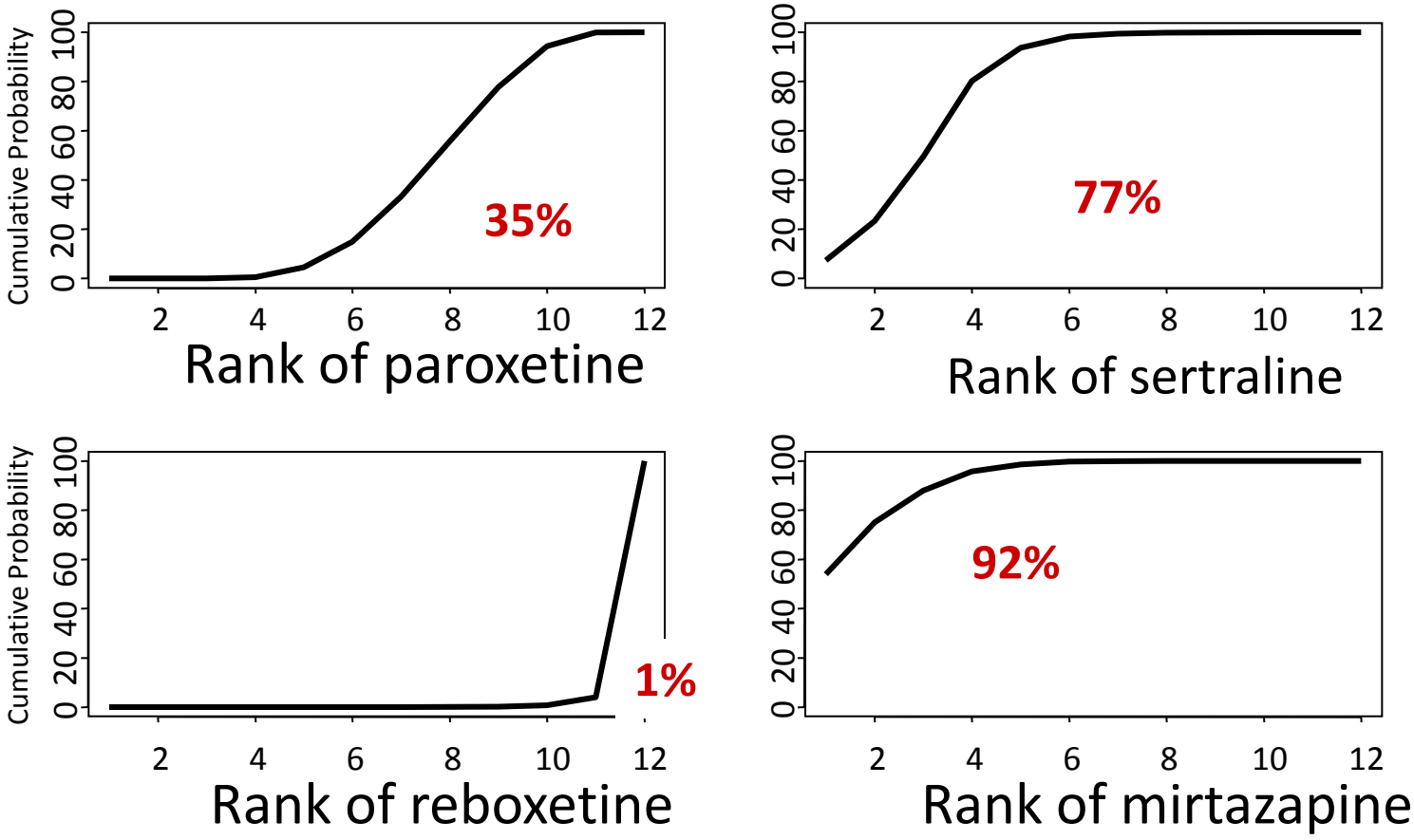
	Fixed effect	
	35 day Mortality %	Probability best
SK	6.5	0%
t-PA	6.4	0%
Acc t-PA	5.6	40%
SK + t-PA	6.2	1%
r-PA	5.8	15%
TNK	5.6	43%

NB: CMIMG MIF award will produce guidance on statistical methods, presentation of results & summarising findings from NMA.



Ranking for efficacy (solid line) and acceptability (dotted line). Ranking: probability to be the best treatment, to be the second best, the third best and so on, among the 12 comparisons).

Surface Under the Cumulative RAnking curve (SUCRA)



Salanti et al JCE 2011

What problems do IC/NMA solve ?

Direct evidence between active treatments B and C is not always available, e.g.

Indirect comparisons AB and AC can be used to infer the efficacy of B relative to C when direct evidence is lacking.

What problems do IC/NMA solve ?

Even when direct evidence is available, there may be not much of it.

NMA allows indirect evidence on BC to be pooled with direct data from BC trials. Reduces uncertainty in treatment effect estimates (increases precision), and inference based on more evidence – more robust.

What problems do IC/NMA solve ?

When SEVERAL treatments A,B, and C are to be compared, evidence that is “direct” for some comparisons is “indirect” for others, and the distinction becomes meaningless.

IC and NMA allows ALL evidence to be combined in a single internally consistent model. Treatments can then be ranked in efficacy, or cost-efficacy.

The important assumption

IC and NMA assume that the “Direct” and “Indirect” evidence estimate the same parameter.

That the treatment effect estimated by the BC trials, would be the same as the treatment effect estimated by the AC and AB trials (if they had included B and C arms).

Nearly all the doubts about the validity of IC and NMA can be traced to this assumption.

Websites of interest

General:

http://cmimg.cochrane.org/welcome

For WinBUGS code:

http://www.bristol.ac.uk/social-community-medicine/projects/mpes/mtc/

(developed by Nicky Welton, Sofia Dias and Tony Ades)

http://www.mtm.uoi.gr/

(developed by Georgia Salanti, Anna Chaimani, Dimitris Mavridis and Julian Higgins)

References

- Salanti G, Ades AE, Ioannidis JP. 2011 Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 64(2):163-71
- Murad H, **Li T**, Puhan M *et al.* 2012 Comparative Effectiveness of Drug Treatments to Prevent Fragility Fractures: A Systematic Review and Network Meta-Analysis *Journal of Clinical Endocrinology & Metabolism.* 97(6)
- Boland A, Dundar Y, Bagust A *et al*. 2003 Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. *Health Technology Assessment*.
- Glazener CM, Evans JH, Peto RE. 2003. Alarm interventions for nocturnal enuresis in children. Cochrane Database Systematic Reviews.
- Glenny AM, Altman DG, Song F *et al.* 2005 Indirect comparisons of competing interventions *Health Technology Assessment*; 9(26)
- Cipriani A, Furukawa TA, Salanti G, et al 2009 Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*. 373(9665):746-58.
- Caldwell DM, Ades AE & Higgins JPT 2005 Simultaneous Comparison of Multiple Treatments: combining direct and indirect evidence *BMJ* 331: 897-

Editorial considerations for reviews that compare multiple interventions

Said Business School, Oxford, UK

March 22, 2013