School of Medicine

ΙΑΤΡΙΚΗ ΣΧΟΛΗ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

European Research Council
(IMMA, Grant nr 260559)

Comparing multiple treatments 2: statistical methods for network meta-analysis

Dimitris Mavridis

University of Ioannina Medical School
Department of Hygiene and Epidemiology
Greece

Argie Veroniki

University of Ioannina Medical School
Department of Hygiene and Epidemiology
Greece

Acknowledgements: Georgia Salanti, Julian Higgins, Andrea Cipriani, Jo McKenzie & many clinicians

We have no actual or potential conflict of interest in relation to this presentation

12 new generation antidepressants

sertraline

milnacipran

reboxetine

paroxetine

mirtazapine

?

duloxetine

fluvoxamine

escitalopram

citalopram

bupropion

venlafaxine

fluoxetine

12 new generation antidepressants

A plethora **meta-analyses** has been published in the last years

"Although **Mirtazapine** is likely to have a faster onset of action than **Sertraline and Paroxetine** no significant differences were observed..."

"...statistically significant differences in terms of efficacy .... between **Fluoxetine and Venlafaxine**, but the clinical meaning of these differences is uncertain..."

"...**meta-analysis** highlighted a trend in favour of **Sertraline** over other **Fluoxetine**"

"**Venlafaxine** tends to have a favorable trend in response rates compared with **duloxetine**"

Fluoxetine: 28€ Venlafaxine:111€ Sertaline: 76 €

12 new generation antidepressants

sertraline

milnacipran

reboxetine

paroxetine

mirtazapine

?

duloxetine

fluvoxamine

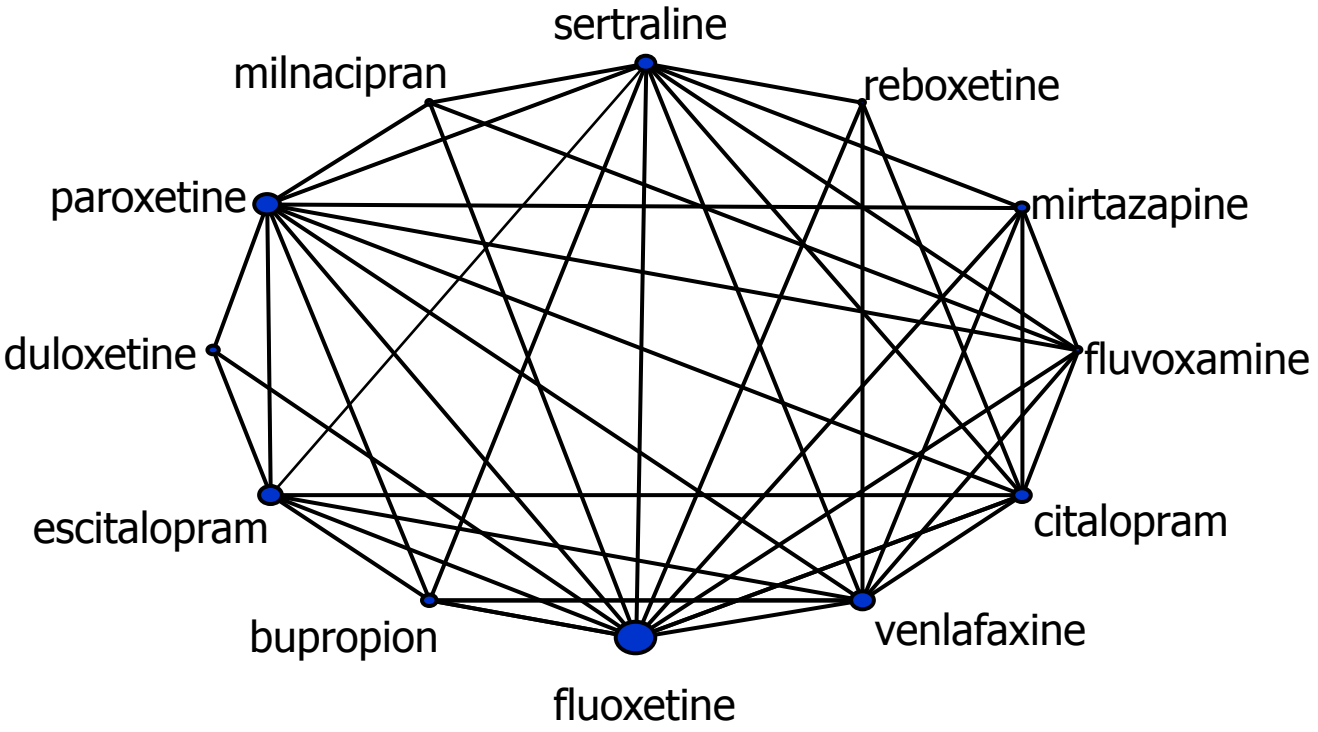
escitalopram

citalopram

bupropion

venlafaxine

fluoxetine



Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis

Andrea Cipriani, Toshiaki A Furukawa, Georgia Salanti, John R Geddes, Julian P T Higgins, Rachel Churchill, Norio Watanabe, Atsuo Nakagawa, Ichiro M Omori, Hugh McGuire, Michele Tansella, Corrado Barbui

Lancet 2009; 373: 746–58

Published Online January 29, 2009 DOI:10.1016/S0140-6736(09)60046-5

See Comment page 700

Department of Medicine and Public Health, Section of Psychiatry and Clinical

Summary

Background Conventional meta-analyses have shown inconsistent results for efficacy of second-generation antidepressants. We therefore did a multiple-treatments meta-analysis, which accounts for both direct and indirect comparisons, to assess the effects of 12 new-generation antidepressants on major depression.

Methods We systematically reviewed 117 randomised controlled trials (25 928 participants) from 1991 up to Nov 30, 2007, which compared any of the following antidepressants at therapeutic dose range for the acute treatment of unipolar major depression in adults: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine. The main outcomes were the proportion

12 new generation antidepressants

paroxetine —— reboxetine

duloxetine —— mirtazapine

escitalopram —— fluvoxamine

milnacipran —— citalopram

sertraline —— venlafaxine

bupropion —— fluoxetine

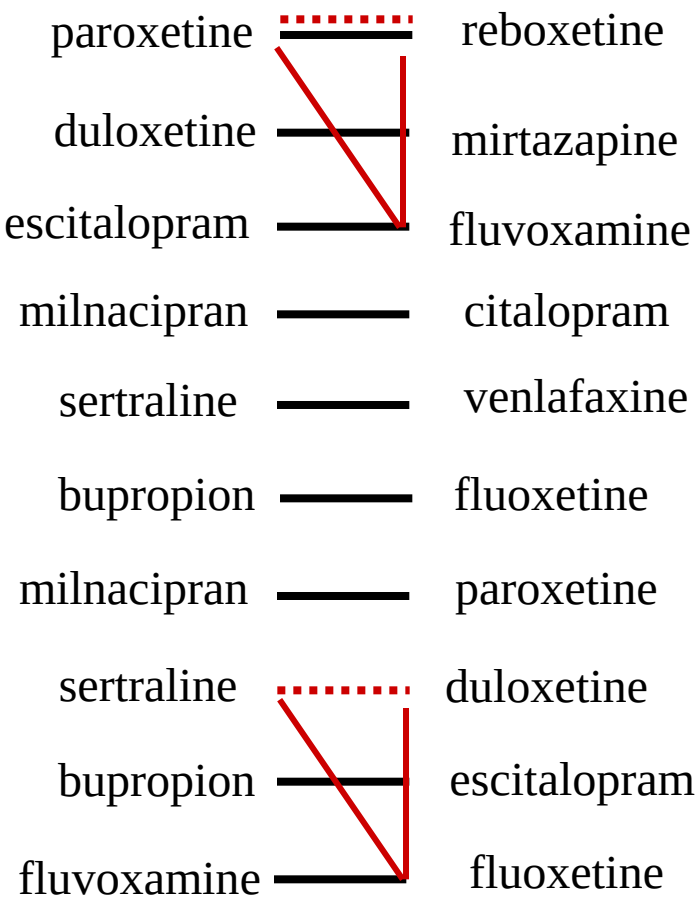
milnacipran —— paroxetine

sertraline ? duloxetine

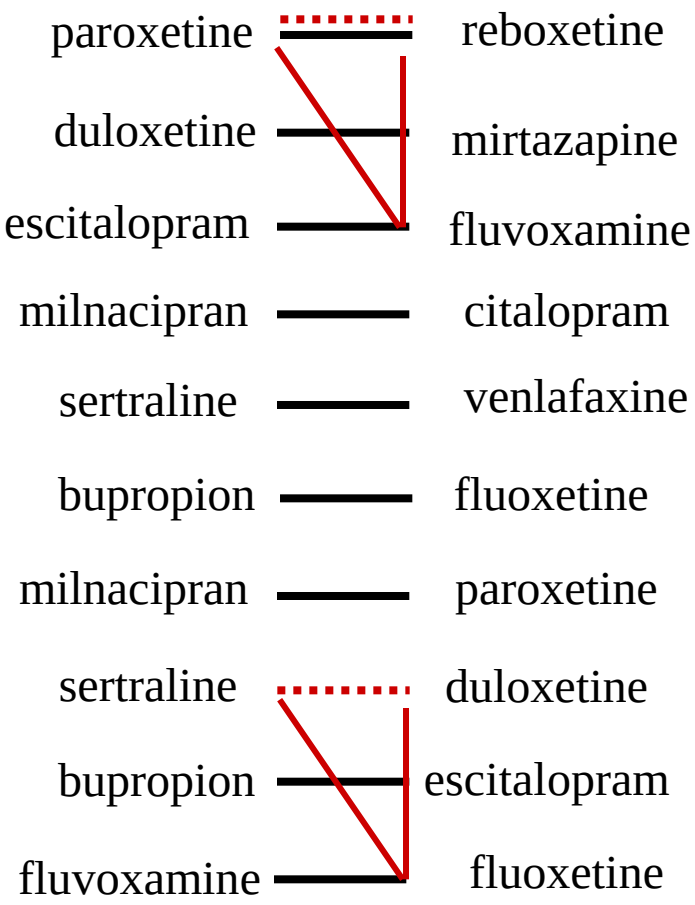
bupropion —— escitalopram

fluvoxamine —— fluoxetine

12 new generation antidepressants



12 new generation antidepressants



paroxetine	0%
sertraline	7%
citalopram	0%
escitalopram	26%
fluoxetine	0%
fluvoxamine	0%
milnacipran	1%
venlafaxine	11%
reboxetine	0%
bupropion	0%
mirtazapine	54%
duloxetine	0%

Probability of being the best

4 Fluoride modalities for preventing dental carries: series of pairwise meta-analyses

Treatment comparison		Studies
Placebo	Toothpaste	69
	Gel	13
	Rinse	31
	Varnish	3
Toothpaste	Rinse	6
Toothpaste	Varnish	1
Gel	Rinse	1
Gel	Varnish	*?*

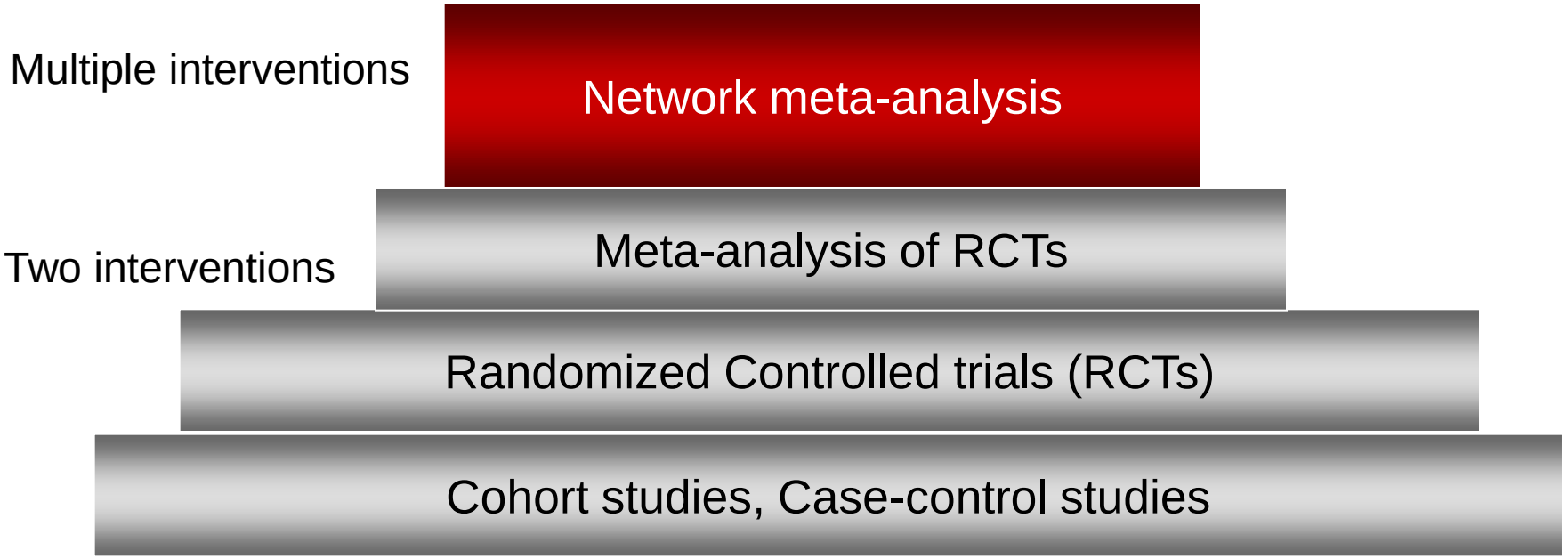
Multiple treatments and series of meta-analyses

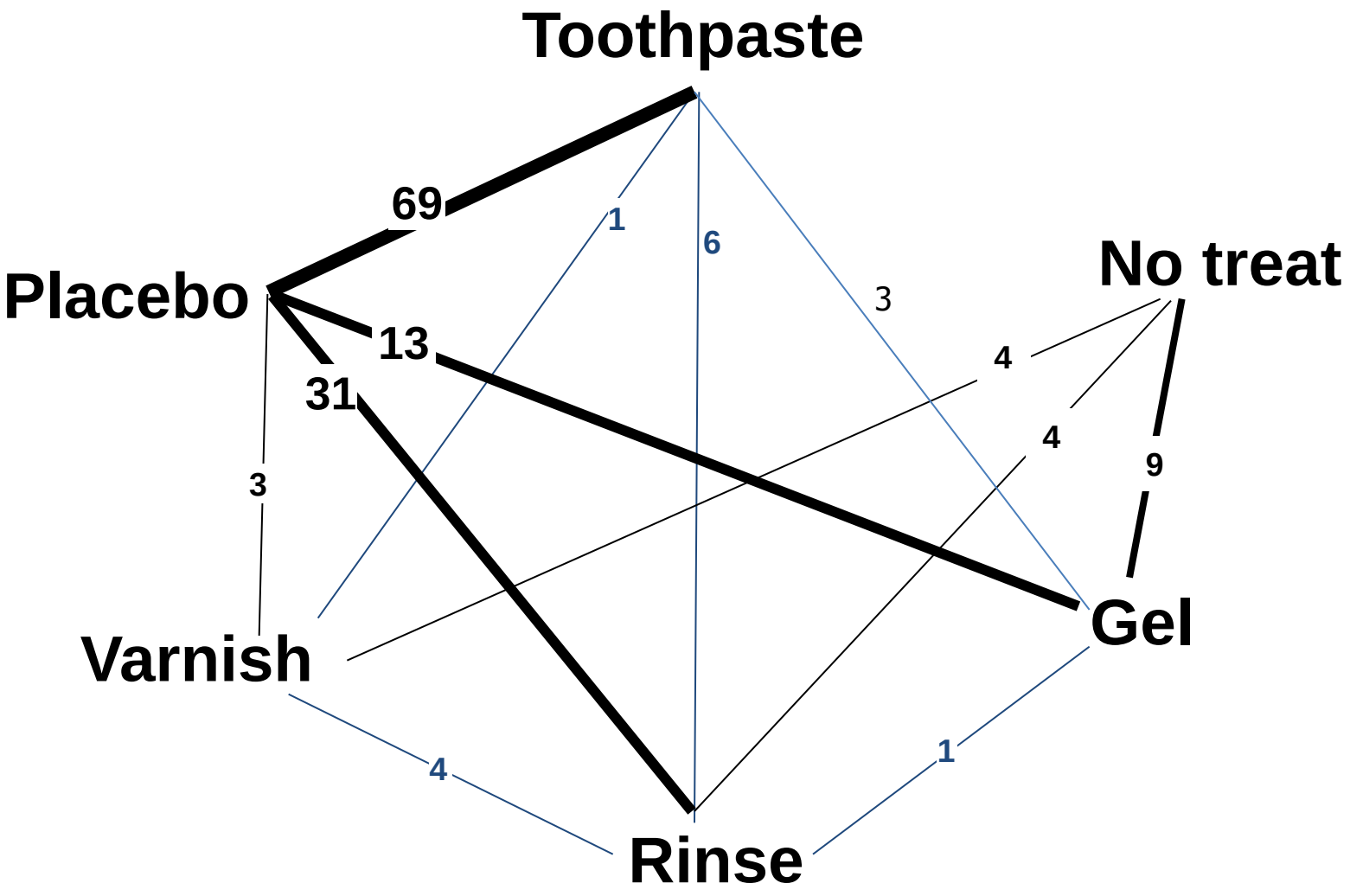
With pairwise meta-analyses we cannot answer the following questions:

- Which fluoride modality is the best?
- What is the ranking of fluoride treatments according to effectiveness?
- Which is better: Gel or Varnish (0 studies)

A new methodological framework

Other names:
Multiple treatments meta-analysis,
Mixed treatment comparisons

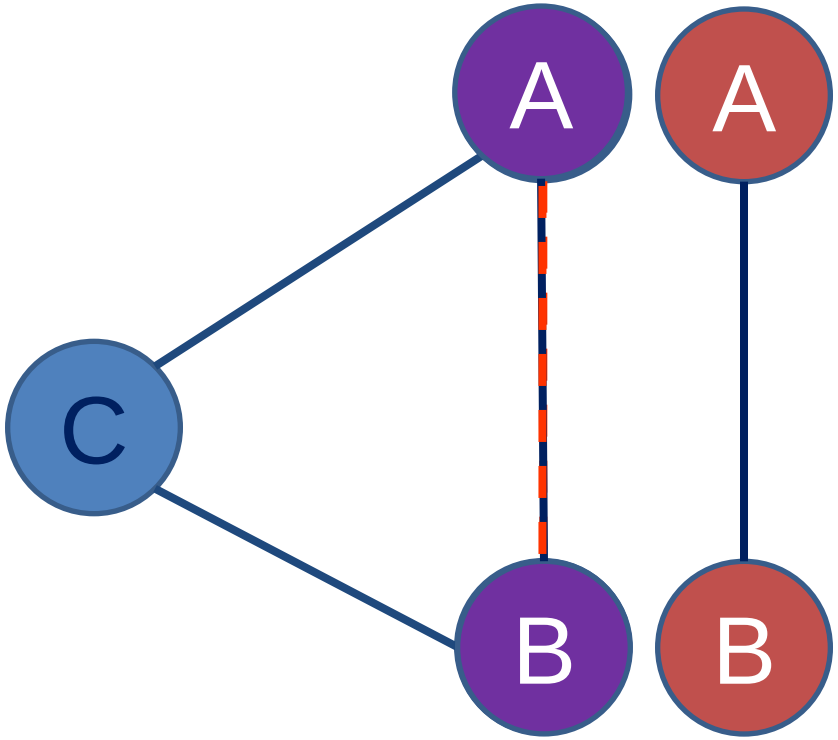



Toothpaste
69
1
6
3
No treat
Placebo
13
4
31
4
3
9
Gel
Varnish
4
1
Rinse

Aims of the workshop

- To explain **indirect** and **mixed** comparison of interventions
  - Assumptions
  - Statistical methods
- To understand the statistical models for **network meta-analysis**
- To discuss presentation of results from network meta-analysis
- To understand inconsistency models

Indirect and mixed effects



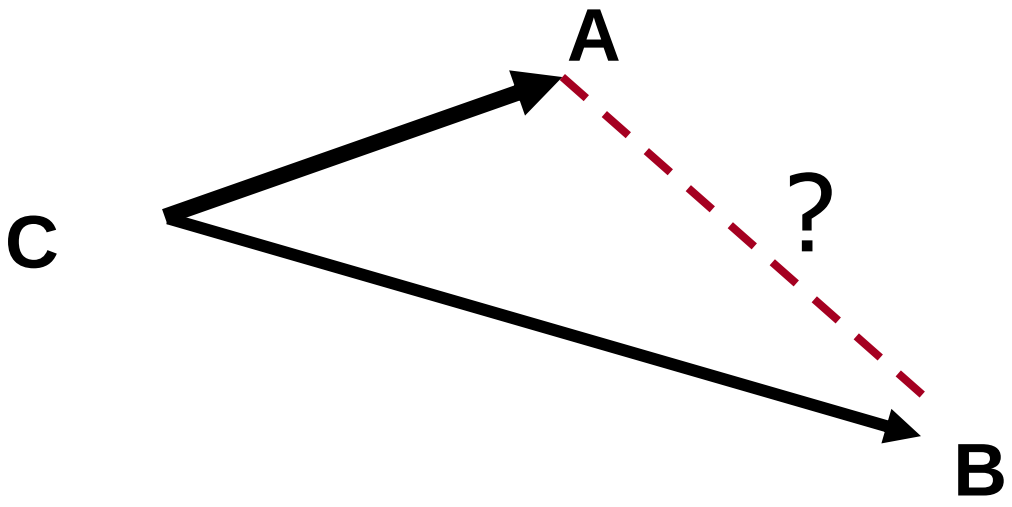
Indirect effect

Direct effect

Mixed effect

Indirect comparison

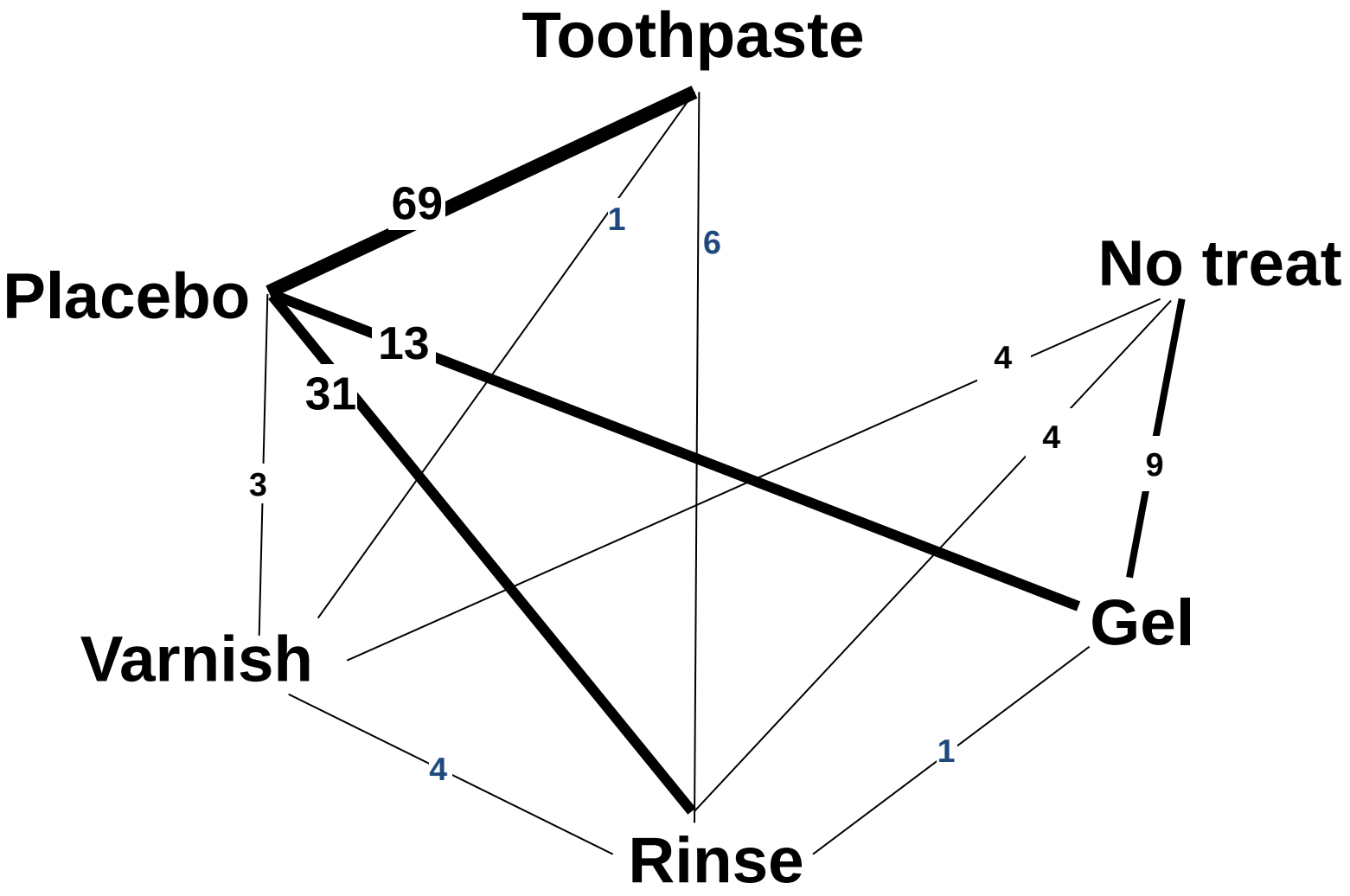
- We can obtain an *indirect* estimate for A vs B from RCTs comparing A vs C and B vs C:



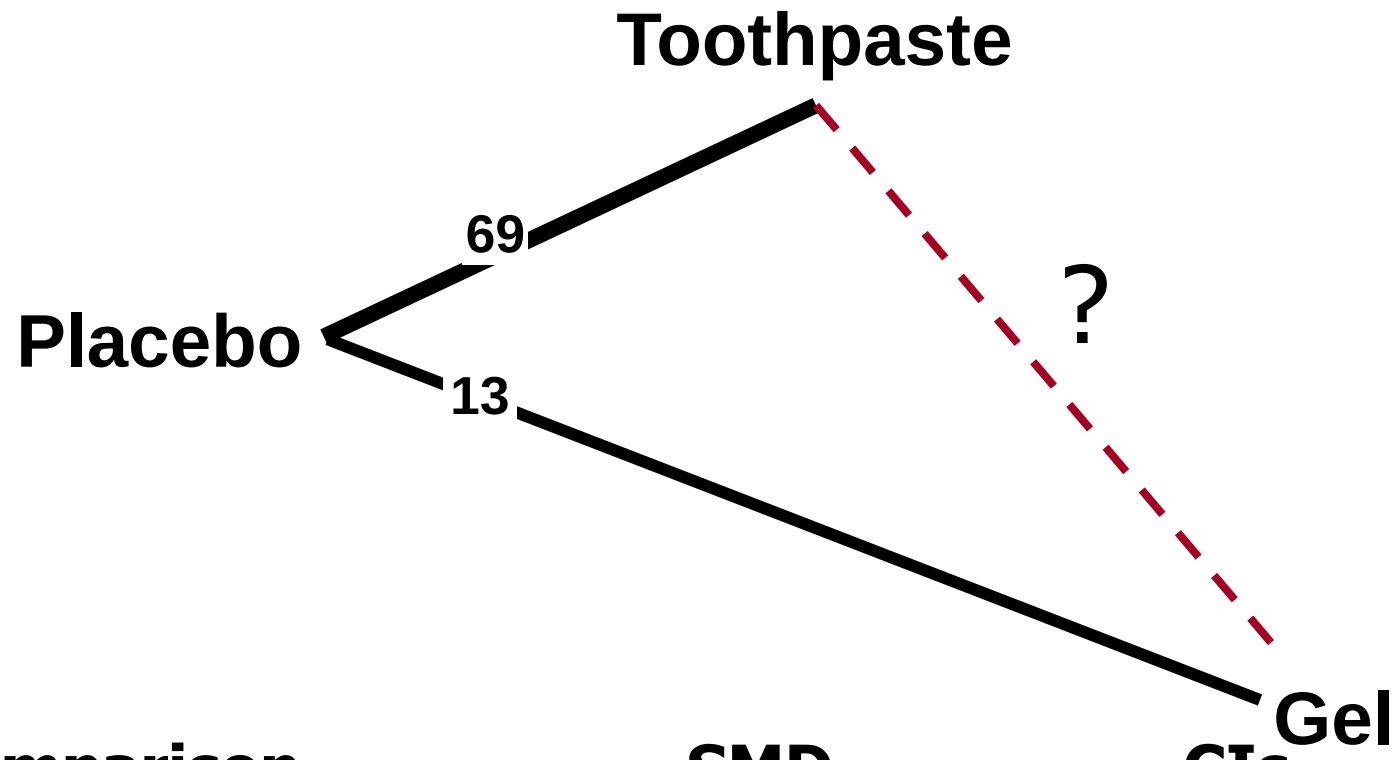
$$SMD_{AB} = SMD_{AC} + SMD_{CB}$$

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

$$Var(SMD_{AB}) = Var(SMD_{AC}) + Var(SMD_{BC})$$


Toothpaste
69
1
6
Placebo
No treat
13
31
4
4
9
3
Gel
Varnish
4
1
Rinse

Example



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

How to compare Gel to Toothpaste?

Exercise

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 SMD_{PvsT} = ((high CI – low CI)/3.92)^2

Variance SMD_{PvsT} = ((–0.28 – (–0.41))/3.92)^2 = 0.0011

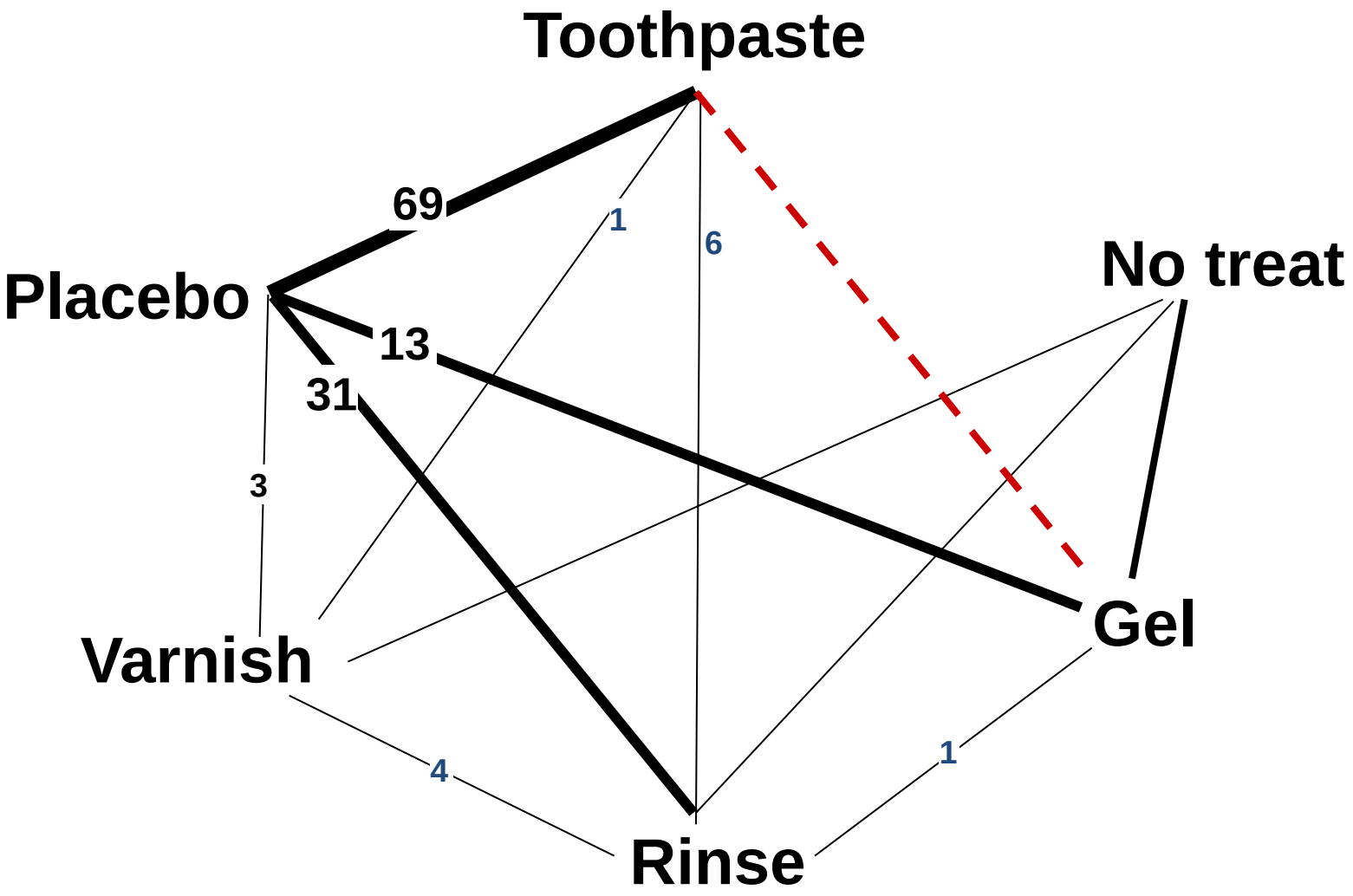
Variance SMD_{GvsT} = ((–0.10 – (–0.30))/3.92)^2 = 0.0026

Variance Indirect SMD_{GvsT} = 0.0011 + 0.0026 = 0.0037

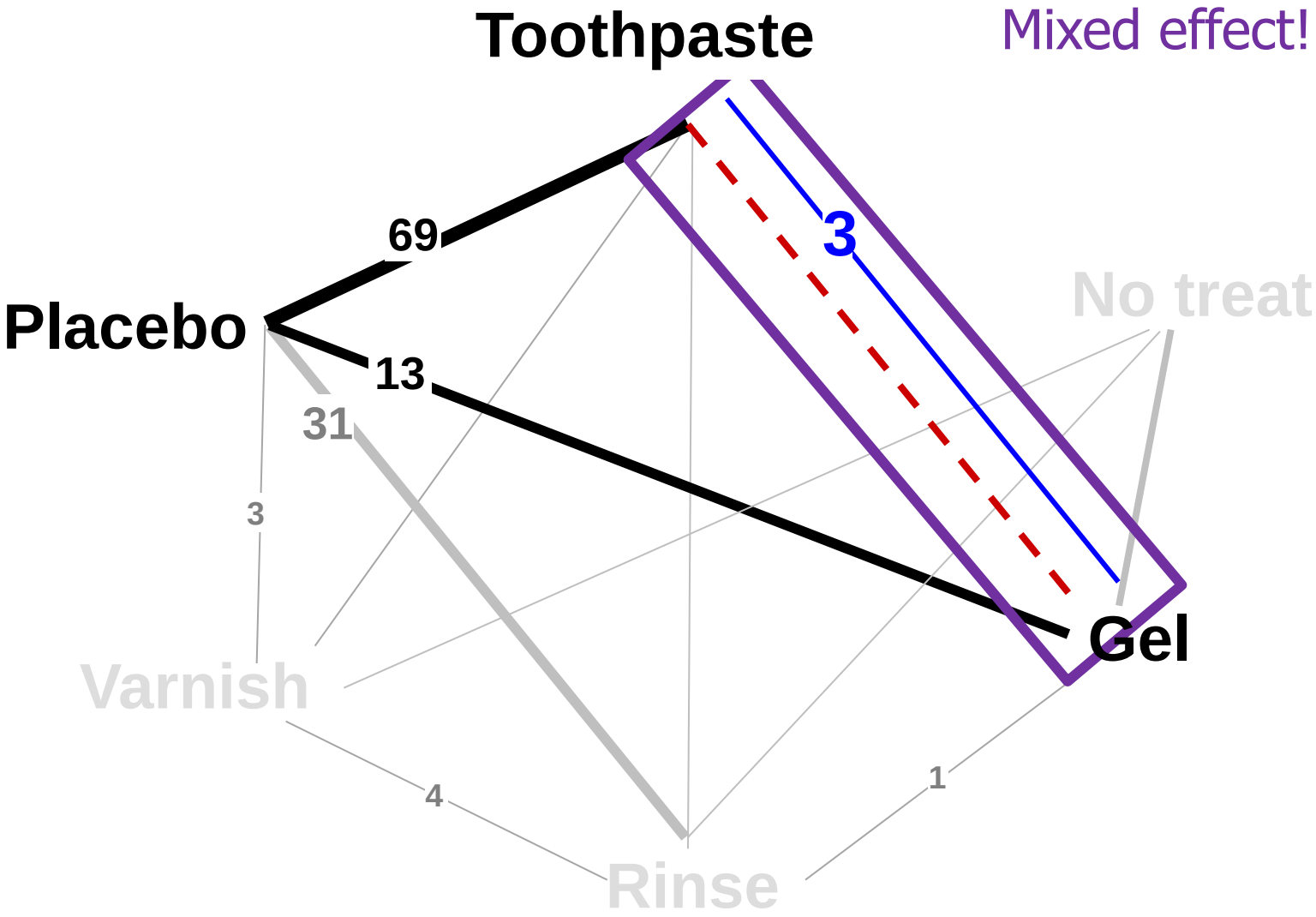
SE Indirect SMD_{GvsT} = sqrt(0.0037) = 0.061

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)



Indirect SMD Gel vs Toothpaste: -0.15 (0.0037)



Indirect SMD Gel vs Toothpaste: -0.15 (0.0037)

Mixed comparison

- Summarize direct and indirect effect size into a single mixed effect

$$\text{Mixed SMD} = \frac{\frac{SMD_{Direct}}{var(SMD_{Direct})} + \frac{SMD_{Indirect}}{var(SMD_{Indirect})}}{\frac{1}{var(SMD_{Direct})} + \frac{1}{var(SMD_{Indirect})}}$$

$$var(\text{Mixed SMD}) = \frac{1}{\frac{1}{var(SMD_{Direct})} + \frac{1}{var(SMD_{Indirect})}}$$

Mixed comparison

Indirect SMD_{GvsT} = - 0.15

Var(Indirect SMD_{GvsT})= 0.004

Direct SMD_{GvsT} = 0.04

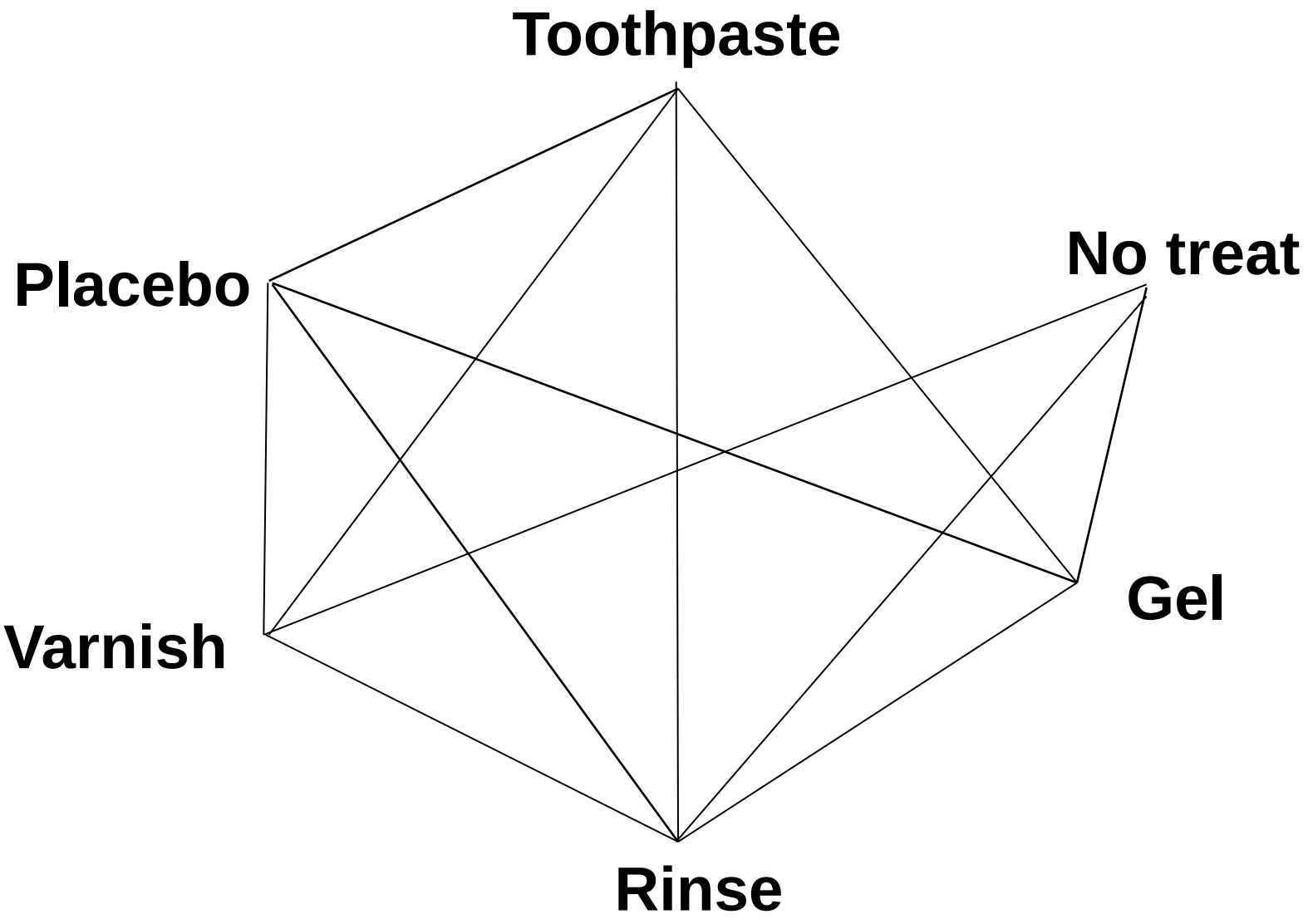
Var(Direct SMD_{GvsT})= 0.011

Mixed SMD_{GvsT} = -0.10

Var(Direct SMD_{GvsT})= 0.003

We gain precision!

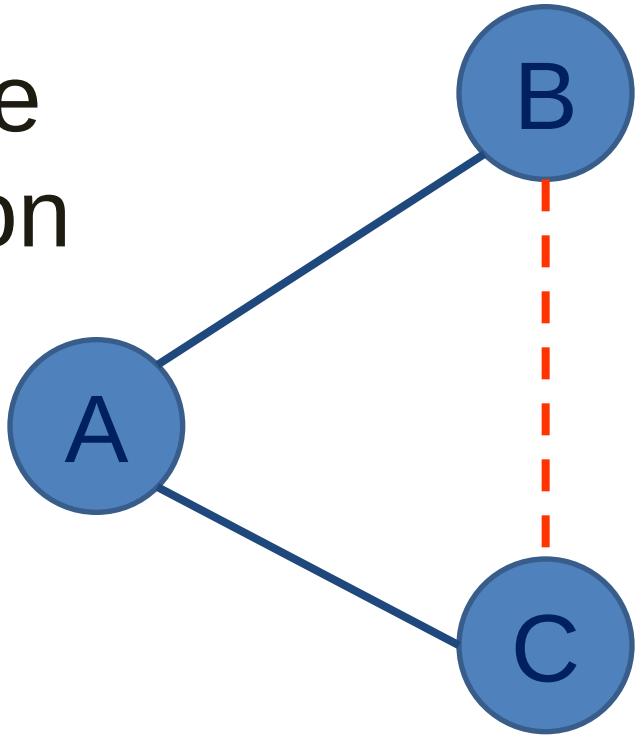
You can do this with any measure... lnOR, lnRR, RD, mean difference, HR, Peto's lnOR etc…



Extend the idea of mixed effect sizes in the entire network

Transitivity

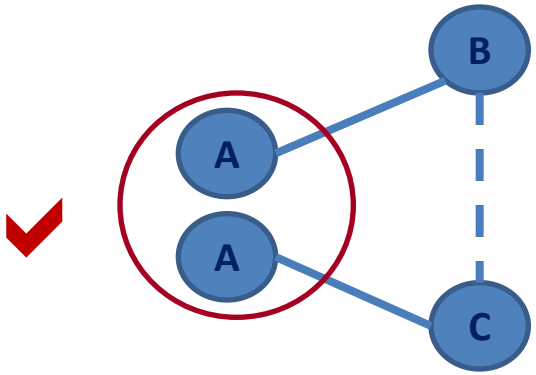
Untestable assumption



The anchor treatment A is 'transitive'

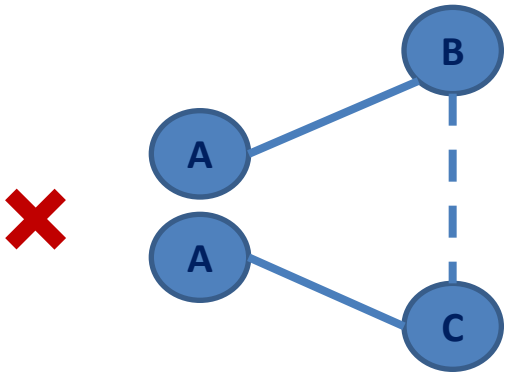
….but you can evaluate clinically and epidemiologically its plausibility

Transitivity requires... (1)



The ‘anchor’ treatment A to be similarly defined when it appears in AB and AC trials.

e.g. a treatment given at different doses but no systematic difference in the *average* dose of A across AB and AC studies

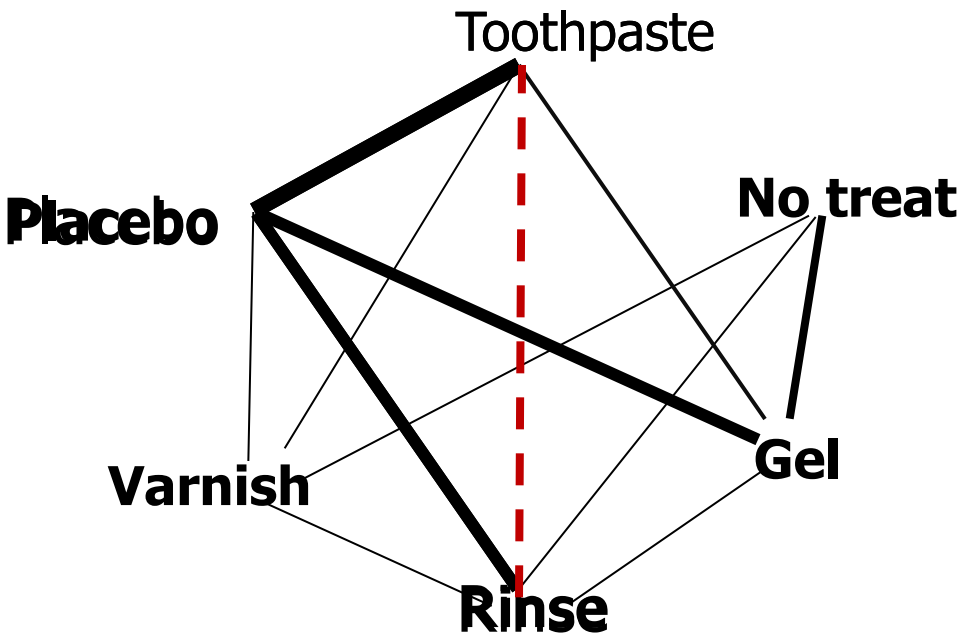


The ‘anchor’ treatment A may be different in AB and AC studies

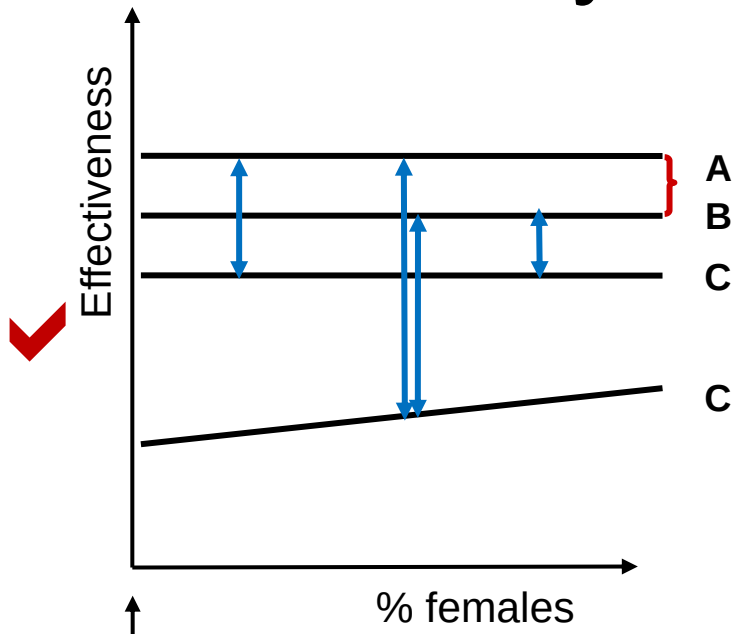
e.g. injection versus pill

Transitivity requires... (1)

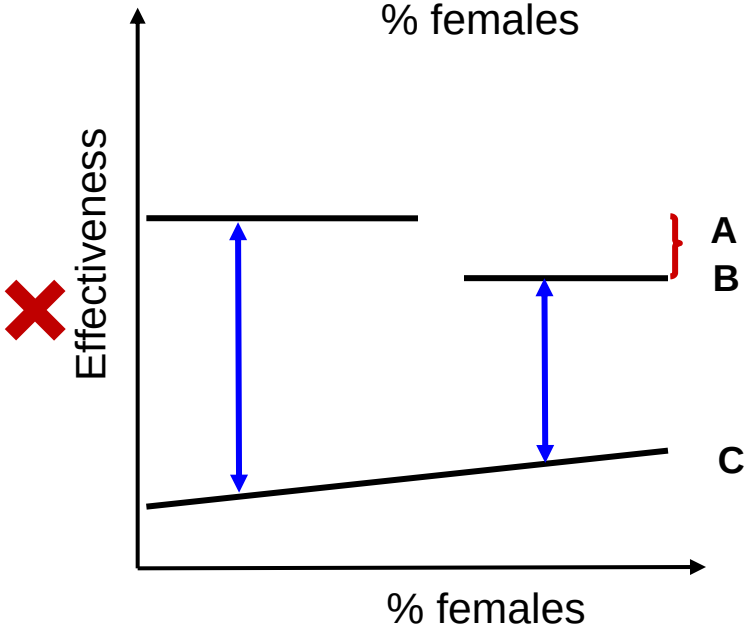
- However, placebo toothpaste and placebo rinse might not be comparable as the mechanical function of brushing might have a different effect on the prevention of caries.
- If this is the case, the transitivity assumption is doubtful (Salanti 2009).



Transitivity means that…



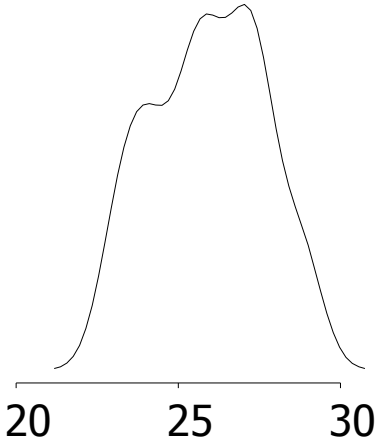
AC and BC trials do not differ with respect to the distribution of effect modifiers



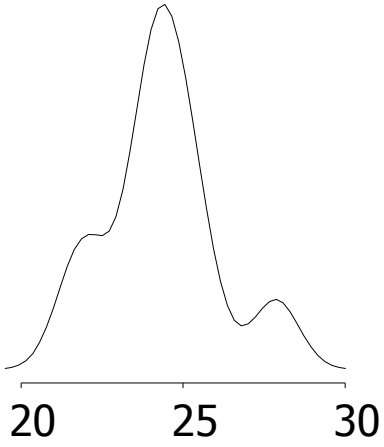
Difficult to defend when you have older and newer treatments, and variables are often unobserved

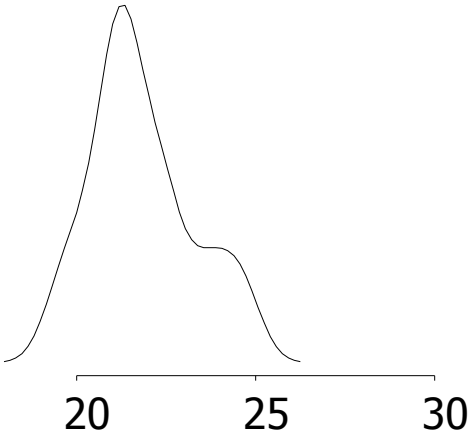
Distribution of mean dose of the active intervention in ten studies

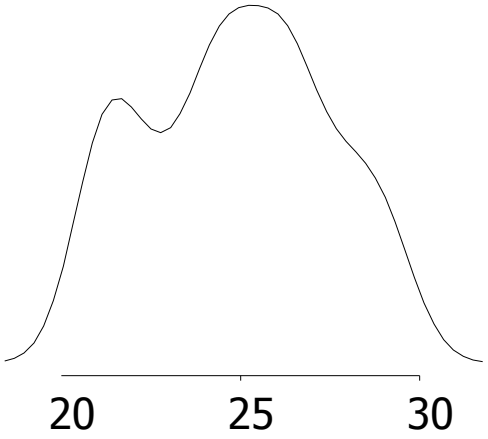
Placebo vs A



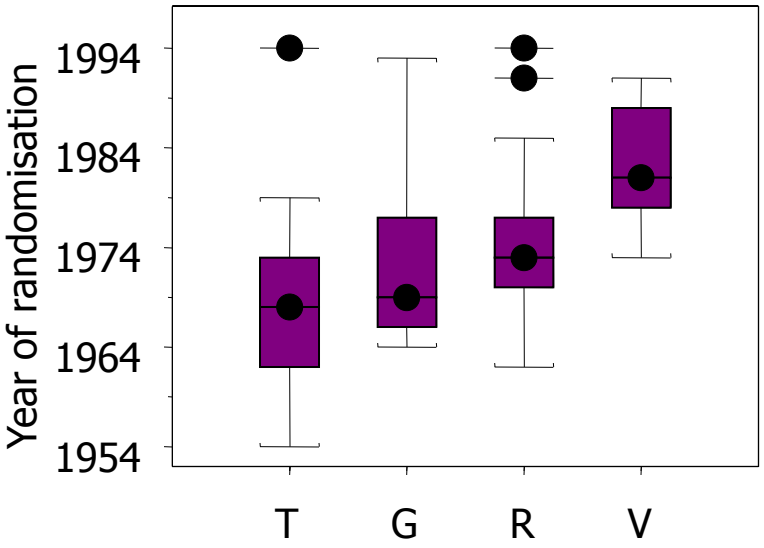
Placebo vs B

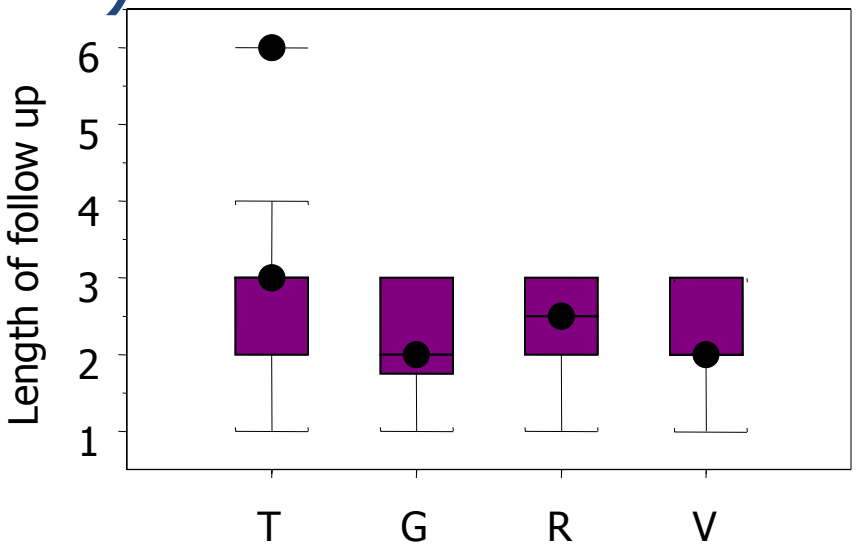


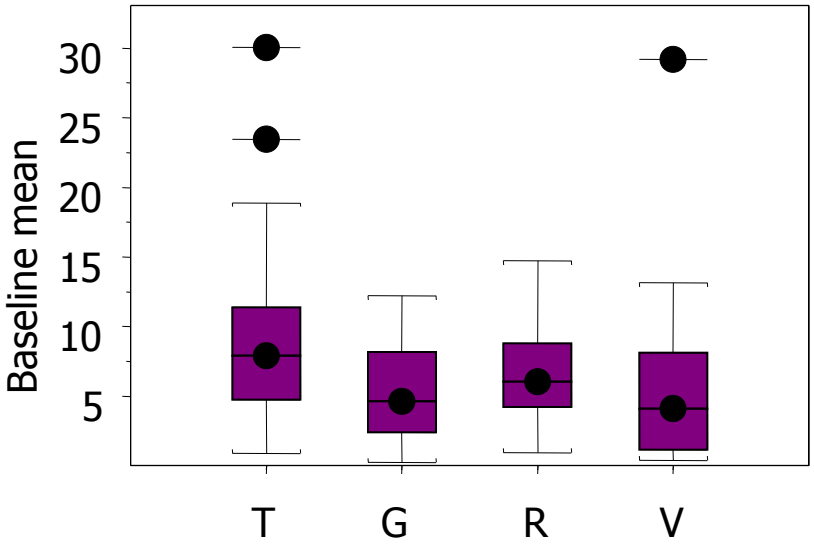


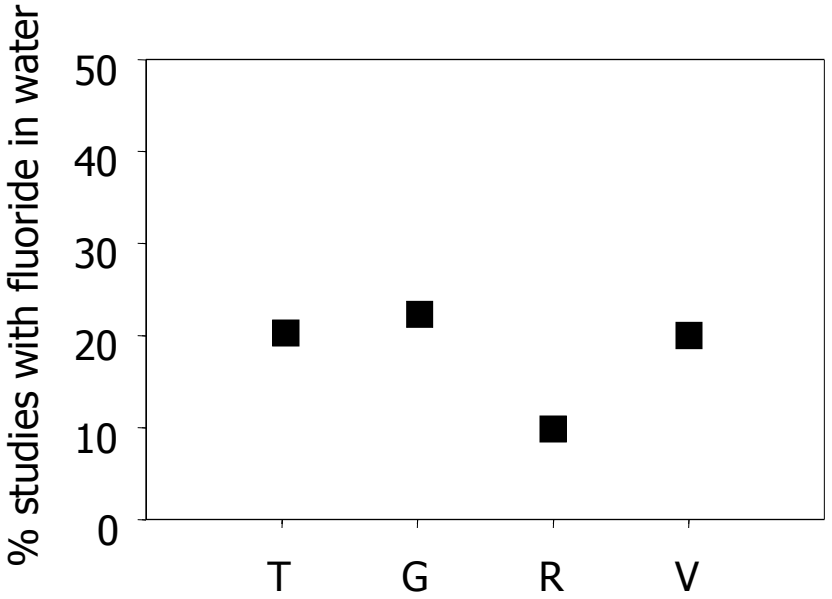


Compare the distribution of important characteristics across treatments (Salanti et al 2009)



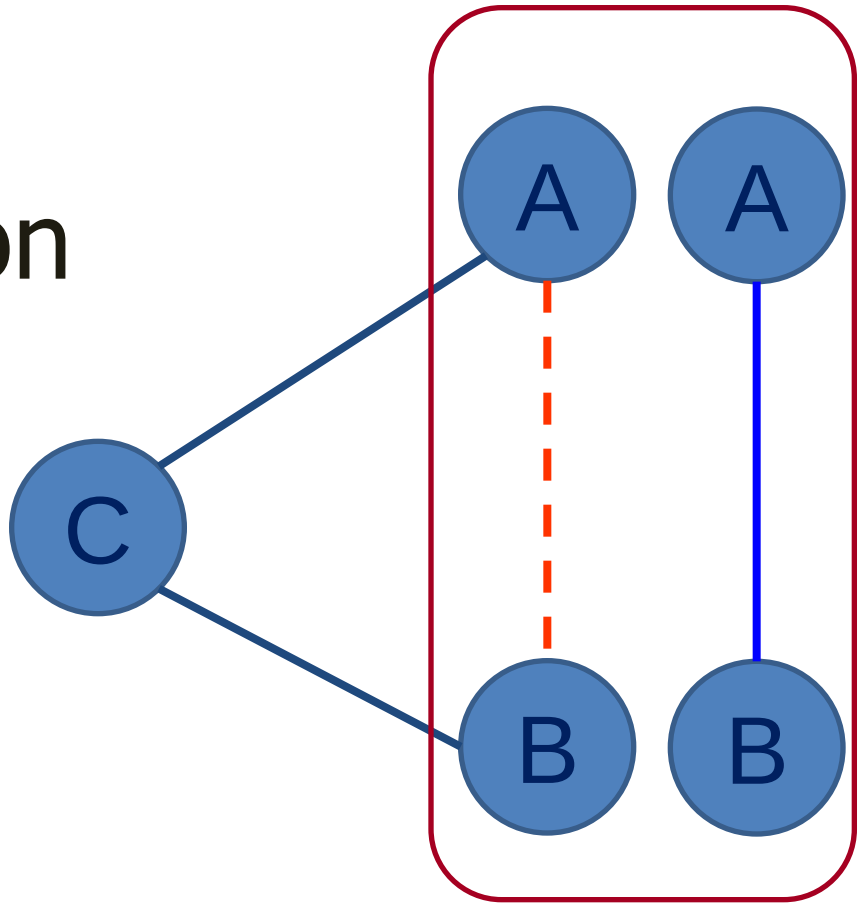






Consistency

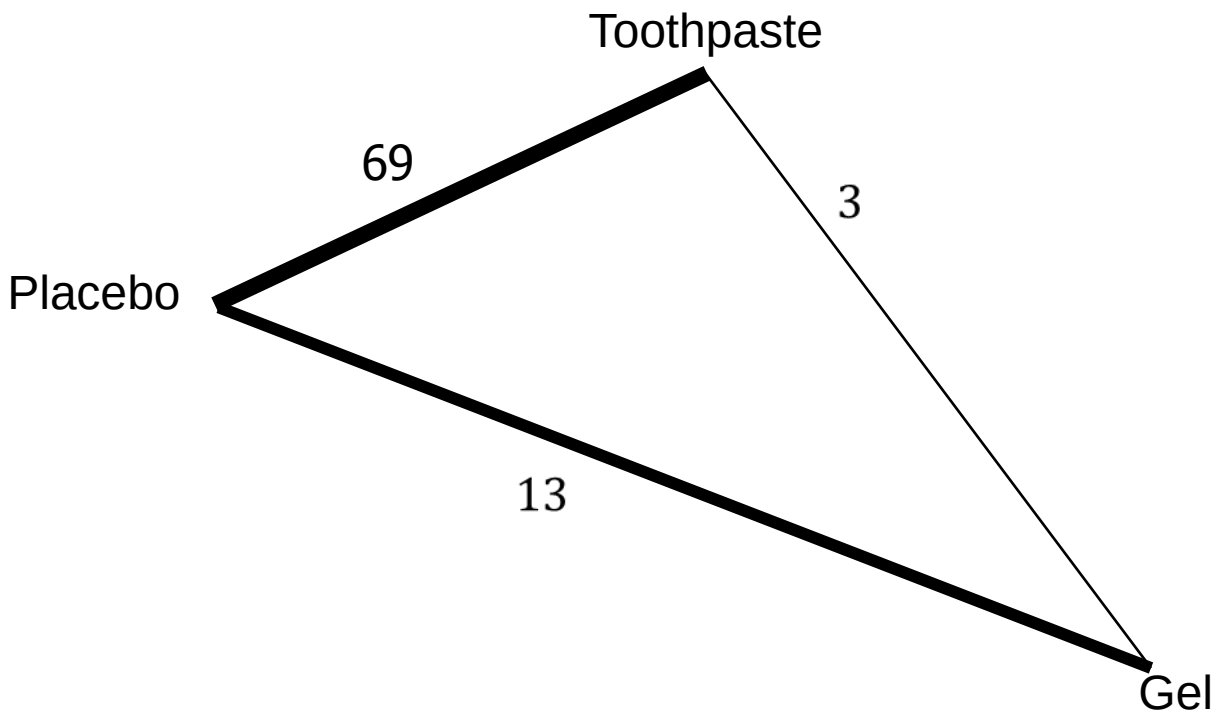
Testable assumption



Direct and indirect evidence are in agreement

If all three A, B and C are transitive then the loop is consistent

Consistency Equation



Inconsistency Factor

Indirect $SMD_{GvsT}^{ind} = -0.15$

$var\left(SMD_{GvsT}^{ind}\right) = 0.004$

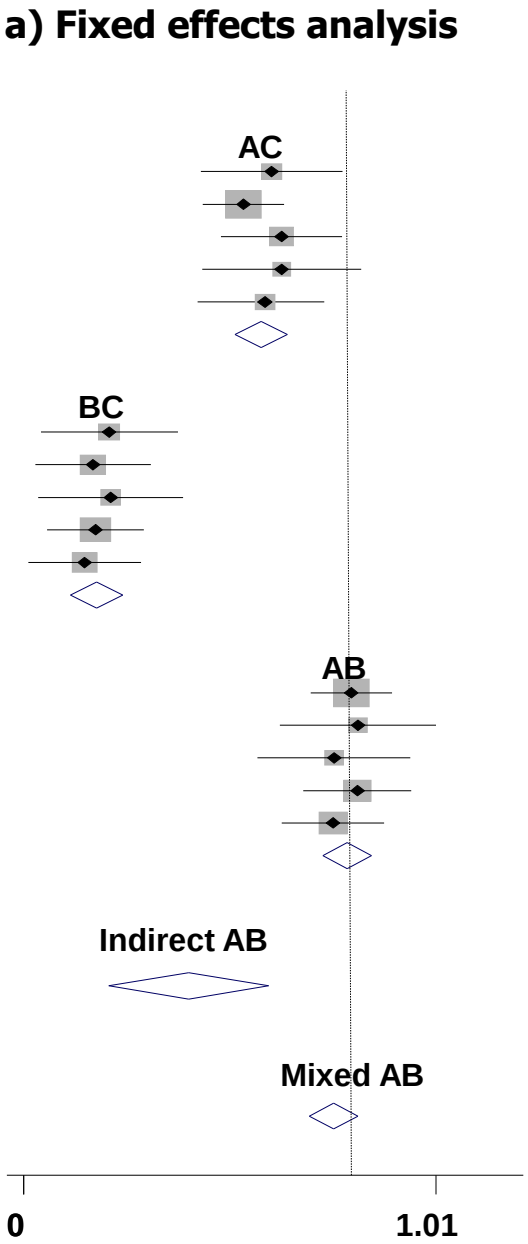
Direct $SMD_{GvsT}^{dir} = 0.04$

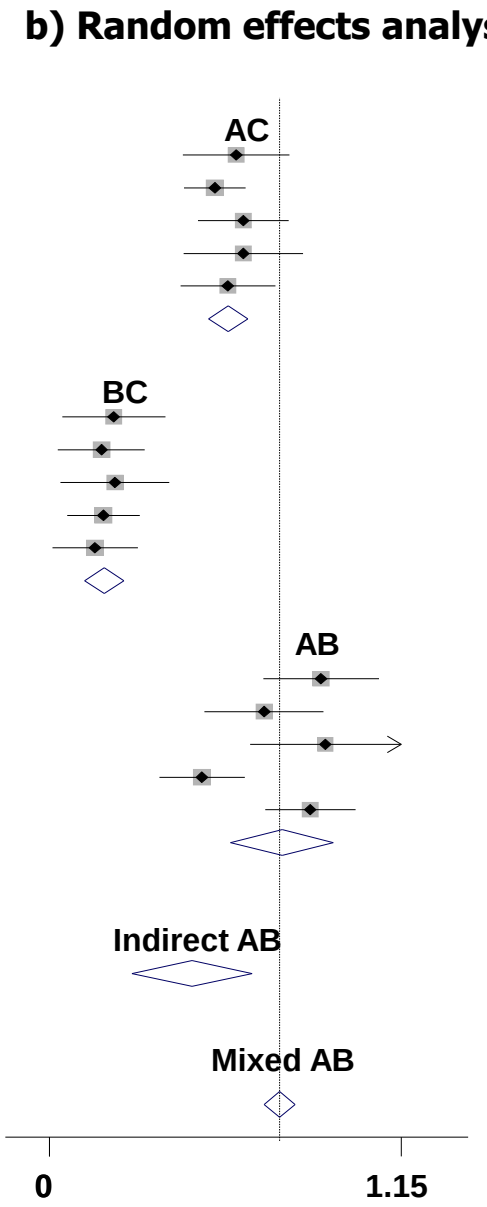
$var\left(SMD_{GvsT}^{dir}\right) = 0.011$

$$IF = \left|SMD_{GvsT}^{dir} - SMD_{GvsT}^{ind}\right|$$
$$= |0.04 - (-0.15)| = 0.19$$

You can do this with any measure... lnOR, lnRR, RD, mean difference, HR e.t.c

Consistency and heterogeneity



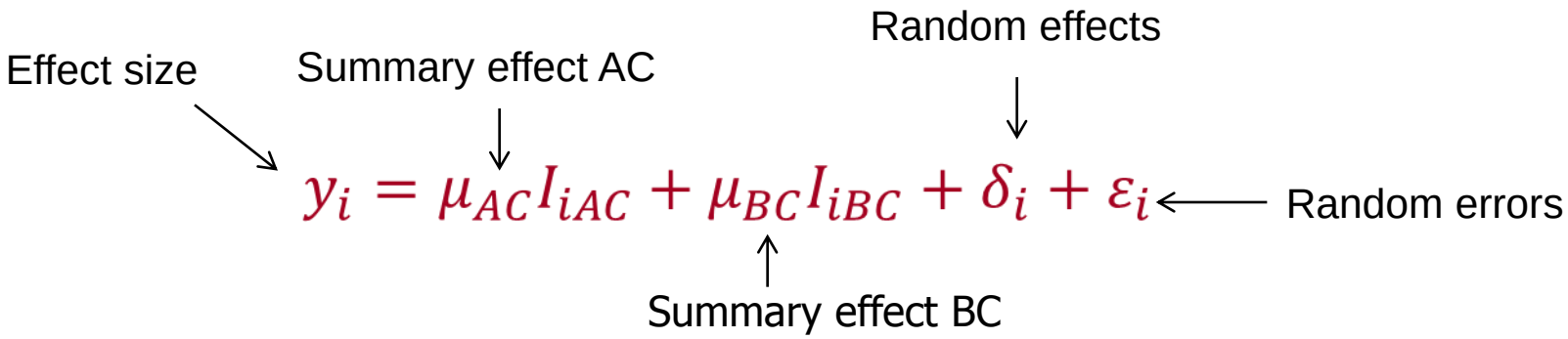


Fit a network meta-analysis model

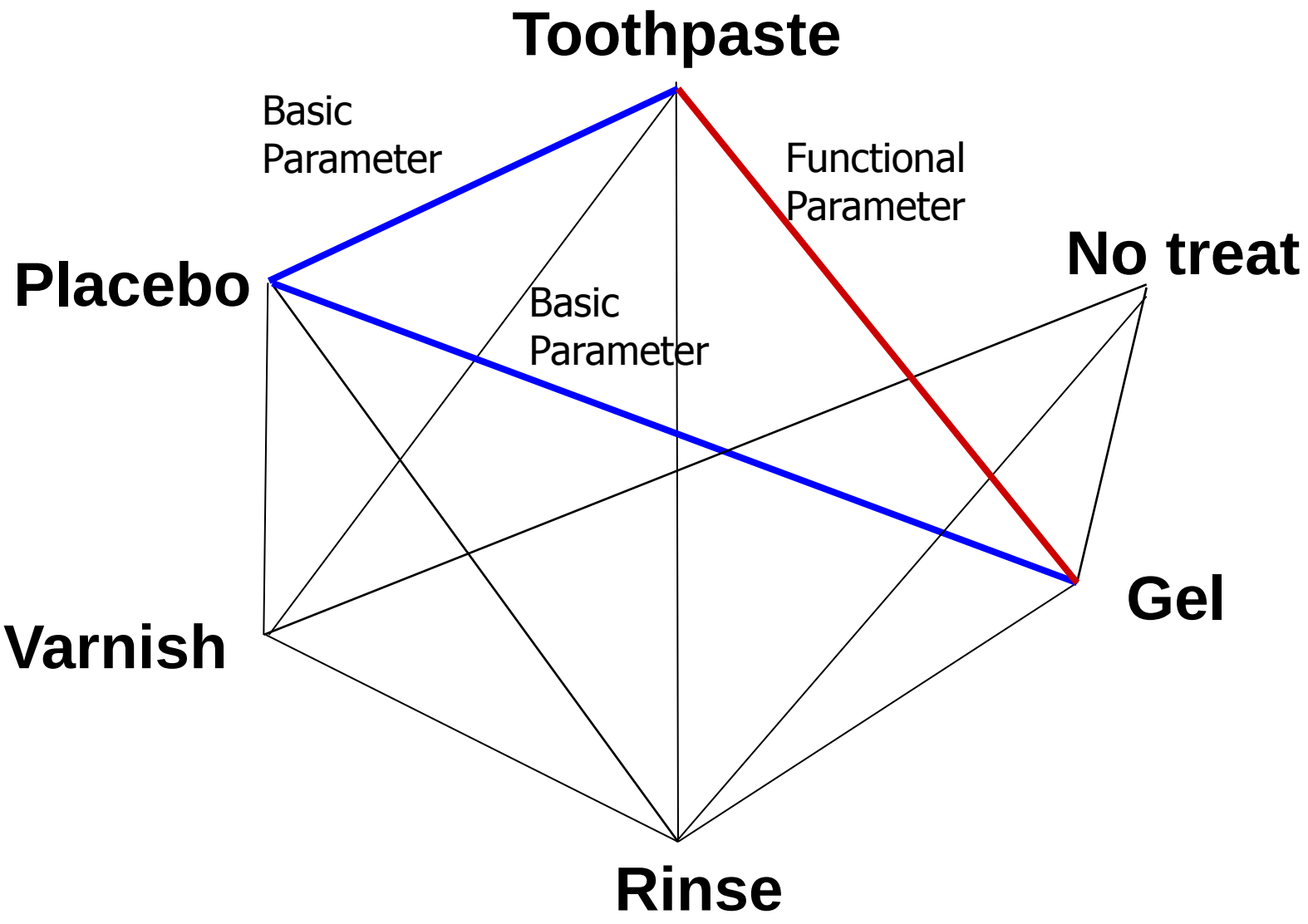
- Meta-analysis is a weighted regression with no covariates
- Network meta-analysis is a weighted regression with dummy variables for the treatments
- You should take into account correlations in multi-arm trials

Network and meta-regression

- Meta-regression using the treatments as ‘covariates’ and without intercept
- With 3 treatments and AC, AB, BC studies, chose C as *reference,* so *AC and BC are basic parameters*

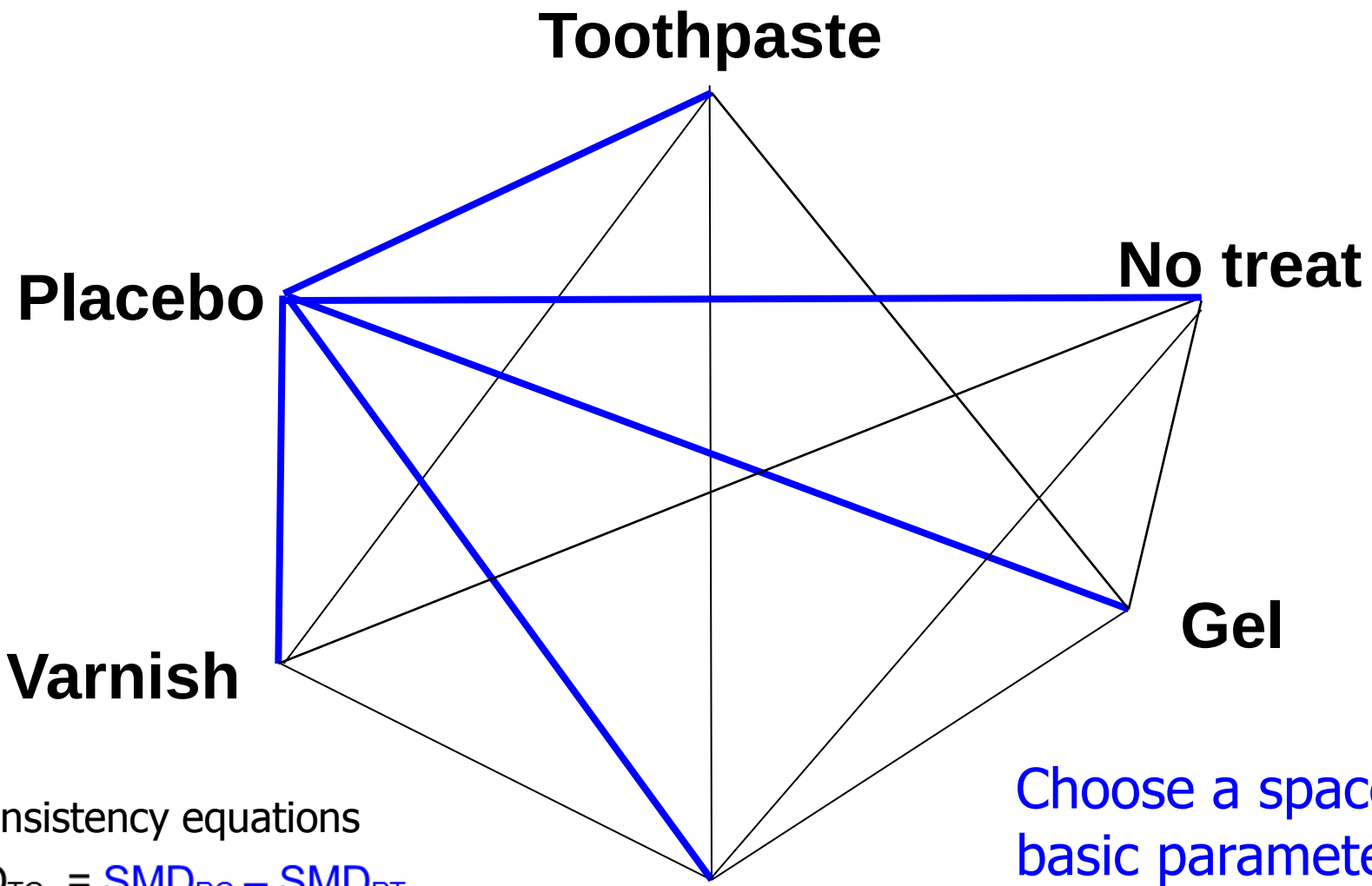


- The AC studies have $(I_{iAC}, I_{iBC}) = (1,0)$, the BC studies $(I_{iAC}, I_{iBC}) = (0,1)$[*basic*]
- AB studies have $(I_{iAC}, I_{iBC}) = (1,-1)$ [*functional*] because AB=AC-BC



$$SMD_{TG} = SMD_{PG} - SMD_{PT}$$

Consistency equation



Consistency equations

$SMD_{TG} = SMD_{PG} - SMD_{PT}$
$SMD_{TR} = SMD_{PR} - SMD_{PT}$
$SMD_{TV} = SMD_{PV} - SMD_{PT}$
$SMD_{RG} = SMD_{PG} - SMD_{PR}$

⋮

Choose a space of basic parameters

How many basic parameters?

$T - 1$

$$y_i = \mu^{PT} T_i + \mu^{PG} G_i + \mu^{PR} R_i + \mu^{PV} V_i + \mu^{PN} N_i$$

Use as ‘covariates’

No. studies	Placebo	Toothpaste	Gel	Rinse	Varnish	NoTreatment
69	-1	1		0	0	0
13	-1	0	1	0	0	0
31	-1	0	0	1	0	0
3	-1	0	0	0	1	0
4	0	0	0	-1	0	1
4	0	0	0	0	-1	1
9	0	0	-1	0	0	1
4	0	0	0	-1	1	0
6	0	-1	0	1	0	0

$$\mathbf{y} = X(\mu^{PT}, \mu^{PG}, \mu^{PR}, \mu^{PV}, \mu^{PN})' + \boldsymbol{\delta} + \varepsilon$$

Matrix of all observations

Design matrix

Vector of summary effects

Random effects

Random errors

$$\varepsilon \sim N(0, diag(v_i))$$

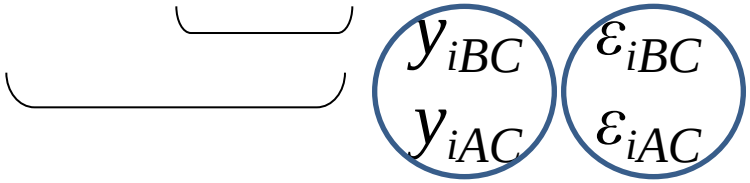
Variances matrix (for the observed SMD)

$$\boldsymbol{\delta} \sim N(\mathbf{0}, diag(\tau^2))$$

We assume equal heterogeneities for all comparisons

What's the problem with multi-arm trials?

- We need to take into account the correlations between the estimates that come from the same study
- A B C



- The random effects that refer to the same trial <u>are correlated</u> as well
- You have to built in *the correlation matrix for the observed effects*, **and** *the correlation matrix for the random effects*

$$\varepsilon \sim N(\mathbf{0}, S) \qquad \boldsymbol{\delta} \sim N(\mathbf{0}, \Delta)$$

Study	No. arms	#	Data	Contrast
i=1	T_1=2	1	$y_{1,1}$, $v_{1,1}$	AB
i=2	T_2=2	1	$y_{2,1}$, $v_{2,1}$	AC
i=3	T_3=2	1	$y_{3,1}$, $v_{3,1}$	BC
i=4	T_4=3	2	$y_{4,1}$, $v_{4,1}$ $y_{4,2}$, $v_{4,2}$ cov($y_{4,1}$, $y_{4,2}$)	AB AC

Meta-regression

$$\begin{pmatrix} y_{1,1} \\ y_{2,1} \\ y_{3,1} \\ y_{4,1} \\ y_{4,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{AB} \\ \mu_{AC} \end{pmatrix} + \begin{pmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \delta_{3,1} \\ \delta_{4,1} \\ \delta_{4,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \\ \varepsilon_{3,1} \\ \varepsilon_{4,1} \\ \varepsilon_{4,2} \end{pmatrix}$$

Study	No. arms	#	Data	Contrast
i=1	$T_1=2$	1	$y_{1,1}, v_{1,1}$	AB
i=2	$T_2=2$	1	$y_{2,1}, v_{2,1}$	AC
i=3	$T_3=2$	1	$y_{3,1}, v_{3,1}$	BC
i=4	$T_4=3$	2	$y_{4,1}, v_{4,1}$ $y_{4,2}, v_{4,2}$ $\text{cov}(y_{4,1}, y_{4,2})$	AB AC

$$\begin{pmatrix} y_{1,1} \\ y_{2,1} \\ y_{3,1} \\ y_{4,1} \\ y_{4,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{AB} \\ \mu_{AC} \end{pmatrix} + \begin{pmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \delta_{3,1} \\ \delta_{4,1} \\ \delta_{4,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \\ \varepsilon_{3,1} \\ \varepsilon_{4,1} \\ \varepsilon_{4,2} \end{pmatrix}$$

Take into account correlation in observations

$$\begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \\ \varepsilon_{3,1} \\ \varepsilon_{4,1} \\ \varepsilon_{4,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} v_{1,1} & 0 & 0 & 0 & 0 \\ 0 & v_{2,1} & 0 & 0 & 0 \\ 0 & 0 & v_{3,1} & 0 & 0 \\ 0 & 0 & 0 & v_{4,1} & \text{cov}\left(y_{4,1}, y_{4,2}\right) \\ 0 & 0 & 0 & \text{cov}\left(y_{4,1}, y_{4,2}\right) & v_{4,2} \end{pmatrix} \right)$$

Study	No. arms	#	Data	Contrast
i=1	$T_1=2$	1	$y_{1,1}, v_{1,1}$	AB
i=2	$T_2=2$	1	$y_{2,1}, v_{2,1}$	AC
i=3	$T_3=2$	1	$y_{3,1}, v_{3,1}$	BC
i=4	$T_4=3$	2	$y_{4,1}, v_{4,1}$ $y_{4,2}, v_{4,2}$ $\text{cov}(y_{4,1}, y_{4,2})$	AB AC

$$\begin{pmatrix} y_{1,1} \\ y_{2,1} \\ y_{3,1} \\ y_{4,1} \\ y_{4,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{AB} \\ \mu_{AC} \end{pmatrix} + \begin{pmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \delta_{3,1} \\ \delta_{4,1} \\ \delta_{4,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{2,1} \\ \varepsilon_{3,1} \\ \varepsilon_{4,1} \\ \varepsilon_{4,2} \end{pmatrix}$$

Take into account correlation in random effects

$$\begin{pmatrix} \delta_{1,1} \\ \delta_{2,1} \\ \delta_{3,1} \\ \delta_{4,1} \\ \delta_{4,2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau^2_{AB} & 0 & 0 & 0 & 0 \\ 0 & \tau^2_{AC} & 0 & 0 & 0 \\ 0 & 0 & \tau^2_{BC} & 0 & 0 \\ 0 & 0 & 0 & \tau^2_{AB} & \text{cov}(\delta_{4,1}, \delta_{4,2}) \\ 0 & 0 & 0 & \text{cov}(\delta_{4,1}, \delta_{4,2}) & \tau^2_{AC} \end{pmatrix} \right)$$

Multivariate meta-analysis

- Studies typically report many outcomes
  - Efficacy and acceptability in antidepressants
- **Multivariate meta-analysis** allows a joint synthesis of the multiple end points
- Different between-treatment contrasts are viewed as different outcomes
- White et al estimate NMA models by expressing them as multivariate random-effects meta-regressions (mvmeta in STATA)

Data: n studies with 2 outcomes

Efficacy/AB

Acceptability/AC

$$\text{Study } 1: y_{11}, y_{12}, \begin{pmatrix} s_{11}^2 & S_{112} \\ S_{112} & s_{12}^2 \end{pmatrix}$$

$$\text{Study } i: y_{i1}, y_{i2}, \begin{pmatrix} s_{i1}^2 & S_{i12} \\ S_{i12} & s_{i2}^2 \end{pmatrix} \qquad S_{i12} = \rho_i s_{\iota 1} s_{\iota 2}$$

$$\text{Study } n: y_{n1}, y_{n2}, \begin{pmatrix} s_{n1}^2 & S_{n12} \\ S_{n12} & s_{n2}^2 \end{pmatrix}$$

Network meta-analysis and multivariate approaches

- We can look at network meta-analysis as either a multivariate meta-regression or a multivariate meta-analysis
- Multivariate meta-regression:
  - extends the meta-regression approach to allow for multi-arm trials
  - dummy 1, -1 and 0 codes for treatments (with a reference in mind)
  - assumes a common heterogeneity variance
- Multivariate meta-analysis:
  - no covariates required
  - Flexible modelling of the between-study variance matrix
  - requires a common reference arm for every study
    - a problem that is surmountable using *data augmentation*

How to fit network meta-analysis?

- R mvmeta, metasem, netmeta
- STATA using metareg (no multi-arm studies)
- STATA mvmeta
- To my knowledge only netmeta in R and mvmeta in STATA model properly the matrix Δ
- Using MCMC (WinBUGS)

Presenting results from network meta-analysis

- With many treatments judgments based on pairwise effect sizes are difficult to make

Antidepressants

Efficacy (response rate) (95% CI) | Comparison | Acceptability (dropout rate) (95% CI)

BUP	1·00 (0·78–1·28)	0·75 (0·55–1·01)	1·06 (0·86–1·32)	0·89 (0·74–1·08)	0·73 (0·53–1·00)	0·87 (0·58–1·24)	0·87 (0·66–1·14)	0·81 (0·65–1·00)	**0·62** (0·45–0·86)	1·01 (0·82–1·27)	0·84 (0·68–1·02)
0·98 (0·78–1·23)	**CIT**	0·75 (0·55–1·02)	1·07 (0·86–1·31)	0·90 (0·73–1·09)	**0·73** (0·54–0·99)	0·87 (0·60–1·24)	0·87 (0·66–1·15)	0·81 (0·65–1·01)	**0·62** (0·45–0·84)	1·02 (0·81–1·28)	0·84 (0·67–1·06)
1·09 (0·83–1·43)	1·12 (0·87–1·44)	**DUL**	**1·43** (1·09–1·85)	1·19 (0·91–1·57)	0·98 (0·67–1·41)	1·16 (0·77–1·73)	1·16 (0·83–1·61)	1·08 (0·84–1·40)	0·83 (0·57–1·22)	**1·36** (1·01–1·83)	1·12 (0·84–1·50)
0·82 (0·67–1·01)	0·84 (0·70–1·01)	**0·75** (0·60–0·93)	**ESC**	0·84 (0·70–1·01)	**0·69** (0·50–0·94)	0·81 (0·55–1·15)	0·81 (0·62–1·07)	**0·76** (0·62–0·93)	**0·58** (0·43–0·81)	0·95 (0·77–1·19)	**0·78** (0·64–0·97)
1·08 (0·90–1·29)	1·10 (0·93–1·31)	0·99 (0·79–1·24)	**1·32** (1·12–1·55)	**FLU**	0·82 (0·62–1·07)	0·97 (0·69–1·32)	0·97 (0·77–1·21)	0·91 (0·79–1·05)	**0·70** (0·53–0·92)	1·14 (0·96–1·36)	0·94 (0·81–1·09)
1·10 (0·83–1·47)	1·13 (0·86–1·47)	1·01 (0·74–1·38)	**1·35** (1·02–1·76)	1·02 (0·81–1·30)	**FVX**	1·18 (0·76–1·75)	1·18 (0·87–1·61)	1·10 (0·84–1·47)	0·85 (0·57–1·26)	**1·38** (1·03–1·89)	1·14 (0·86–1·54)
1·07 (0·77–1·48)	1·09 (0·78–1·50)	0·97 (0·69–1·38)	1·30 (0·95–1·78)	0·99 (0·74–1·31)	0·97 (0·68–1·37)	**MIL**	0·99 (0·69–1·53)	0·94 (0·68–1·31)	0·72 (0·48–1·10)	1·17 (0·84–1·72)	0·97 (0·69–1·40)
0·79 (0·72–1·00)	0·80 (0·63–1·01)	**0·72** (0·54–0·94)	0·96 (0·76–1·19)	**0·73** (0·60–0·88)	**0·71** (0·55–0·92)	0·74 (0·53–1·01)	**MIR**	0·93 (0·75–1·17)	0·72 (0·51–1·03)	1·17 (0·91–1·51)	0·97 (0·76–1·23)
1·06 (0·87–1·30)	1·08 (0·90–1·30)	0·97 (0·78–1·20)	**1·30** (1·10–1·53)	0·98 (0·86–1·12)	0·96 (0·76–1·23)	1·00 (0·74–1·33)	**1·35** (1·11–1·64)	**PAR**	0·77 (0·56–1·05)	**1·25** (1·04–1·52)	1·03 (0·86–1·24)
1·60 (1·20–2·16)	**1·63** (1·25–2·14)	**1·46** (1·05–2·02)	**1·95** (1·47–2·59)	**1·48** (1·16–1·90)	**1·45** (1·03–2·02)	**1·50** (1·03–2·18)	**2·03** (1·52–2·78)	**1·50** (1·16–1·98)	**REB**	**1·63** (1·19–2·24)	1·34 (0·99–1·83)
0·87 (0·72–1·05)	0·88 (0·72–1·07)	0·79 (0·62–1·01)	1·06 (0·88–1·27)	**0·80** (0·69–0·93)	0·79 (0·61–1·01)	0·81 (0·60–1·11)	1·10 (0·90–1·36)	**0·82** (0·69–0·96)	**0·54** (0·41–0·71)	**SER**	0·82 (0·67–1·00)
0·85 (0·70–1·01)	0·86 (0·71–1·05)	**0·77** (0·60–0·99)	1·03 (0·86–1·24)	**0·78** (0·68–0·90)	**0·77** (0·59–0·99)	0·79 (0·58–1·08)	1·08 (0·87–1·33)	**0·79** (0·67–0·94)	**0·53** (0·40–0·69)	0·98 (0·82–1·16)	**VEN**

OR>1 means the treatment in top-left is better

Probabilities

- Estimate for each treatment ***the probability of being the best***
- Rankings are contructed by drawing the coefficients a large number of times from their approximate posterior density
- For each draw, the effect sizes are estimated and the largest effect size is noted

12 new generation antidepressants

paroxetine	0%
sertraline	7%
citalopram	0%
escitalopram	26%
fluoxetine	0%
fluvoxamine	0%
milnacipran	1%
venlafaxine	11%
reboxetine	0%
bupropion	0%
mirtazapine	54%
duloxetine	0%

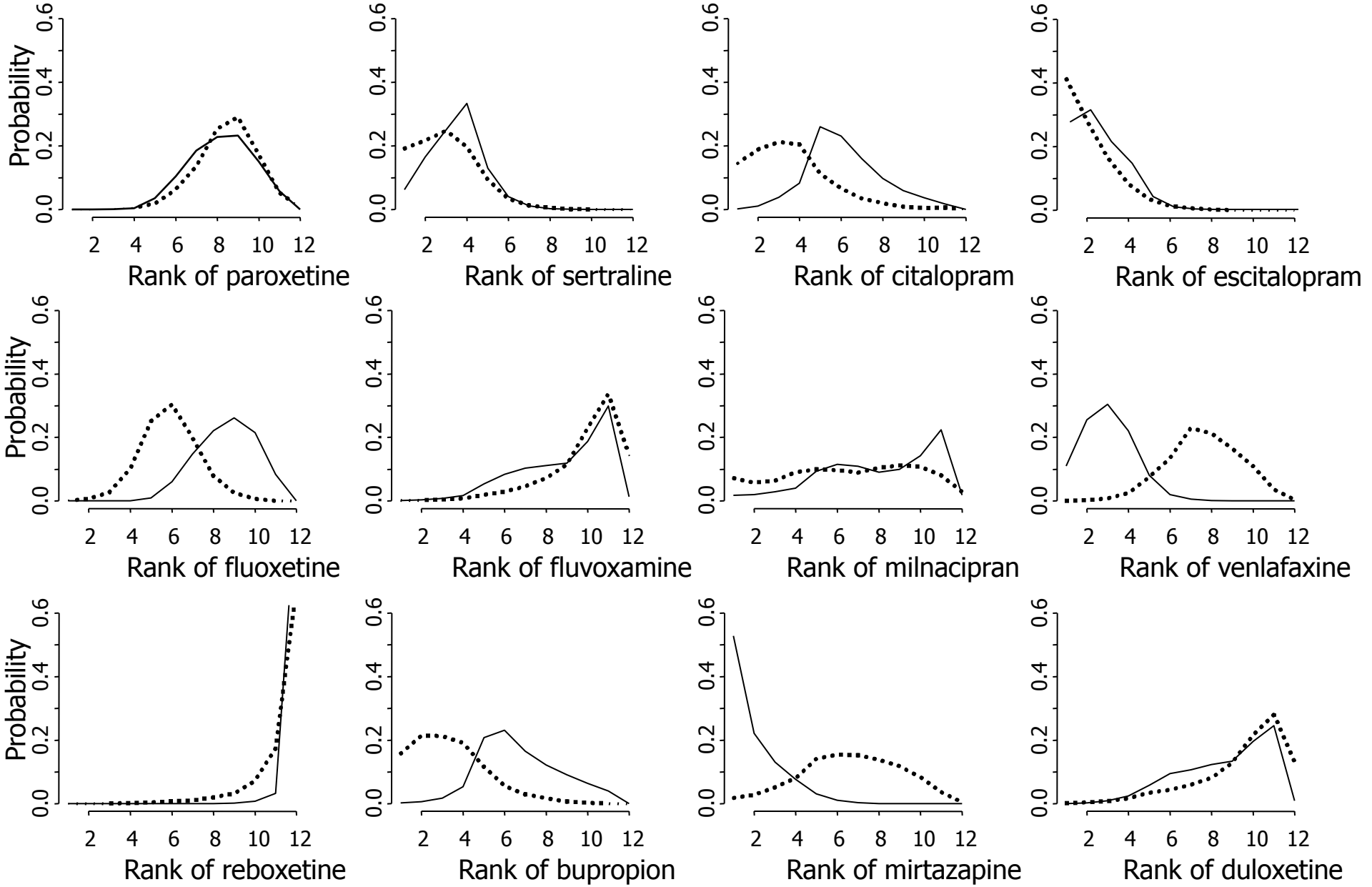
Probability of being the best

The probability of being the best does not convey the spread of the rank probabilities....

% probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00

% probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00
j=2	0.25	0.25	0.50	0.00
j=3	0.25	0.25	0.25	0.25
j=4	0.25	0	0	0.75

i the treatment
j the rank



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

Surface under the cumulative ranking curve

Use posterior probabilities for each treatment to be among the *n* -best options

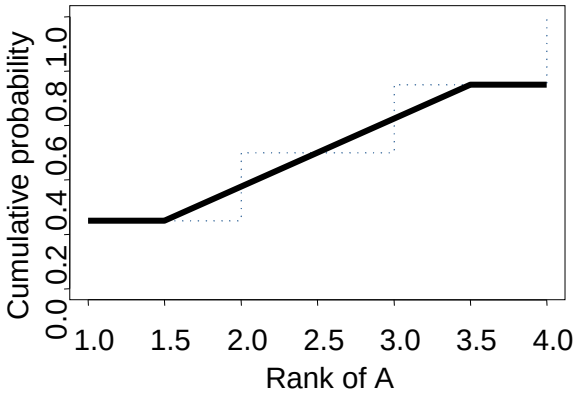
$$\text{Cumulative ranking curve}_i = \frac{\sum_{j=1}^{T-1} p_{ji}}{T-1}$$

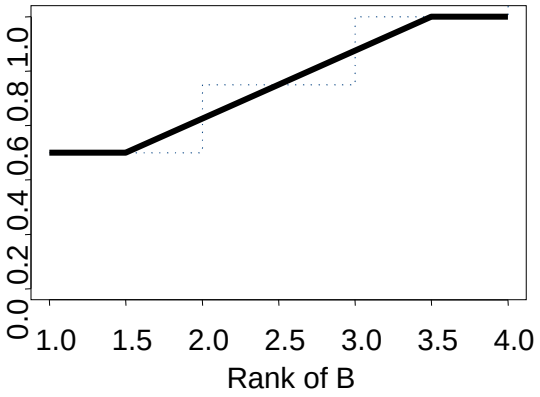
T Total number of treatments

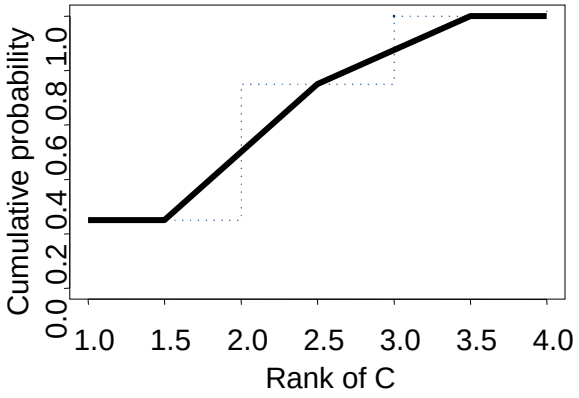
[J Clin Epidemiol. 2010 Salanti et al]

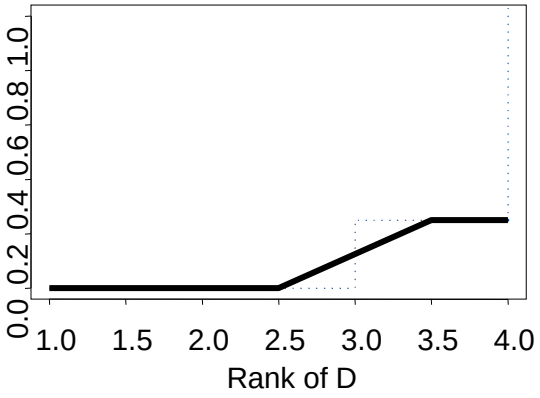
% probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00
j=2	0.50	0.75	0.75	0.00
j=3	0.75	1.00	1.00	0.25
j=4	1.00	1.00	1.00	1.00

i the treatment
j the rank



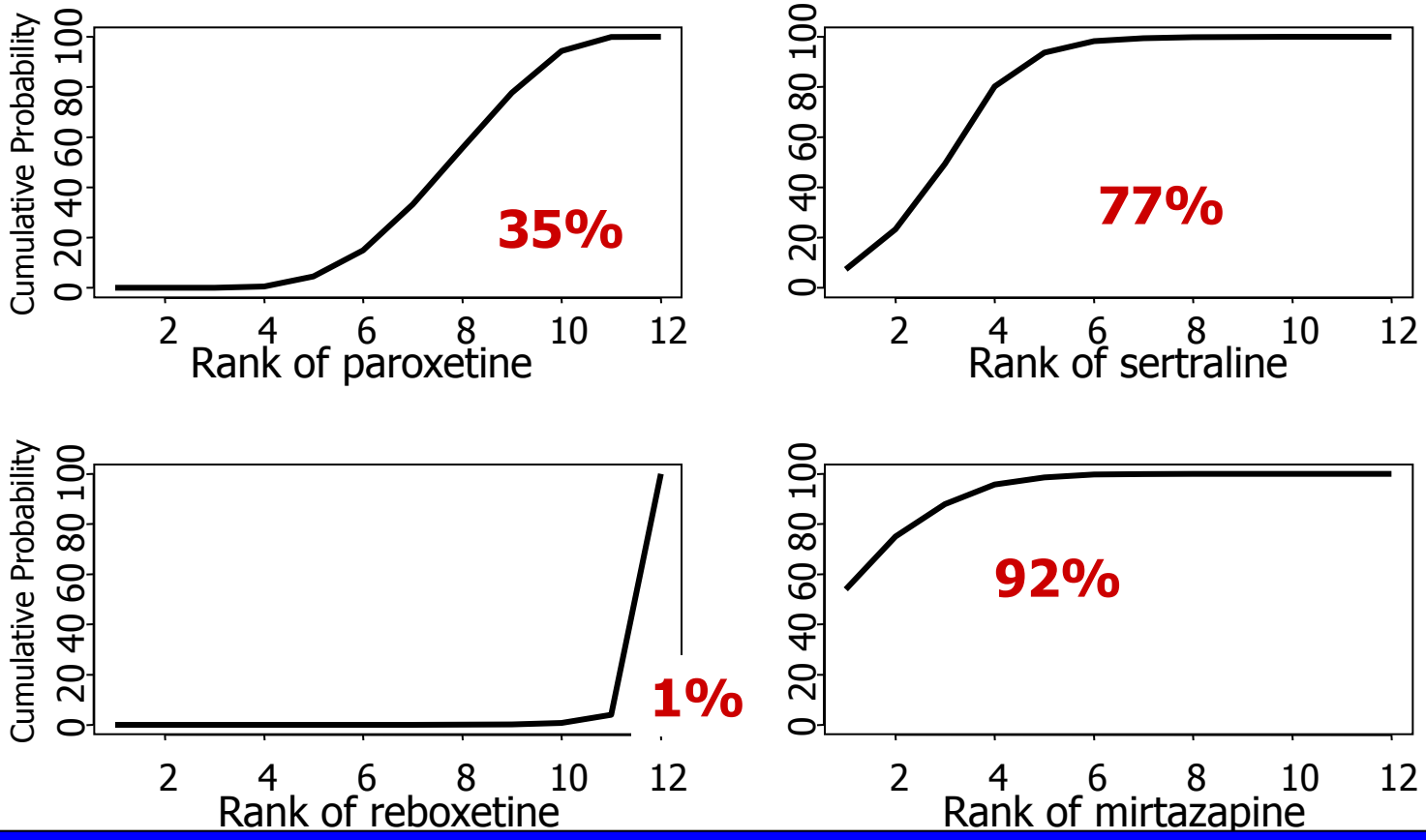






The areas under the cumulative curves for the four treatments of the example above are
A=0.5
B=0.75
C=0.67
D=0.08

Surface under the cumulative ranking curve



Warning: measures based on probabilities are attractive, but can be unstable and should be presented along with the effect sizes!

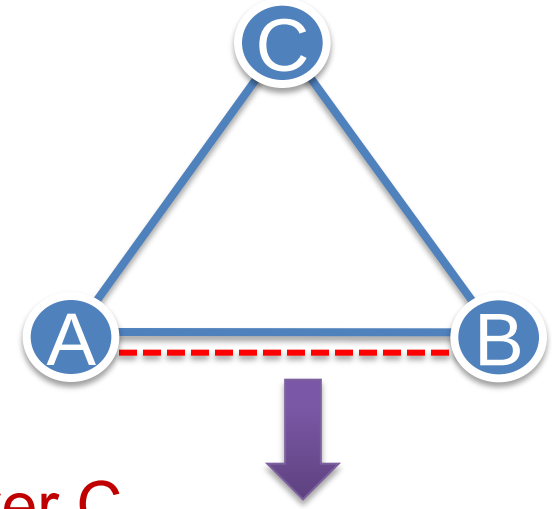
INCONSISTENCY

Validity of network meta-analysis

- The validity of a network meta-analysis depends on *transitivity* of effect size parameters:
- For any pair A and B,

  typical (or mean) advantage of A over B =

  advantage of A over C − advantage of B over C



Evaluate the assumption of consistency

- In a simple indirect comparison, we cannot test this assumption empirically.
- In a network meta-analysis, we sometimes can.
- We call this looking at *inconsistency*.

What is inconsistency?

- *Consistency* = The data fit together according to the laws of transitivity

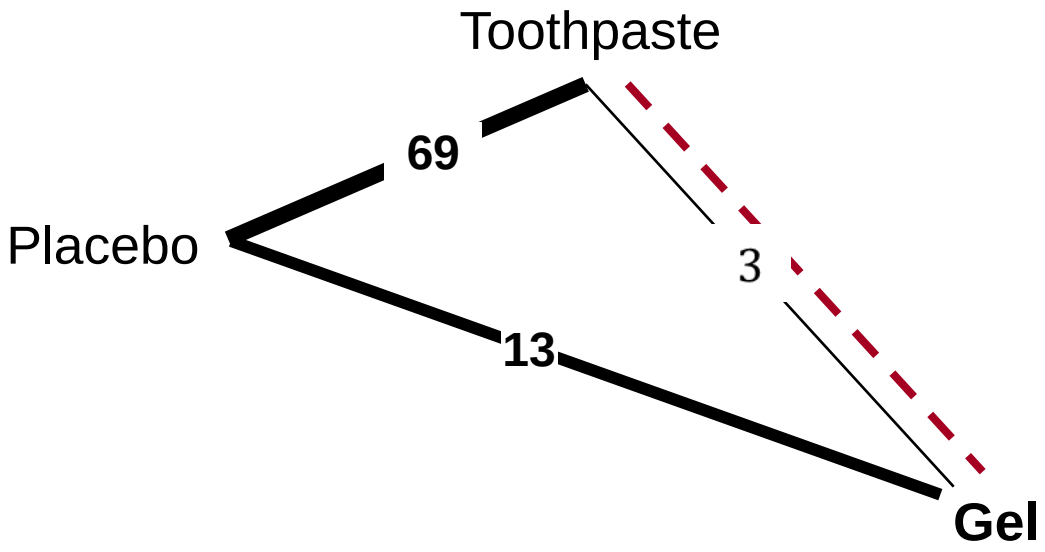
- i.e.
  - for each pair of interventions A and B, all sources of evidence about A vs B agree with each other
    - (this means direct evidence (if available) and different routes to indirect evidence)

- *Inconsistency* = Lack of consistency

- Only *closed loops* can tell us about (in)consistency

Example: a simple loop of treatments

Direct Evidence	Summary Treatment effect	Variance	Confidence Interval
SMD_{GvsT}	0.04	0.011	(–0.17, 0.25)



Indirect Evidence	Summary Treatment effect	Variance	Confidence Interval
SMD_{PvsT}	-0.34	0.002	(–0.43, –0.25)
SMD_{PvsG}	-0.19	0.002	(–0.28, –0.10)
Indirect comparison			
SMD_{GvsT_ind}	-0.15	0.004	(–0.27, –0.03)

How much inconsistency?

- Taking into account the previous evidence,
- the difference between direct and indirect estimates is

$$IF = \left|SMD_{GvsT}^{dir} - SMD_{GvsT}^{ind}\right| = |0.04 - (-0.15)| = 0.19$$

- and we add the variances (since the sources of evidence are independent):

Var(difference between direct and indirect) =

$$var(IF) = var\left(SMD_{GvsT}^{dir}\right) + var\left(SMD_{GvsT}^{ind}\right) = 0.004 + 0.011 = 0.015$$

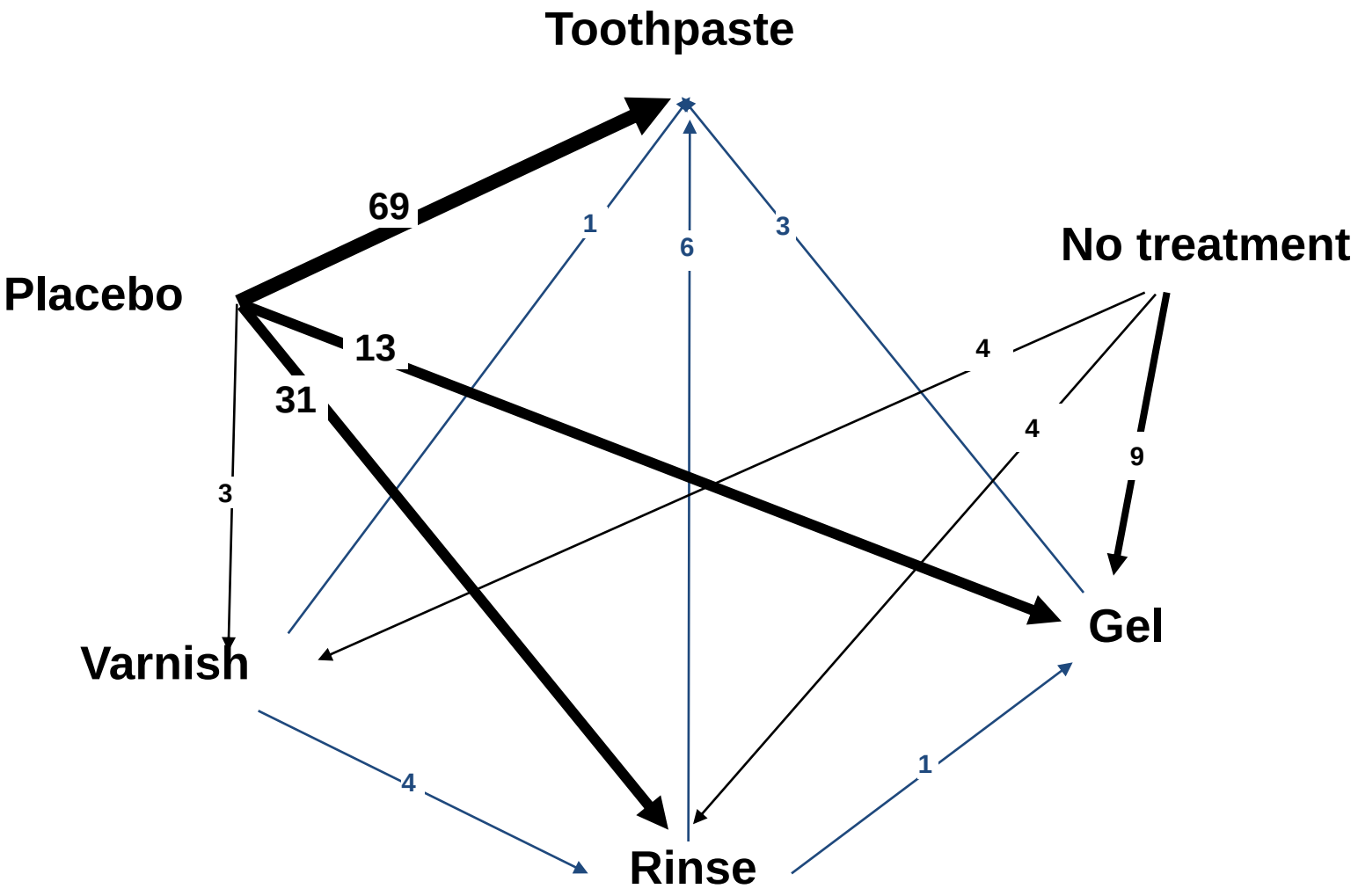
How much inconsistency?

$$Z = \frac{IF}{\sqrt{var(IF)}} \sim N(0,1)$$

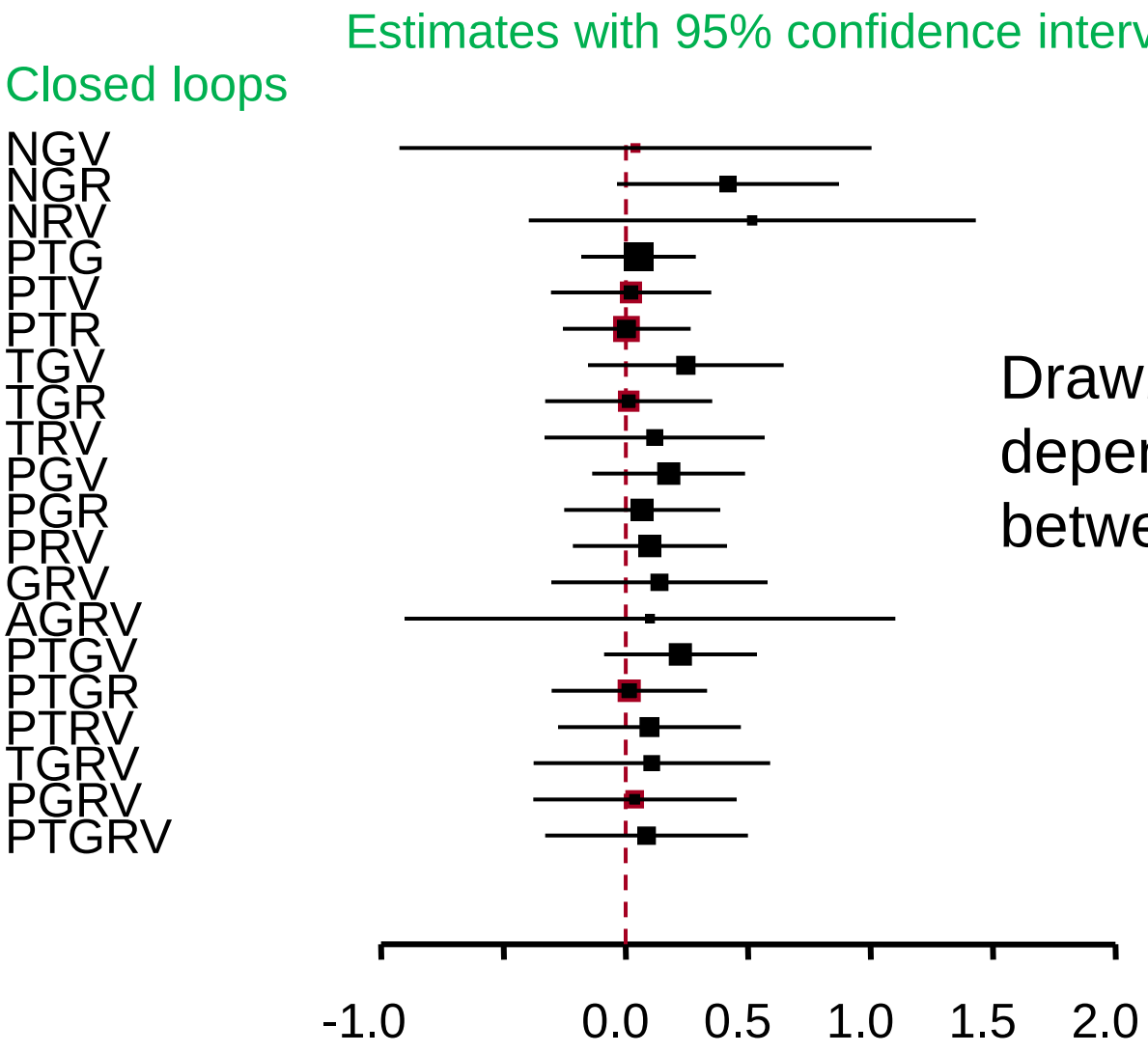
95% Confidence Interval for Inconsistency

$$IF \pm 1.96\sqrt{var(IF)}$$
$$0.19 \pm 1.96\sqrt{0.015}$$
$$0.19 \pm 0.24$$
$$(-0.05, 0.43)$$

Example: fluoride treatments



Evaluation of consistency within closed loops



Drawback: dependence between loops

R routine in www.mtm.uoi.gr/howotodoanmtm.html

[Clin Epidemiol 2009, Salanti et al]

Are networks typically inconsistent?

Triangular networks

- Song et *al* BMJ 2011 found 16/112 (14%) inconsistent triangles
  - The same authors evaluated the assumption of consistency in Cochrane Reviews separately (Xiong et *al* JCE 2013) and found 16/94 (17%) triangles inconsistent

Complex networks

- Veroniki et *al* (IJE 2013) published network meta-analyses with binary data that involve at least 4 treatments and at least one closed loop
  - so far 40 networks, 303 loops
  - Inconsistency was detected in between 2% and 10% of the tested loops, depending on the effect measure and heterogeneity estimation method
  - About one eighth of the networks was found to be inconsistent.

Approaches for exploring inconsistency

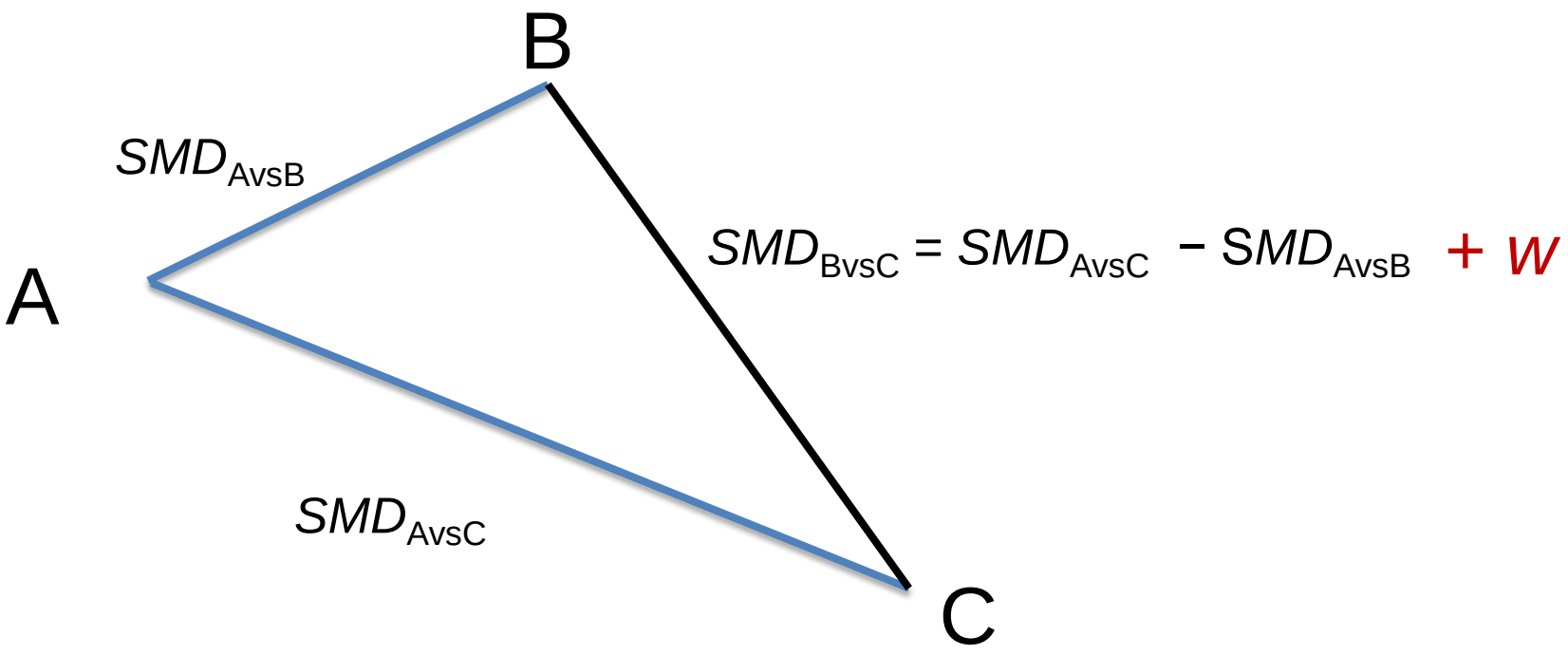
Evaluation of local inconsistency

- Loop-Specific : examine each closed loop separately
- Node-splitting (Dias et *al* Stat Med 2010)

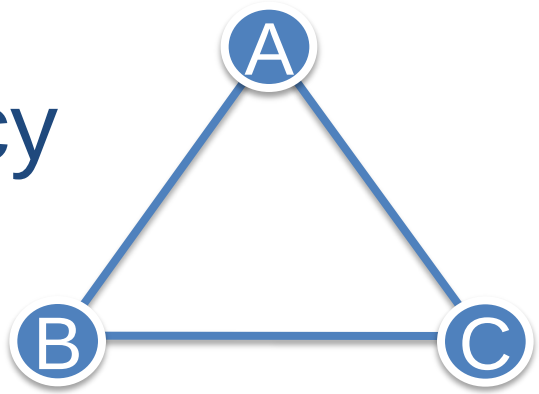
Evaluation of global inconsistency

- Use a network meta-analysis model that allows for inconsistency (Lu & Ades JASA 2005)
- Compare model fit between consistency and inconsistency models
- Apply a 'design by treatment' interaction model (White et *al* RSM 2012, Higgins et *al* RSM 2012)

Inconsistency models: introduction

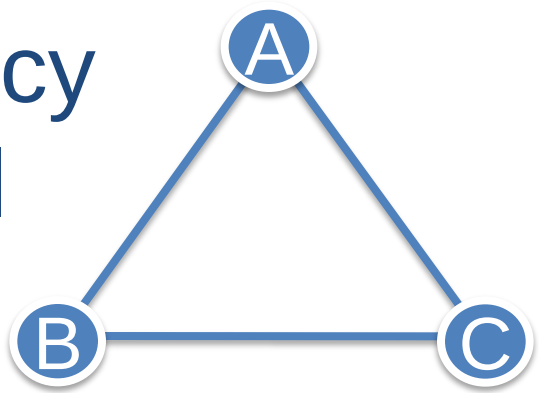


Model for consistency



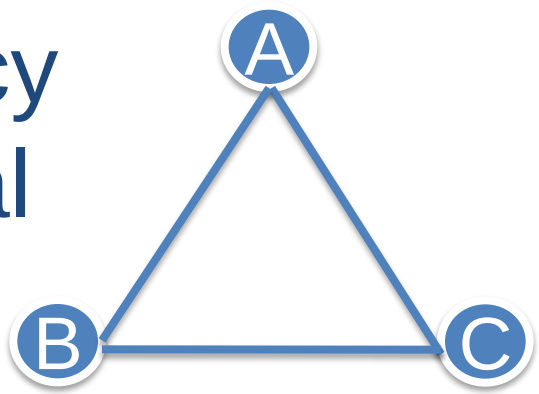
Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
AB	*ref*	$\mu_{AB} + \delta_i$	
AC	*ref*		$\mu_{AC} + \delta_i$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$

Model for inconsistency
Lu and Ades model



Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
AB	*ref*	$\mu_{AB} + \delta_i$	
AC	*ref*		$\mu_{AC} + \delta_i$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i + w$

Model for consistency with a three-arm trial



Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
ABC	*ref*	$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$
AB	*ref*	$\mu_{AB} + \delta_i$	
AC	*ref*		$\mu_{AC} + \delta_i$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$

Issues with the Lu and Ades model

- In the presence of multi-arm trials, the Lu and Ades inconsistency model is not uniquely defined

- Problems arise because multi-arm trials *must* be consistent, so a network with multi-arm trials will have a mixture of **consistent** and **inconsistent** loops

Lu and Ades model

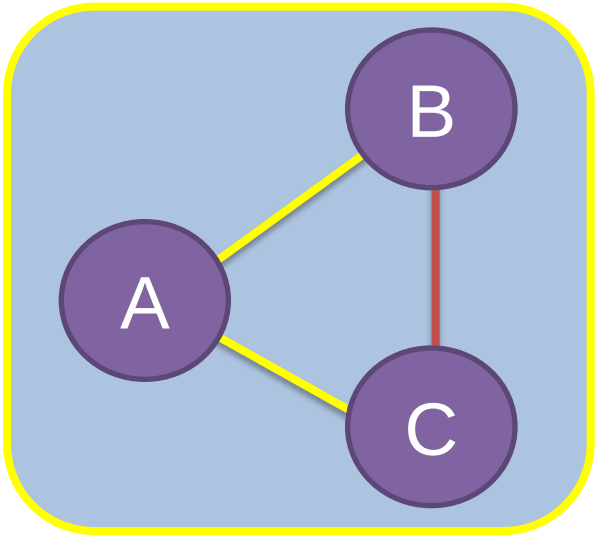
Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
ABC	*ref*	$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$
AB	*ref*	$\mu_{AB} + \delta_i$	
AC	*ref*		$\mu_{AC} + \delta_i$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i + w$

Design by treatment interaction model

- A model that is completely general is one that allows for all types of inconsistency
  - inconsistency within loops made up of different trials
  - inconsistency between two-arm and three-arm trials
  - and beyond...
- Such a model has been termed a design-by-treatment interaction model

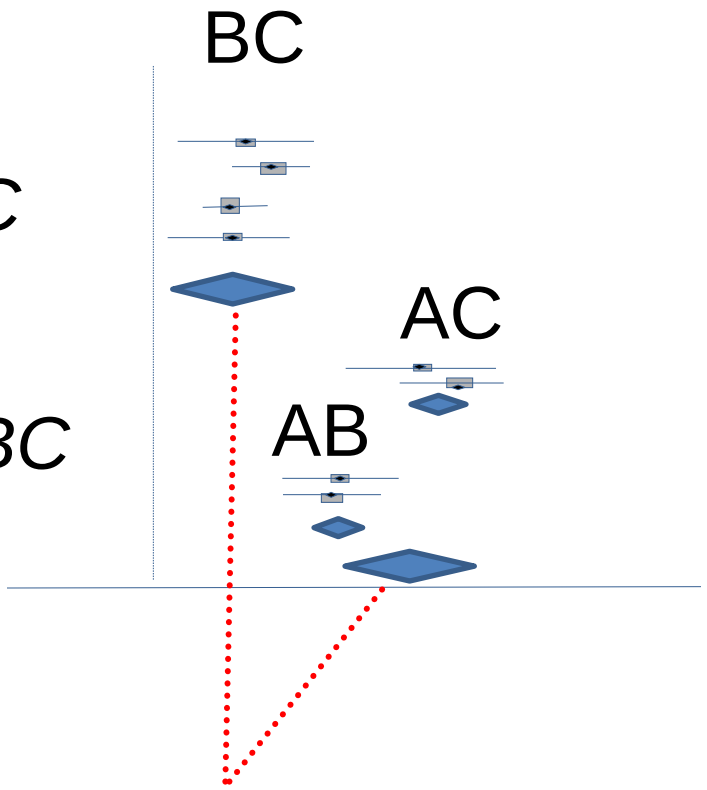
Forms of Inconsistency

Loop Inconsistency



Direct BC

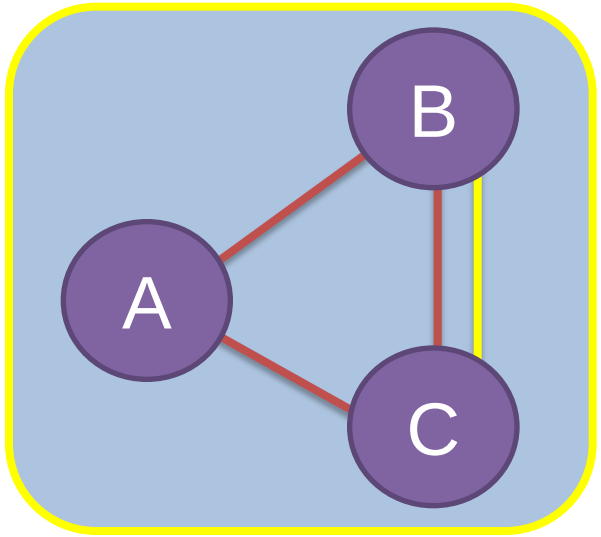
Indirect BC



If they statistically differ : *Inconsistency!*

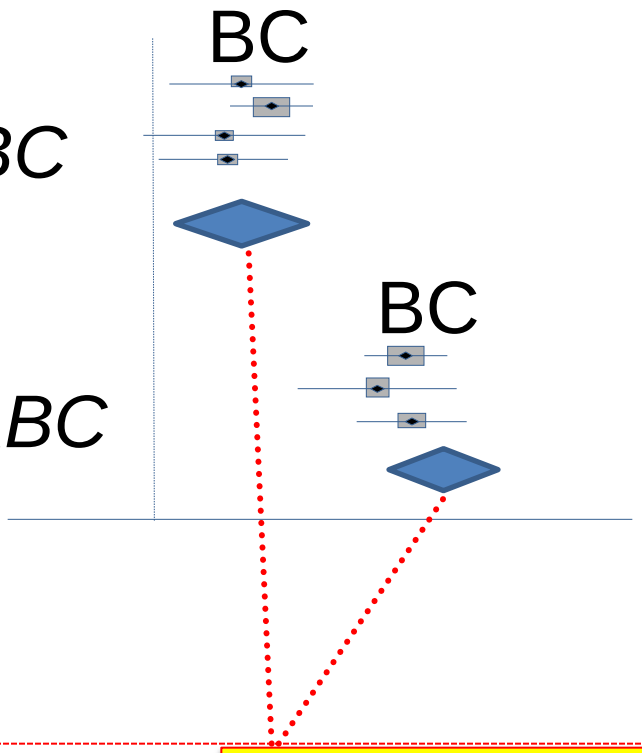
Forms of Inconsistency

Design Inconsistency



Design BC

Design ABC



If they statistically differ : *Inconsistency!*

Design-by-treatment interaction model

Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
ABC	*ref*	$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$
AB	*ref*	$\mu_{AB} + \delta_i + w_{AB}$	
AC	*ref*		$\mu_{AC} + \delta_i + w_{AC}$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i + w_{BC}$

[Higgins et *al* RSM 2012], [White et *al* RSM 2012]

Lu and Ades model for inconsistency with a three-arm trial

Modelled log odds ratios (basic parameters μ_{AB} and μ_{AC}); δ_i is the heterogeneity random effect			
Trial	*A*	*B*	*C*
ABC	*ref*	$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i$
AB	*ref*	$\mu_{AB} + \delta_i$	
AC	*ref*		$\mu_{AC} + \delta_i$
BC		$\mu_{AB} + \delta_i$	$\mu_{AC} + \delta_i + w$

Modelling the *w* parameters

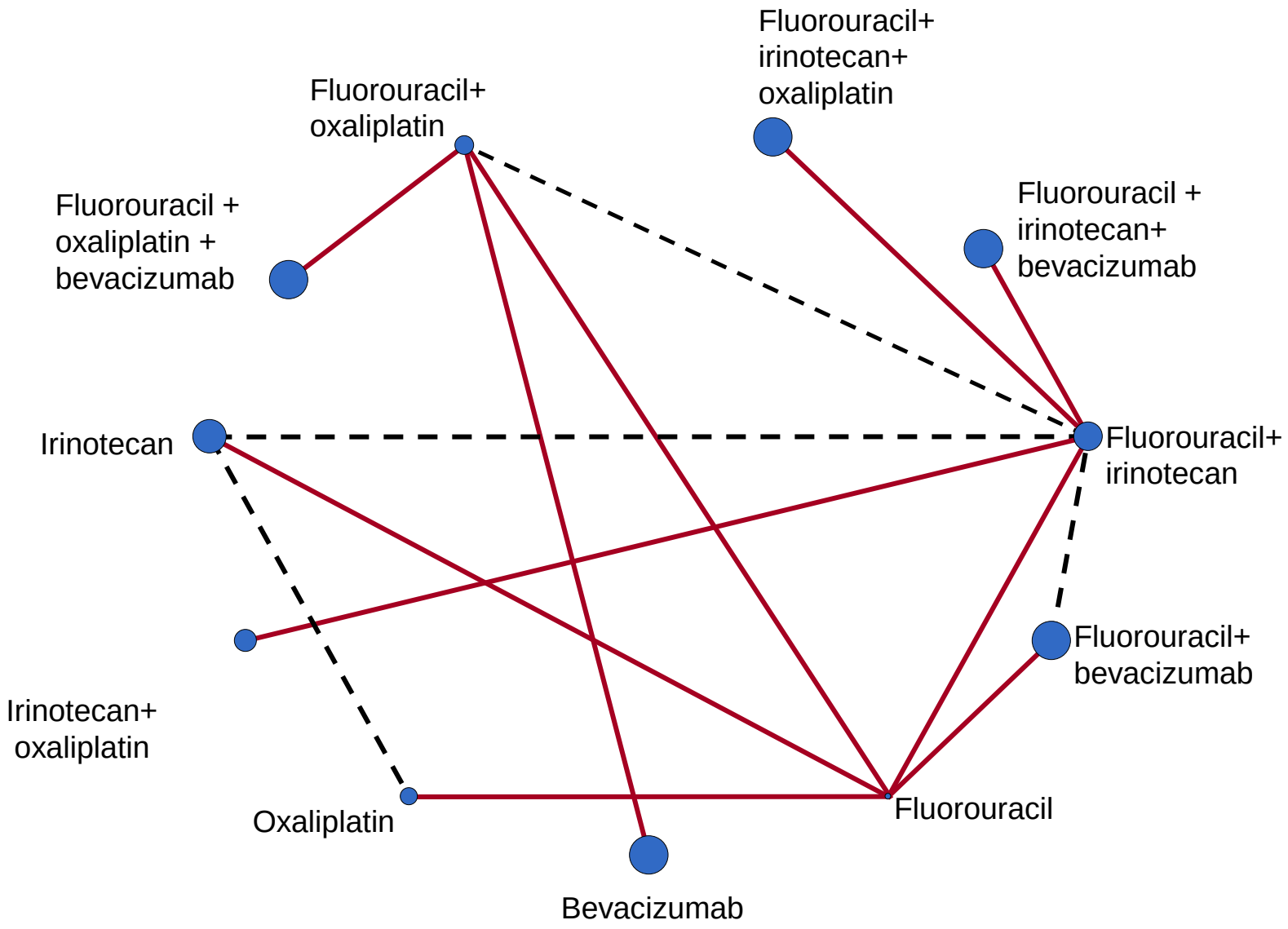
- When we have several inconsistency (*w*) parameters, we could let them have a random-effects distribution across comparisons

$$w_j \sim N(0,\sigma^2)$$

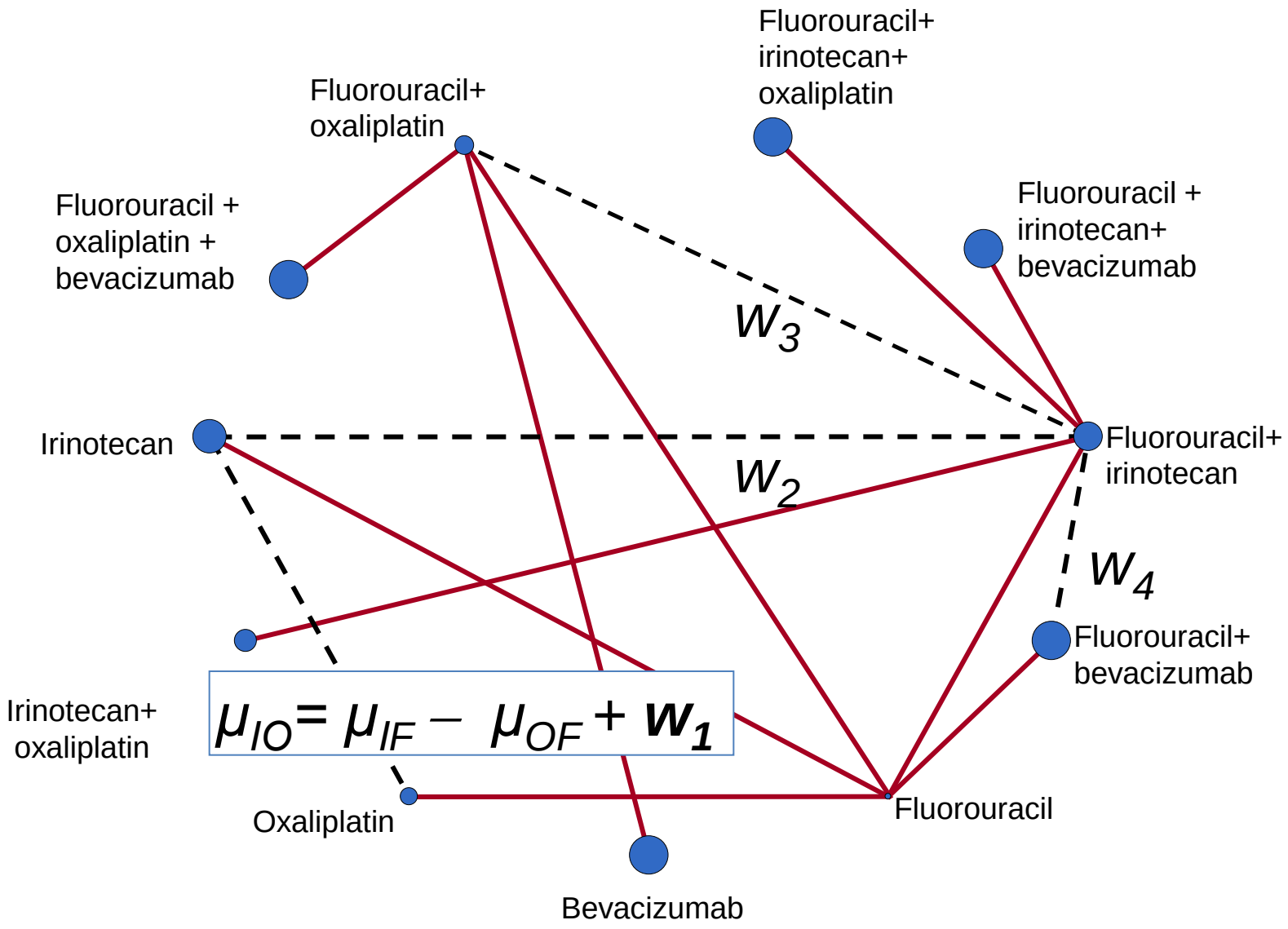
- Comparing σ^2 with τ^2 *(heterogeneity)* allows us to assess the magnitude of the inconsistency

- I prefer to use fixed effects for the *w* parameters
  - can interpret them individually
  - and it's easier to fit the model using Stata

Example: Survival with 11 chemotherapy regimens in colorectal cancer



Lu and Ades model for colorectal cancer



Results: colorectal cancer network

- $w_1 = -0.08$, $w_2 = -0.07$, $w_3 = -0.06$, $w_4 = -0.03$
  - No loop is remarkably inconsistent
- $\sigma^2 = 0.11$ (SD 0.04), $\tau^2 = 0.19$ (SD 0.18)
- $P(\sigma^2 > \tau^2) = 0.41$
  - No important changes in posterior HRs or fit

What if we find inconsistency?

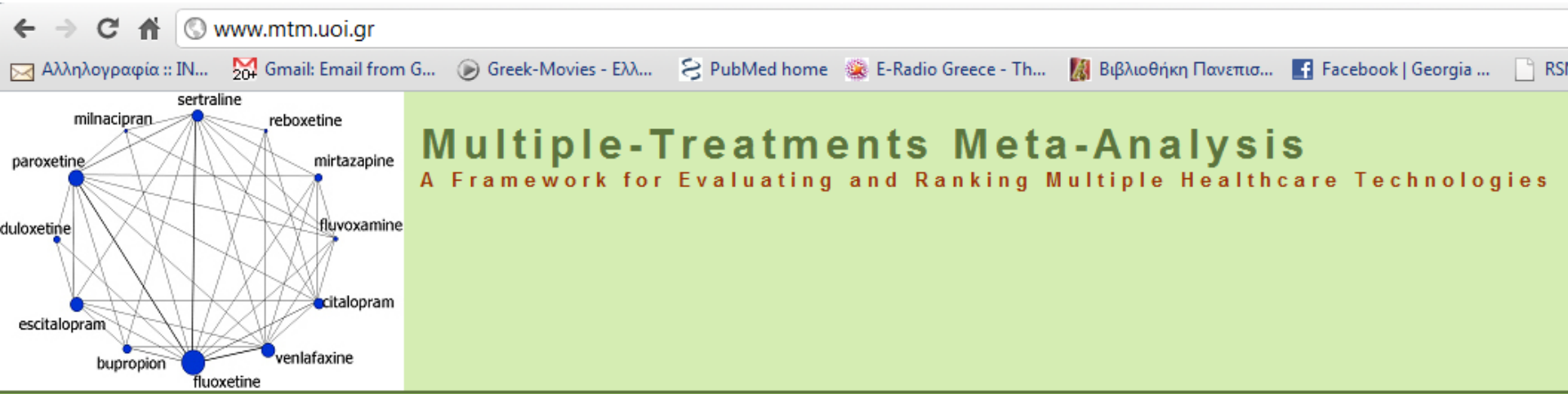
- Try to explain inconsistency!
- Use network meta-regression
- Might consider
  - presenting results from the inconsistency model
  - presenting a variety of separate direct, indirect and mixed comparisons
- Be careful! Selective inclusion of evidence pieces might lead to bias

Comparison of assumptions (random effects models)

Meta-analysis	Network meta-analysis
Similarity of participants, interventions and outcomes	Similarity of participants, outcomes; 'random selection' of interventions
Appropriate modelling of study data (within-study variances often assumed known, uncorrelated with effects)	Appropriate modelling of study data (within-study variances often assumed known, uncorrelated with effects)
Normal distribution for random effects	Normal distribution for random effects
Possibly covariates to explain heterogeneity	Possibly covariates to explain heterogeneity and/or inconsistency
	Possible assumptions about different τ values for different comparisons
	Possible extra parameters to allow for inconsistency across comparisons

Hands on

www.mtm.uoi.gr



Hands on

- www.mtm.uoi.gr
- Go to ‘how to do an MTM’ tab
- Use R routine `mtmnetwork.plot` to plot a network
- Use the R routine `ifplot.fun` to plot inconsistency in all closed loops
- In WinBUGS: read the description of models (e.g. www.mtm.uoi.gr/3.binarymodeldescription.pdf) download the data and the WinBUGS code
- Use the R routine `sucraplot.fun` to get rankograms and SUCRA
- Go to ‘STATA routines for Network Meta-Analysis’ tab for an implementation of network meta-analysis

Research Synthesis Methods

The official journal of the Society for Research Synthesis Methodology
http://www.srsm.org/

A special issue for Network Meta-analysis published in 2012

The Cochrane Collaboration

A new methods group has been recently established to support reviews that aim to compare multiple interventions cmimg.cochrane.org

Bucher HC, Guyatt GH, Griffith LE, Walter SD. 1997. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 50:683-691.
Caldwell DM, Ades AE, Higgins JP. 2005. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 331:897-900.
Caldwell DM, Gibb DM, Ades AE. 2007. Validity of indirect comparisons in meta-analysis. Lancet 369:270.
Caldwell DM, Welton NJ, Ades AE. 2010. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. J Clin Epidemiol.
Cooper NJ, Sutton AJ, Ades AE, Paisley S, Jones DR. 2007. Use of evidence in economic decision models: practical issues and methodological challenges. Health Economics 16:1277-1286.
Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. 2009. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Statistics in Medicine 28:1861-1881.
Dias S, Welton N, Marinho V, Salanti G, Ades A. 2010. Estimation and adjustment of Bias in randomised evidence using Mixed Treatment Comparison Meta-analysis. Journal of the Royal Statistical Society (A) 173.
Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, Bradburn M, Eastwood AJ. 2005. Indirect comparisons of competing interventions. Health Technol Assess 9:1-iv
Higgins JP, Whitehead A. 1996. Borrowing strength from external trials in a meta-analysis. Statistics in Medicine 15:2733-2749.
Mavridis D, Salanti G. A practical introduction to multivariate meta-analysis. *Stat Methods Med Res.* 2012 Jan 23.
Salanti G, Ades AE, Ioannidis JP Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2010 Aug 3.
Salanti G, Dias S, Welton NJ, Ades AE et al Evaluating novel agent effects in multiple-treatments meta-regression. Stat Med. 2010 Oct 15;29(23):2369-83.
Salanti G, Higgins JP, Ades AE, Ioannidis JP. 2008a. Evaluation of networks of randomized trials. Statistical Methods in Medical Research 17:279-301.
Salanti G, Kavvoura FK, Ioannidis JP. 2008b. Exploring the geometry of treatment networks. Ann Intern Med 148:544-553.
Salanti G, Marinho V, Higgins JP. 2009. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. Journal of Clinical Epidemiology 62:857-864.
Song F, Altman DG, Glenny AM, Deeks JJ. 2003. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ 326:472.
Song F, Harvey I, Lilford R. 2008. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. Journal of Clinical Epidemiology 61:455-463.
Sutton A, Ades AE, Cooper N, Abrams K. 2008. Use of indirect and mixed treatment comparisons for technology assessment. Pharmacoeconomics 26:753-767.
Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. Int J Epidemiol. 2013 Feb;42(1):332-45
Welton NJ, Cooper NJ, Ades AE, Lu G, Sutton AJ. 2008. Mixed treatment comparison with multiple outcomes reported inconsistently across trials: Evaluation of antivirals for treatment of influenza A and B. Statistics in Medicine 29:5620-5639.

Thank you!
Questions?