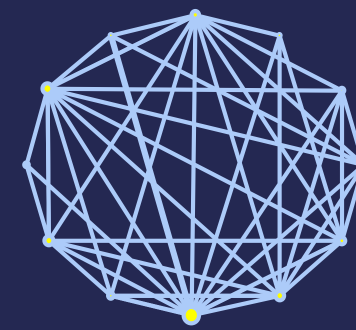
Addressing consistency in networks of randomized trials



Areti Angeliki Veroniki, MSc, PhD

<u>Prepared for</u>: 44th Annual Meeting of the SSC

June 1, 2016

E-mail: VeronikiA@smh.ca

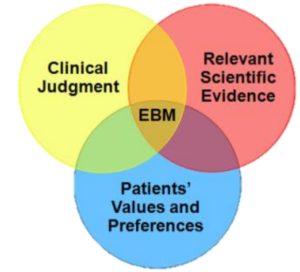
Knowledge Translation Program,
Li Ka Shing Knowledge Institute,
St. Michael's Hospital,
Toronto, Canada

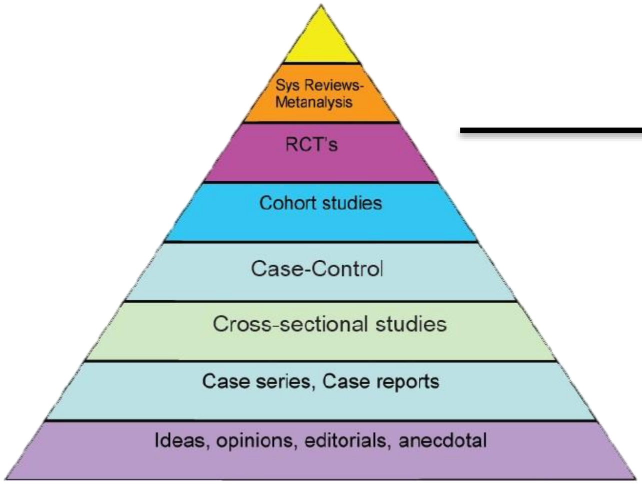
St. Michael's
Inspired Care.
Inspiring Science.

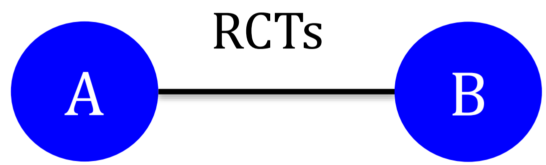
UNIVERSITY OF
TORONTO

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

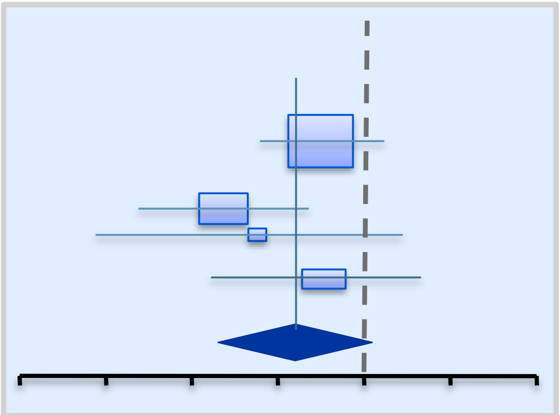
Evidence-Informed Medicine



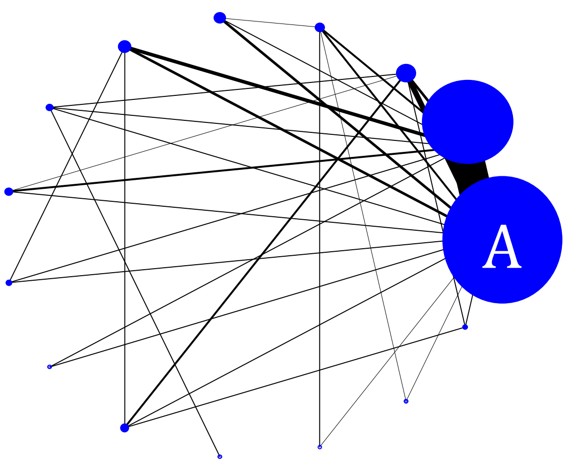


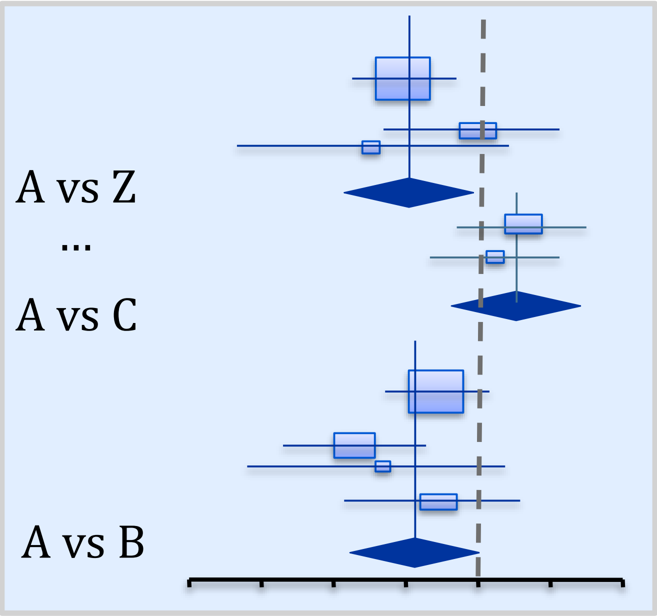


Meta-analysis



Network Meta-analysis





If multiple comparisons are available

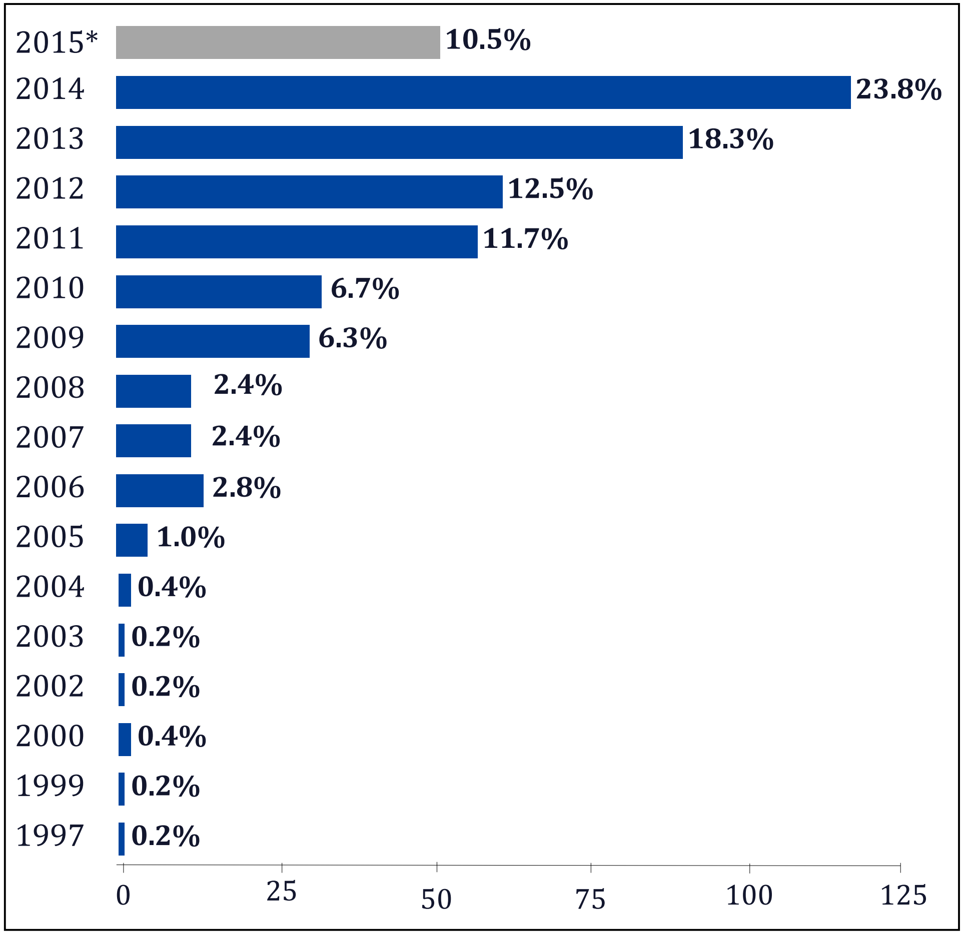
Network Meta-analysis (NMA)

NMA has become increasingly popular over the last two decades with ~500 publications

- When policy makers are considering what interventions to cover through health plans or what safety labels to put on medications, they need evidence from an NMA because this method uses all available RCTs for a specific clinical topic

- The validity of the results from NMA rests on the assumption of transitivity, requiring that the pairwise comparisons are similar in factors which could affect the relative treatment effects.

Temporal distribution of publications (n=494)

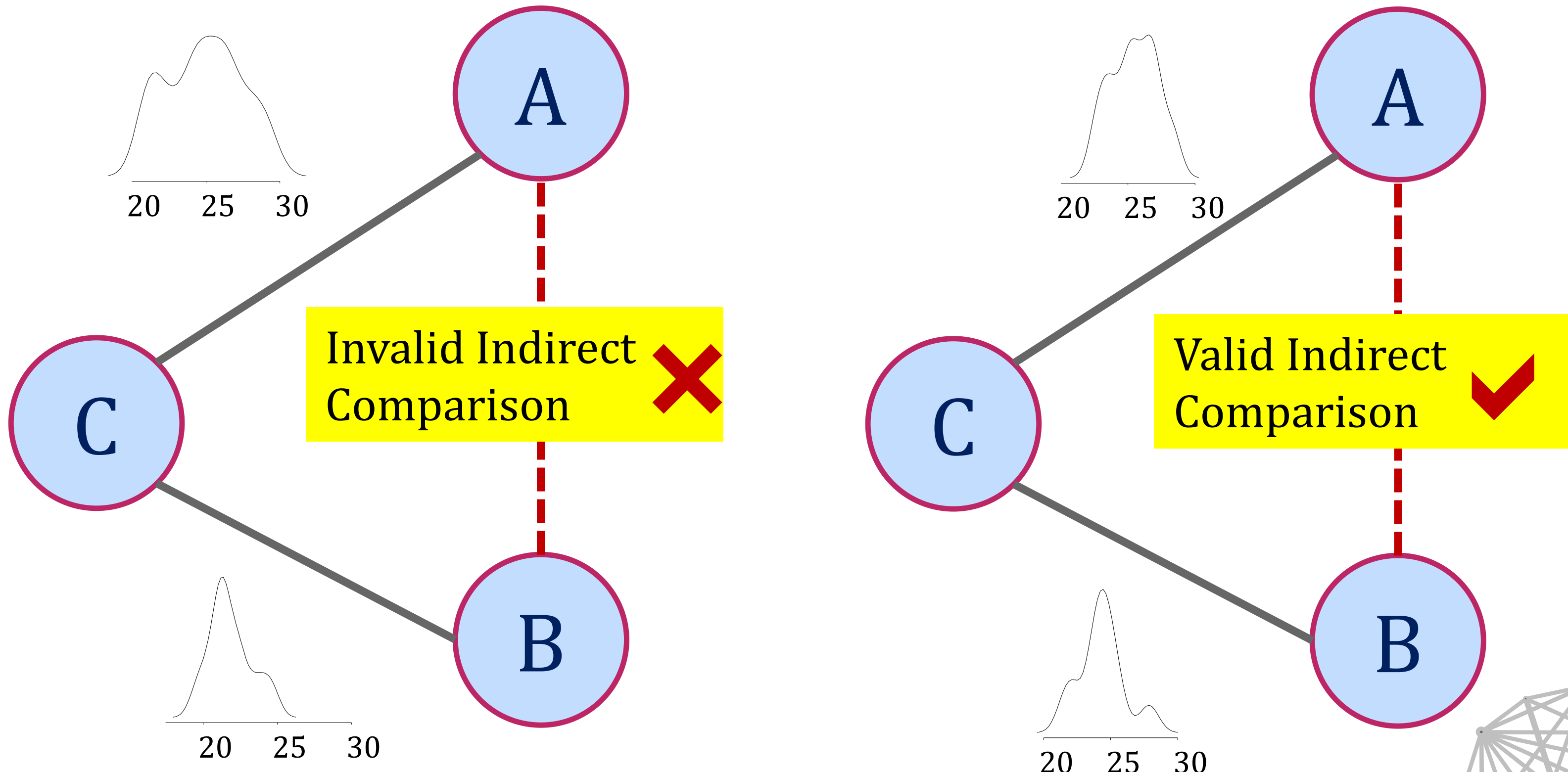


*2015 sample is not complete

Transitivity assumption

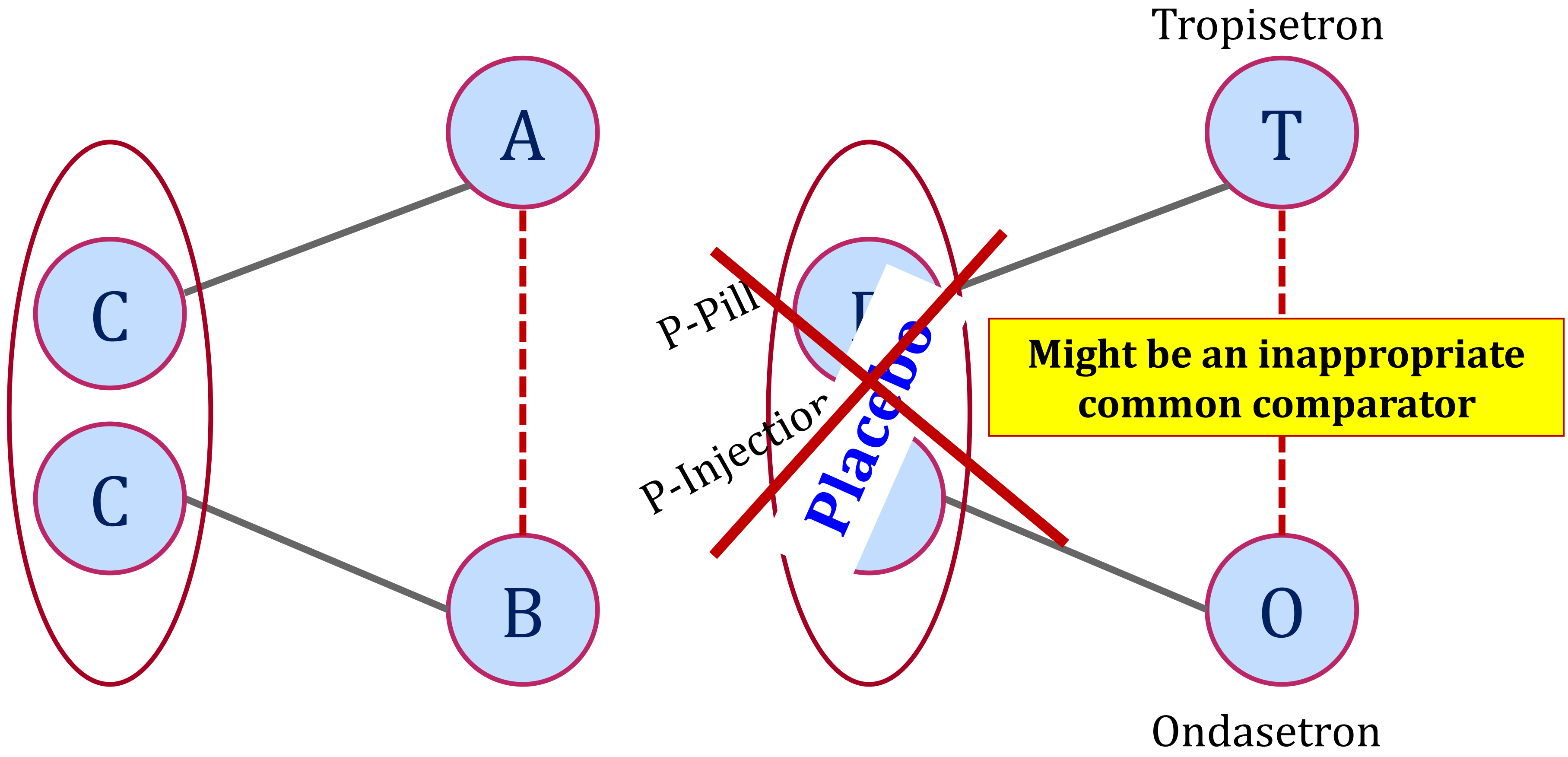
The two sets of trials AC and CB do not differ with respect to the distribution of effect modifiers.

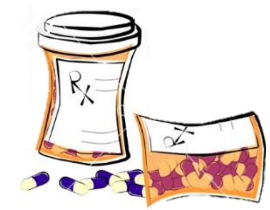
Age effect modifier



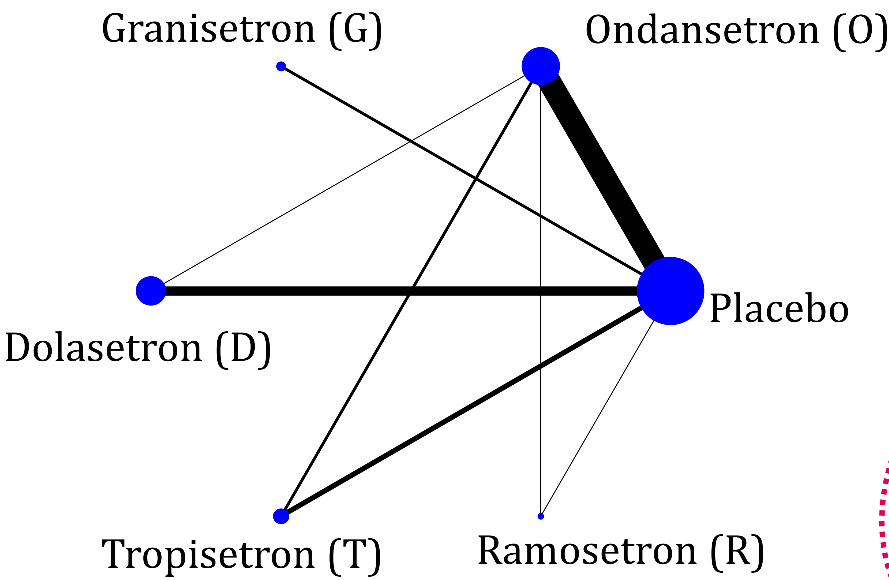
Transitivity assumption

Treatment C should be similar when it appears in AC and BC trials

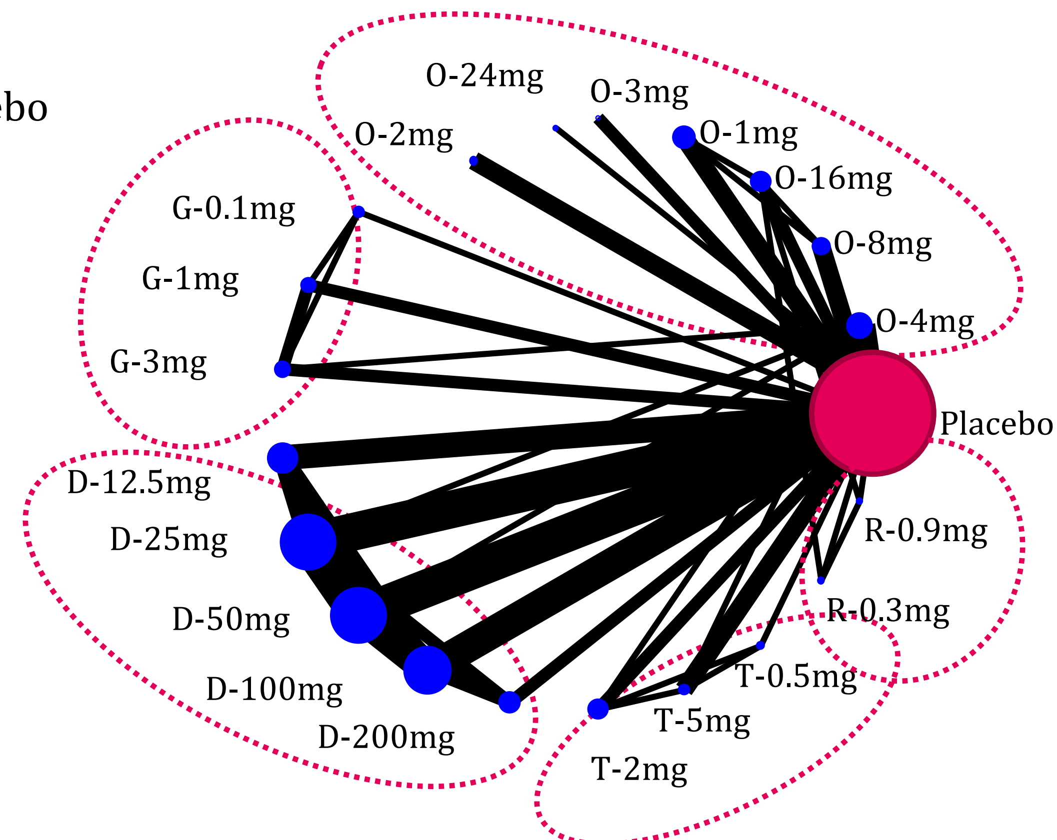




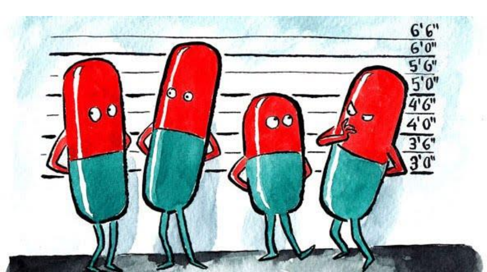
Transitivity assumption



Lumping or splitting nodes?

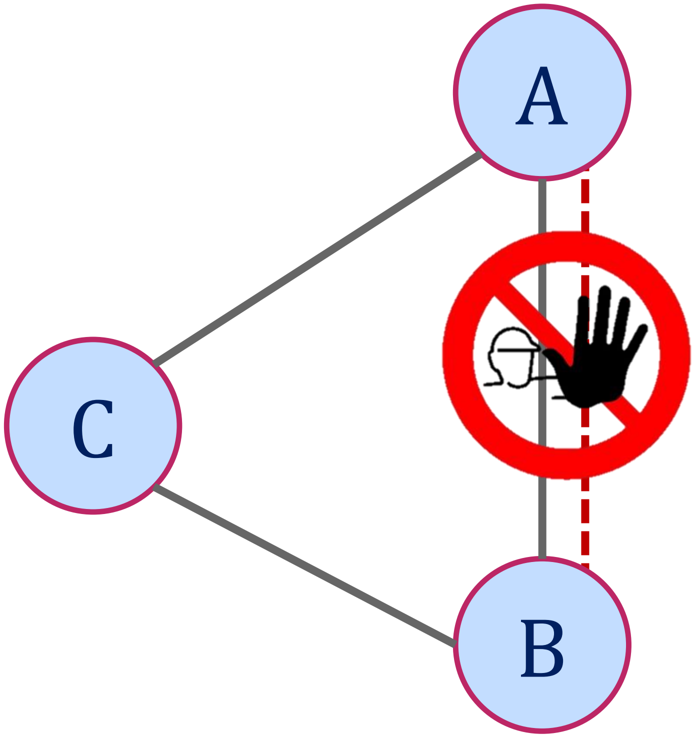


Our initial decisions on the network structure might affect validity of NMA results!



Consistency assumption

When the common comparator is transitive, it allows a valid indirect comparison of the treatments to which it is linked



BUT
Lack of transitivity can create statistical disagreement between direct and indirect evidence

Inconsistency!

Assumption underlying indirect comparison and NMA (in addition to considering homogeneity)

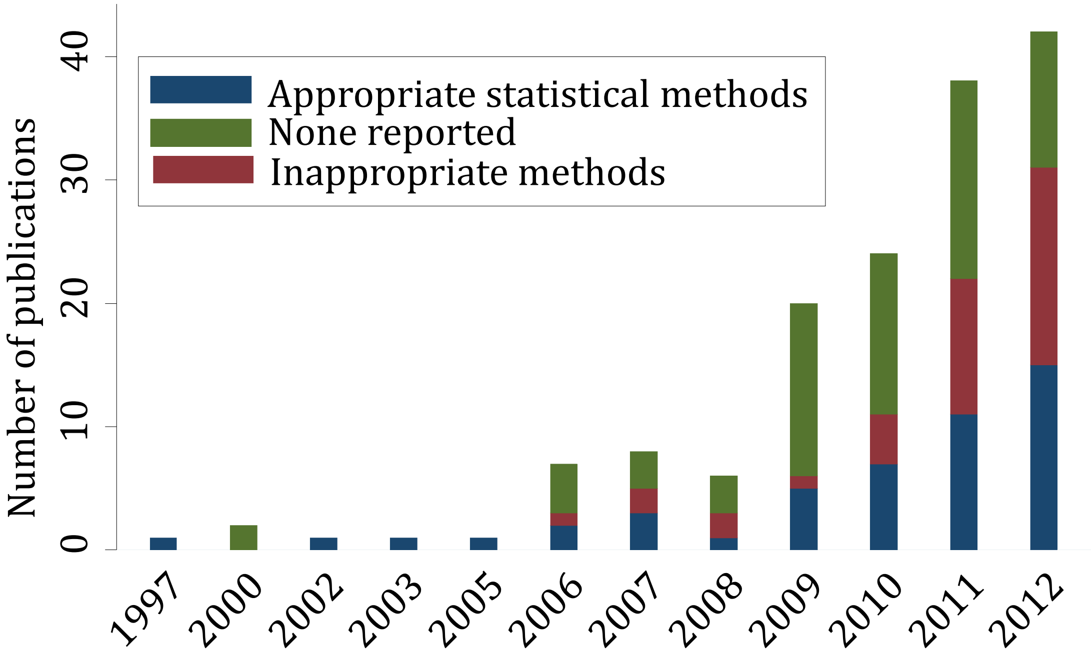
- Assumption for indirect and mixed comparison
  - Conceptual definition (Transitivity)
    - Clinical
    - Methodological
  - Property of parameters and data (Consistency)
    - Statistical

Cipriani et al Ann of Int Medicine 2013

Network Meta-analysis (NMA)

A database of 186 NMAs showed that…

- In 24% of the networks the authors used inappropriate methods to evaluate consistency
  - Comparison of direct with NMA estimates
  - Comparison of previous meta-analyses with NMA results
- In 44% of the networks the authors did not report a method to evaluate consistency

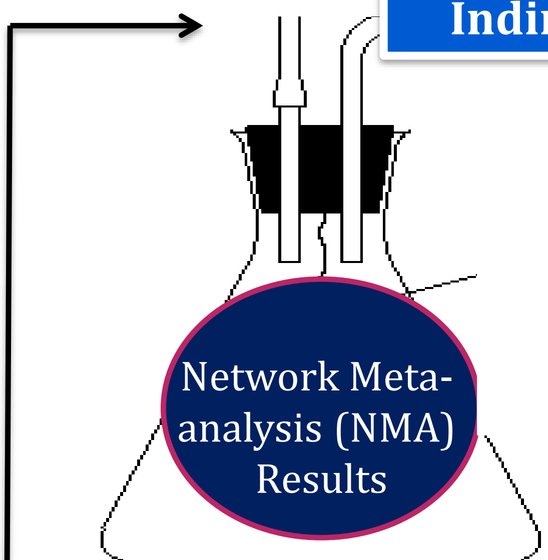


Nikolakopoulou et al PLoS One 2013

Consistency assumption

Direct evidence

Indirect evidence



Are the results valid?

REMEMBER!

Evaluate the assumption of consistency

- *What is the extent of inconsistency in complex networks?*
- *Which factors control its statistical significance?*
- *Which would be the most appropriate approach to employ?*

Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2013;**42**:332–345
doi:10.1093/ije/dys222

Evaluation of inconsistency in networks of interventions

Areti Angeliki Veroniki,[1] Haris S Vasiliadis,[2,3] Julian PT Higgins[4,5] and Georgia Salanti[1]*

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

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

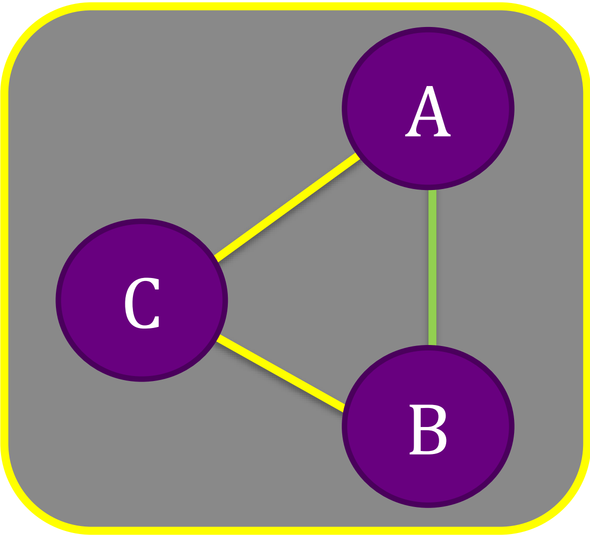
Accepted 21 November 2012

Background The assumption of consistency, defined as agreement between direct and indirect sources of evidence, underlies the increasingly popular method of network meta-analysis. No evidence exists so far

Aim: To update our previous empirical evaluation by employing different statistical approaches to detect inconsistency and to estimate empirically the prevalence of inconsistency in a set of published networks.

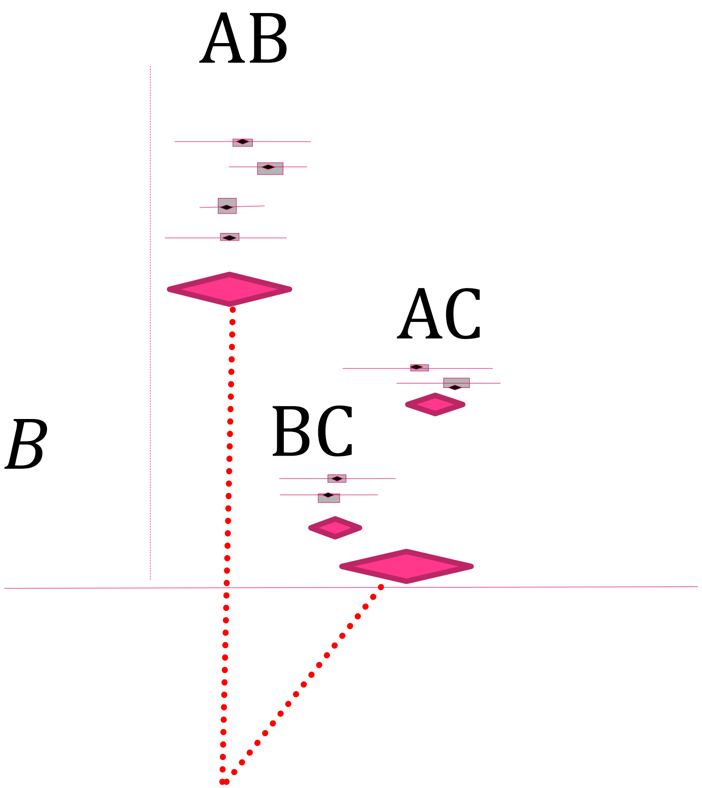
Forms of Inconsistency

Loop Inconsistency



Direct AB

Indirect AB

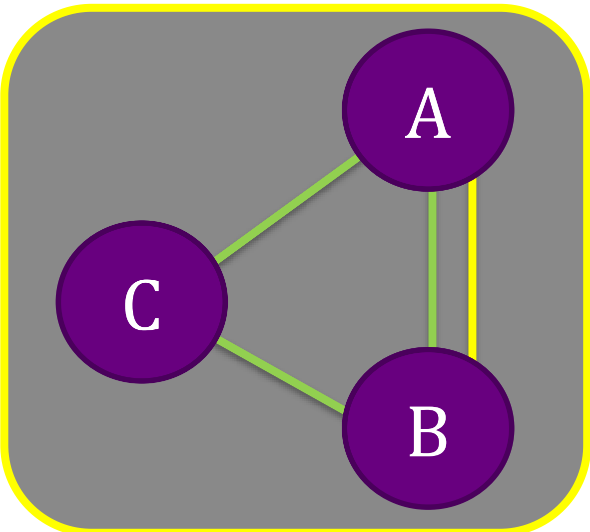


If they statistically differf : *Inconsistency!*

Lu and Ades JASA 2006

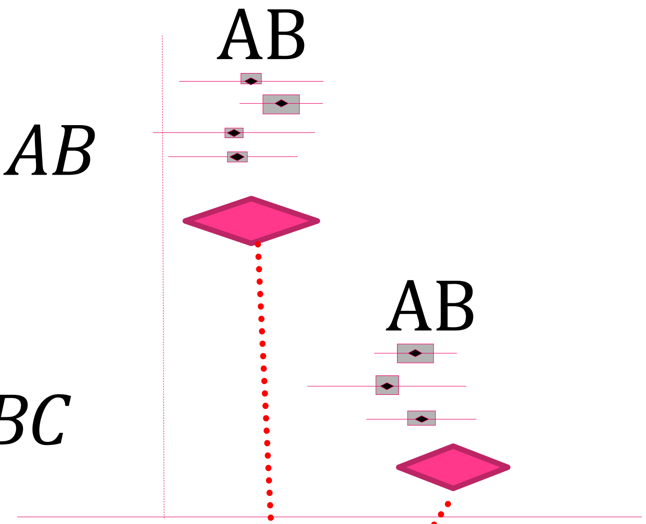
Forms of Inconsistency

Design Inconsistency



Design AB

Design ABC



If they statistically differer : *Inconsistency!*

White et al RSM 2012
Higgins et al RSM 2012

Approaches for evaluating…

LOCAL INCONSISTENCY

- Loop-Specific (LS)
- Node-splitting / Separating Indirect and Direct Evidence (SIDE)
- Separating One Design from the Rest (SODR)

GLOBAL INCONSISTENCY

- Composite test for inconsistency
- Lu and Ades (LA)
- Design by treatment interaction (DBT)

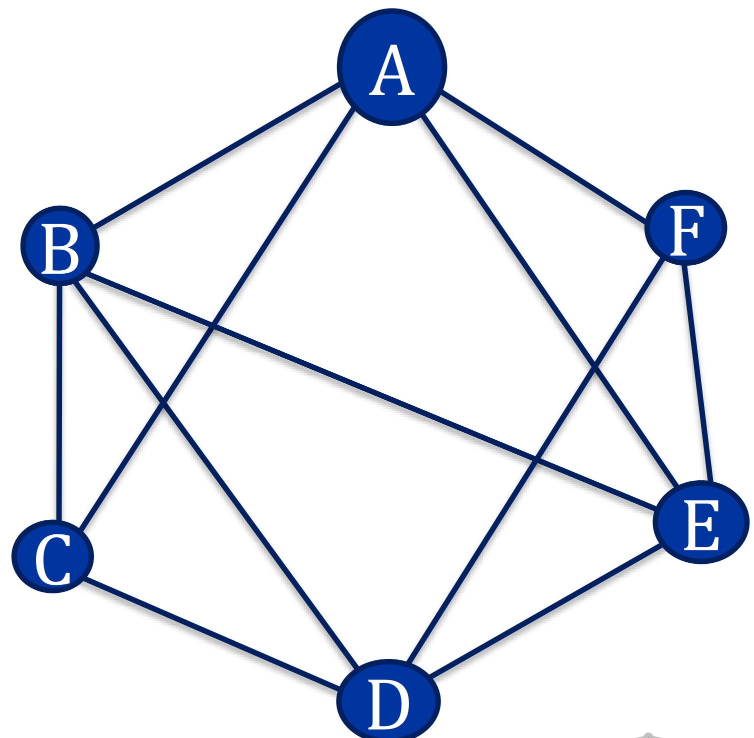
❑ Note: There is also **Comparison of model fit and parsimony between consistency and inconsistency models** approach

- Requires Bayesian framework – uses the measures of model fit & parsimony (e.g. DIC)
- Does **not** provide inconsistency estimates
- Infers on global inconsistency

Approaches for inconsistency

Fictional Dataset

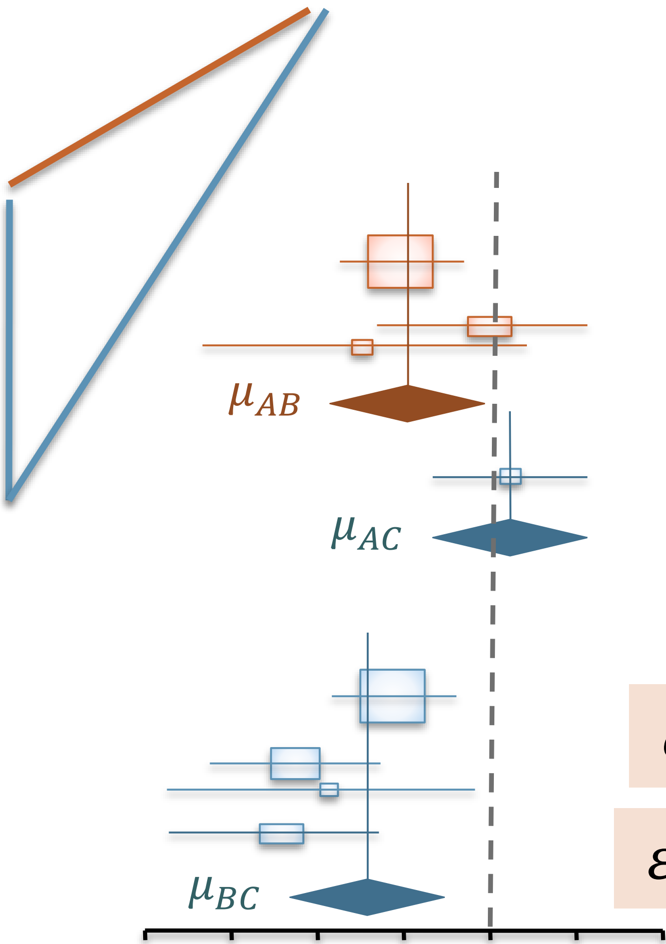
studies	A	B	C	D	E	F
3	■				■	
2	■	■				
1					■	■
1				■		■
7		■			■	
6				■	■	
3	■					■
1	■		■			
4		■	■			
1		■		■		
1			■	■		



Approaches for local inconsistency

Loop-Specific (LS) Method

Multiple tests evaluating inconsistency within each closed loop



$$y_{i,AB} = \mu_{AB} + \delta_{i,AB} + \varepsilon_{i,AB}$$
$$y_{i,AC} = \mu_{AC} + \delta_{i,AC} + \varepsilon_{i,AC}$$
$$y_{i,BC} = \mu_{BC} + \delta_{i,BC} + \varepsilon_{i,BC}$$

$$\delta_{i,AB} \sim N(0, \tau^2_{AB})$$

$$\varepsilon_{i,AB} \sim N(0, v_{i,AB})$$

$$\delta_{i,AC} \sim N(0, \tau^2_{AC})$$

$$\delta_{i,BC} \sim N(0, \tau^2_{BC})$$

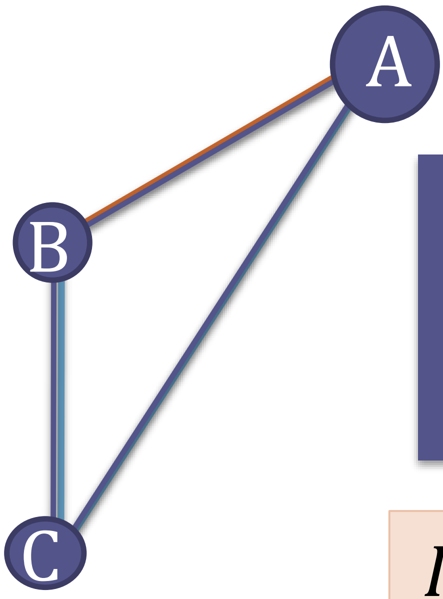
$$\varepsilon_{i,AC} \sim N(0, v_{i,AC})$$

$$\varepsilon_{i,BC} \sim N(0, v_{i,BC})$$

Bucher et al. JCE 1997; Veroniki et al. Int.J.E. 2013

Approaches for local inconsistency

Loop-Specific (LS) Method



Multiple tests evaluating inconsistency within each closed ***loop***

$$y_{i,AB} = \mu_{AB} + \delta_{i,AB} + \varepsilon_{i,AB}$$
$$y_{i,AC} = \mu_{AC} + \delta_{i,AC} + \varepsilon_{i,AC}$$
$$y_{i,BC} = \mu_{BC} + \delta_{i,BC} + \varepsilon_{i,BC}$$

$$IF_{ABC}^{LS} = |\mu_{AB} - (\mu_{AC} - \mu_{BC})|$$

Statistical Evaluation: $H_0 : IF_{ABC}^{LS} = 0$

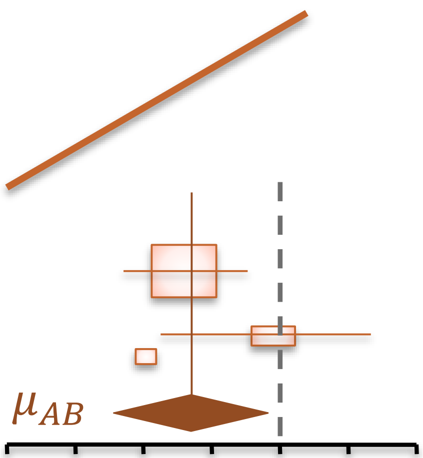
$$W_{LS} = \frac{\widehat{IF}_{ABC}^{LS}}{\sqrt{\widehat{Var}(IF_{ABC}^{LS})}} \sim N(0,1)$$

Bucher et al. JCE 1997; Veroniki et al. Int J Epidemiol 2013

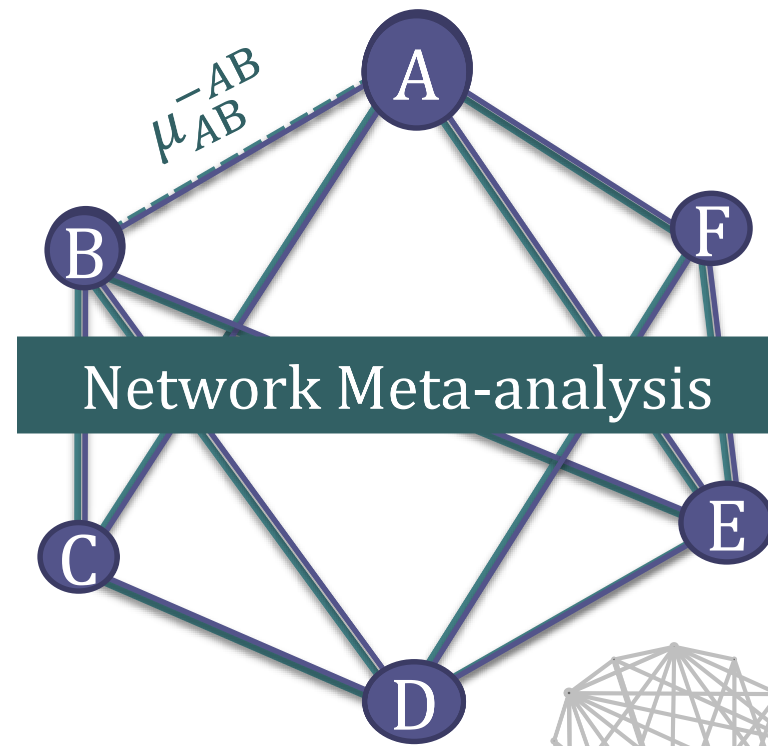
Approaches for local inconsistency

Node-Splitting or Separating Indirect and Direct Evidence (SIDE)

Multiple tests evaluating inconsistency for each comparison in the network



$$IF^{SIDE}_{AB} = \left|\mu_{AB} - \mu^{-AB}_{AB}\right|$$

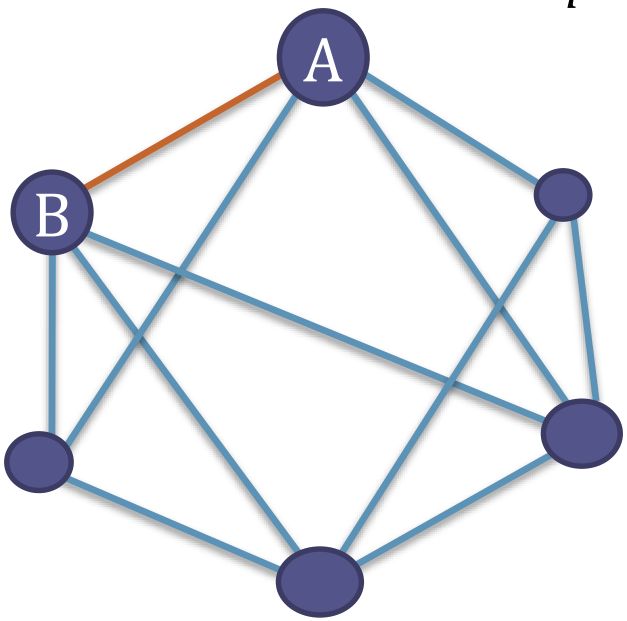


Dias et al. Stat Med 2010

Approaches for local inconsistency

Node-Splitting or Separating Indirect and Direct Evidence (SIDE)

Multiple tests evaluating inconsistency for each comparison in the network



observed treatment effect

inconsistency factor

within-study error

$$y_{i,AT} = \mu_{AT}^{-AB} + IF_{\mathrm{AB}}^{SIDE} + \delta_{i,AT} + \varepsilon_{i,AT}$$

mean of the distribution of the study-spe... the netwo... comparison... rem...

...ts

Statistical Evaluation: $H_0: IF_{AB}^{SIDE} = 0$

$$W_{AB}^{SIDE} = \frac{\widehat{IF}_{AB}^{SIDE}}{\sqrt{\widehat{Var}(IF_{AB}^{SIDE})}} \sim N(0,1)$$

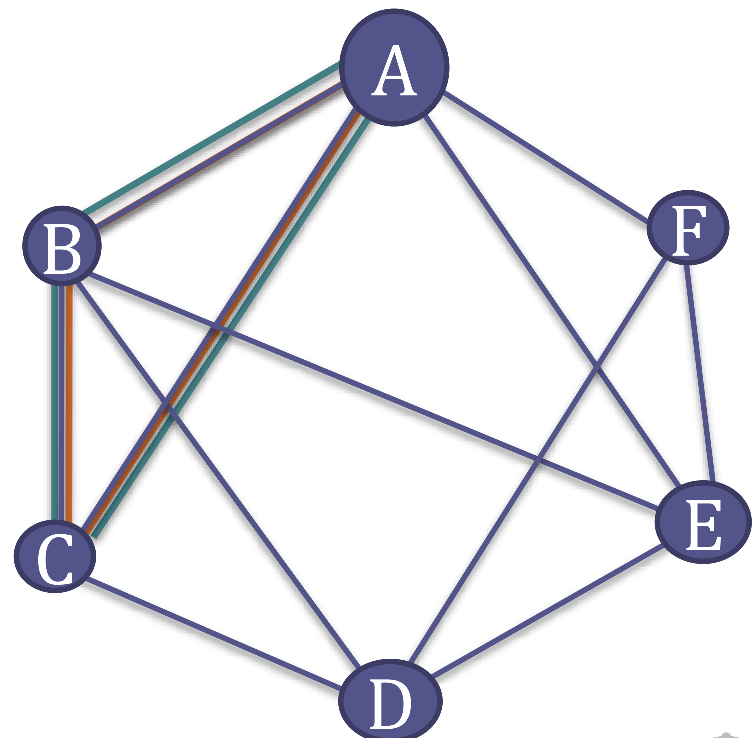
Dias et al. Stat Med 2010

Approaches for global inconsistency

Design By Treatment Interaction (DBT) Model

Global Test assessing both design and loop inconsistency

A	B	C	D	E	F



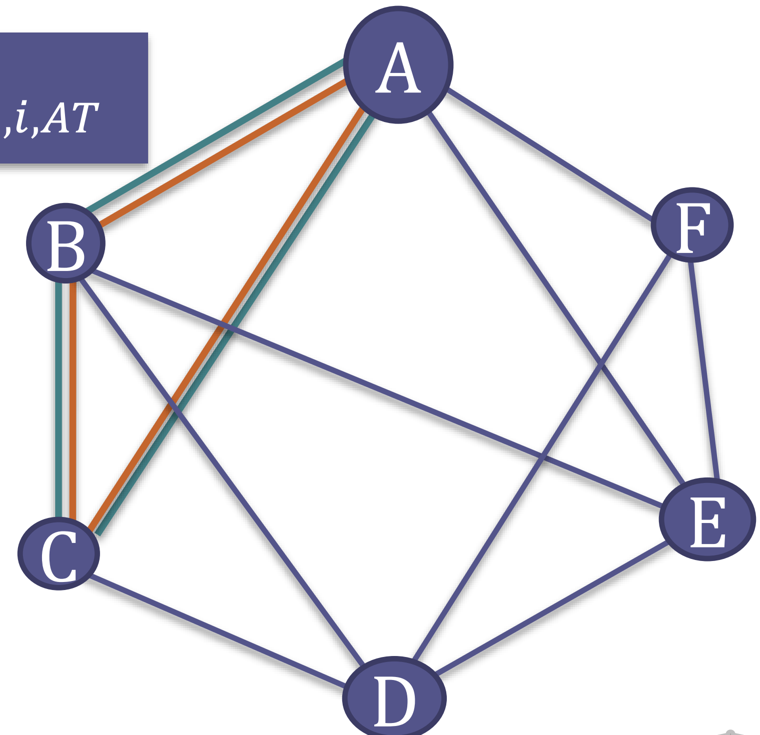
White et al RSM 2012
Higgins et al RSM 2012

Approaches for global inconsistency

Design By Treatment Interaction (DBT) Model

Global Test assessing both design and loop inconsistency

$$y_{d,i,AT} = \mu_{AT} + IF_{d,AT} + \delta_{d,i,AT} + \varepsilon_{d,i,AT}$$



Statistical Evaluation: $H_0: \boldsymbol{IF} = \mathbf{0}$

$$W^{DBT} = \boldsymbol{IF}'\boldsymbol{\Sigma}^{-1}\boldsymbol{IF} \sim \chi_l^2$$

The *joint significance* of all *IFs* is evaluated using the Global χ^2 test

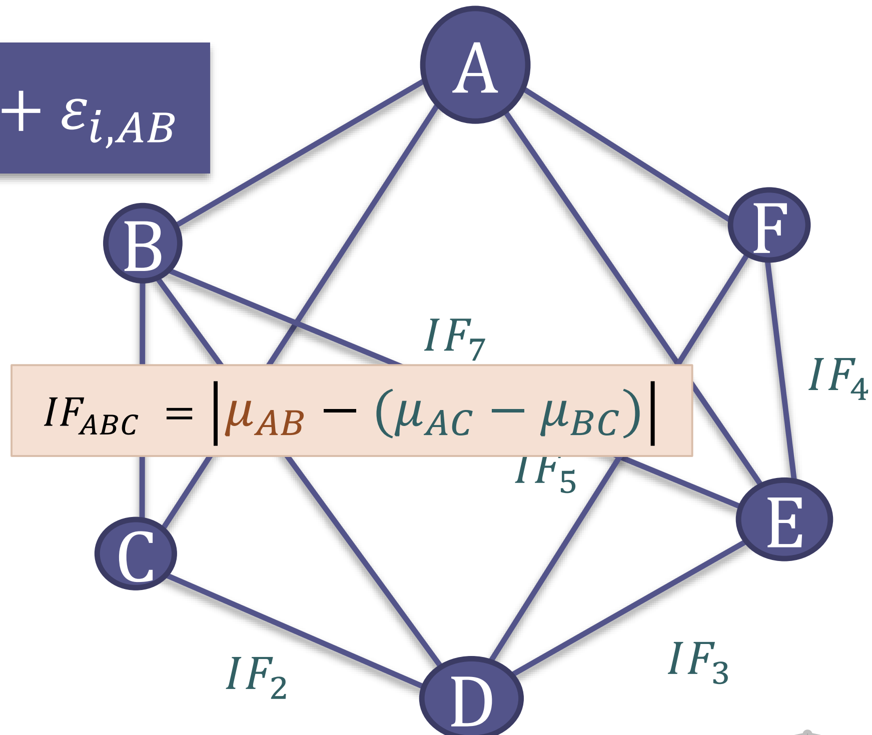
White et al RSM 2012
Higgins et al RSM 2012

Approaches for global inconsistency

Lu and Ades (LA) model

Global test assessing loop inconsistency

$$y_{i,AB} = \mu_{AC} - \mu_{BC} + IF_{ABC} + \delta_{i,AB} + \varepsilon_{i,AB}$$



Statistical Evaluation: $H_0: \boldsymbol{IF} = \mathbf{0}$

$$\mathrm{W}^{LA} = \boldsymbol{IF}'\boldsymbol{\Sigma}^{-1}\boldsymbol{IF} \sim \chi_f^2$$

The *joint significance* of all *IFs* is evaluated using the Global χ^2 test

Lu and Ades JASA 2006

Properties of the inconsistency approaches

	Loop-Specific	SIDE/ Node splitting	SODR	LA	DBT
Simple to compute	✓	✗	✗	✗	✗
Insensitive to parameterization of multi-arm studies	✗	✗	✓	✗	✓
Indirect estimate derived from the entire network	✗	✓	✓	✓	✓
Does not suffer from multiple testing	✗	✗	✗	✓	✓
Power	✗	✗	?	✗	?

Song et al BMC Med Res Methodol 2012, Veroniki et al BMC Med Res Methodol 2014

Database with 40 networks of interventions

Loop-Specific (LS) Method

Out of the 303 loops :

☑ The loop inconsistency rate ranges between 2% and 10%

☑ Statistical inconsistency does **not** importantly differ between the four *effect measures*:

	OR	RRH	RRB	RD
Within-loop heterogeneity	8%	9%	10%	10%
Within-network heterogeneity	5%	6%	6%	5%

Veroniki et *al* IJE (2013)

Database with 40 networks of interventions

Loop-Specific (LS) Method

☑ The different assumptions and estimators for heterogeneity can importantly **impact** on the assessment of inconsistency:

	OR	RRH	RRB	RD
Within-loop heterogeneity	8%	9%	10%	10%
Within-network heterogeneity	5%	6%	6%	5%

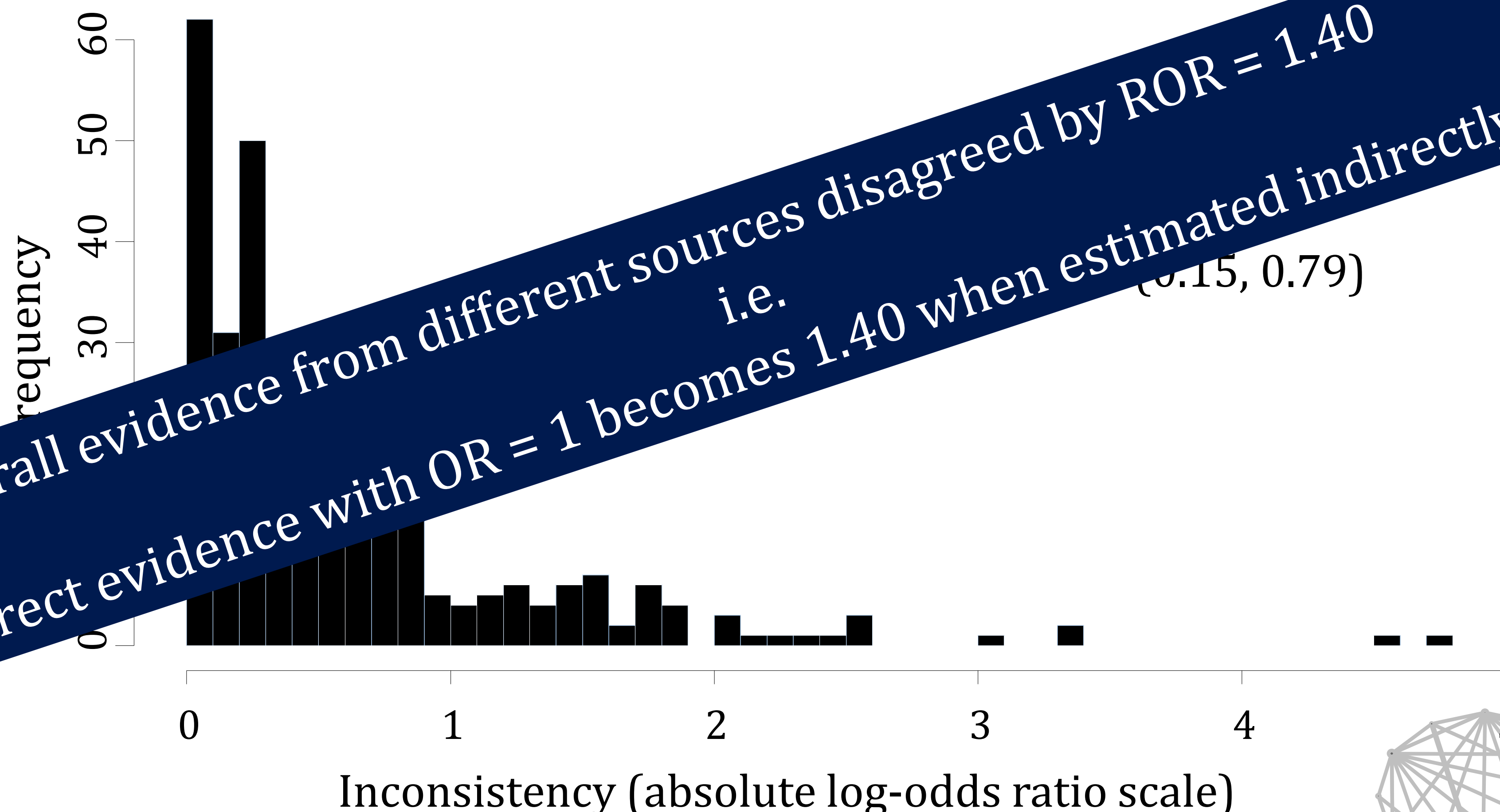
	DL	REML	SJ
Within-loop heterogeneity	8%	7%	5%

☑ Evidence loops that include comparisons informed by a **single** study are **more likely** to show inconsistency:

- 2% to 9% depending on the estimator and assumption for heterogeneity

Distribution of IF in loops

Loop-Specific (LS) Method



Overall evidence from different sources disagreed by ROR = 1.40

i.e.

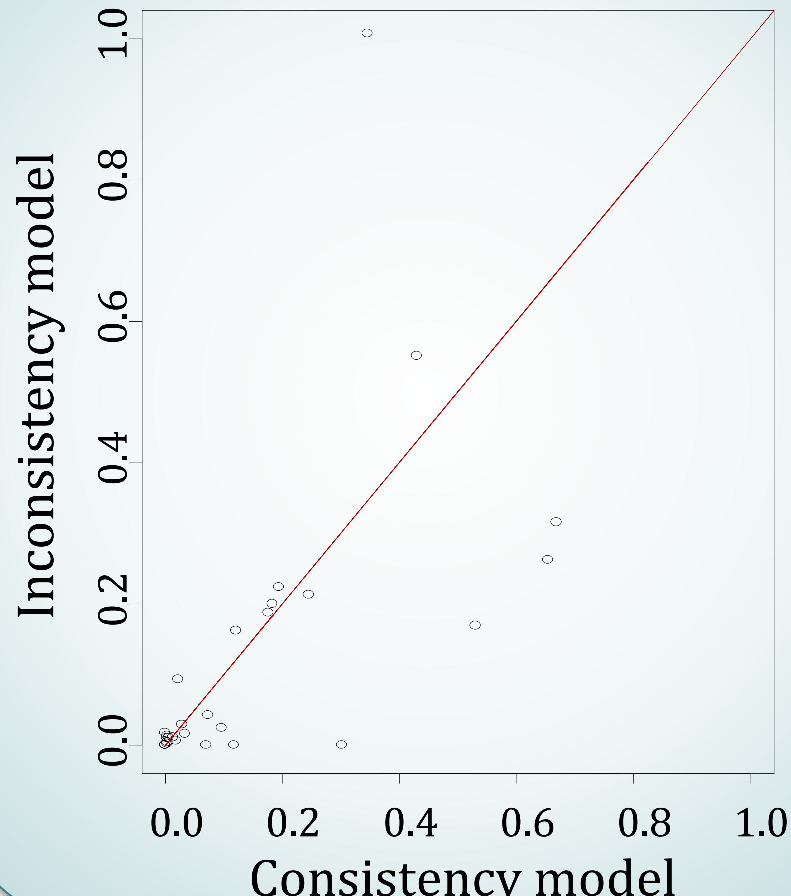
Direct evidence with OR = 1 becomes 1.40 when estimated indirectly

Veroniki et *al* IJE 2013

Inference on networks

Design by treatment interaction (DBT) model

common-within network τ^2
with REML estimator



The *consistency* models display <u>higher</u> heterogeneity accounting probably for *inconsistency* in the data

Veroniki et *al* IJE 2013

In summary…

- Inconsistency can occur in one in ten of the loops and in one in eight of the networks.
- Lower statistical heterogeneity is associated with more chances to detect inconsistency but the estimated magnitude of inconsistency is lower
- Care is needed when interpreting the results of a consistency test as issues of heterogeneity and power may limit its usefulness
- Results and inferences on the prevalence of inconsistency are sensitive to the estimation method of the heterogeneity.
- A sensitivity analysis in the assumptions of heterogeneity may be needed before concluding the absence of statistical inconsistency, particularly in networks with few studies.

References…

1. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol 1997* Jun;**50**(6):683-91.
2. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-944.
3. Donegan S, Williamson P, D'Alessandro U, Tudur-Smith C. Assessing key assumptions of network meta-analysis: a review of methods Research Synthesis Methods 2013.
4. Lu GB, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 2006;101:447-459
5. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G, Characteristics of networks of interventions: a description of a database of 186 published networks. *PLoS One*. 2014;**9**(1):e86754.
6. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconstency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods* 2012;3:98-110.
7. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* 2012;**3**(2):80-97.
8. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, et al. Inconsistency between direct and indirect comparisons of competing interventions: Meta-epidemiological study. BMJ. 2011;343:d4909.
9. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013;42:332-345.
10. Veroniki AA, Mavridis D, Higgins JP, Salanti G. Statistical evaluation of inconsistency in a loop of evidence: a simulation study informed by empirical evidence. *BMC Med. Res. Methodol.* 2014; **14**, 106.
11. Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J. Clin. Epidemiol.* 2016; pii: S0895-4356(16)00153-0.
12. White IR, Barret JK, Jackson D, Higgins JPT. Consistency and inconsistency in multiple treatments meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods* 2012;**3**(2):111-25.

E-mail: VeronikiA@smh.ca

UNIVERSITY OF TORONTO

St. Michael's

Inspired Care.
Inspiring Science.