



Lecture 3

Assumptions underlying indirect and mixed treatment comparisons

Georgia Salanti

Inconsistent indirect comparison

Canary Wharf vs Empire State = 140m

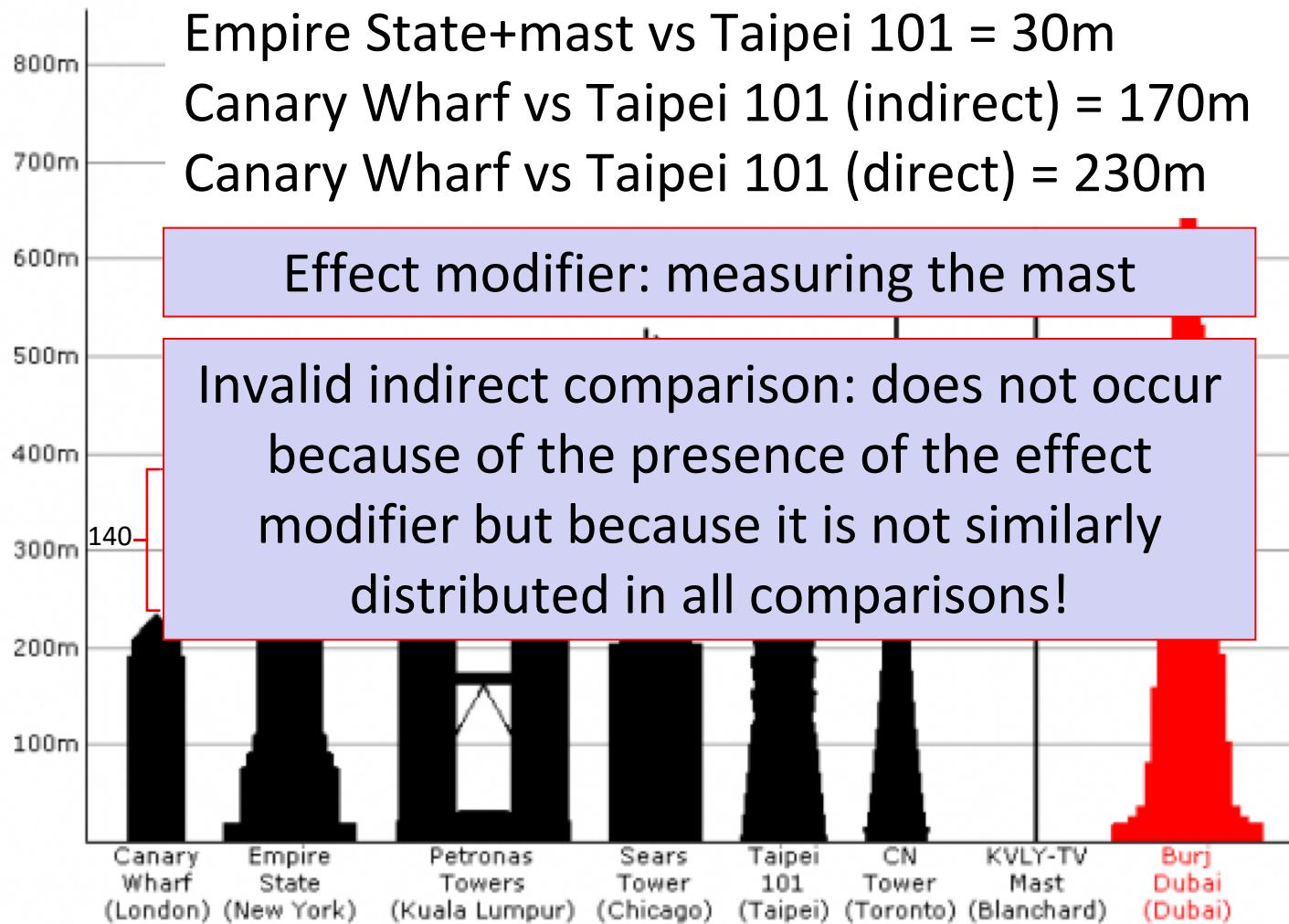
Empire State+mast vs Taipei 101 = 30m

Canary Wharf vs Taipei 101 (indirect) = 170m

Canary Wharf vs Taipei 101 (direct) = 230m

Effect modifier: measuring the mast

Invalid indirect comparison: does not occur because of the presence of the effect modifier but because it is not similarly distributed in all comparisons!



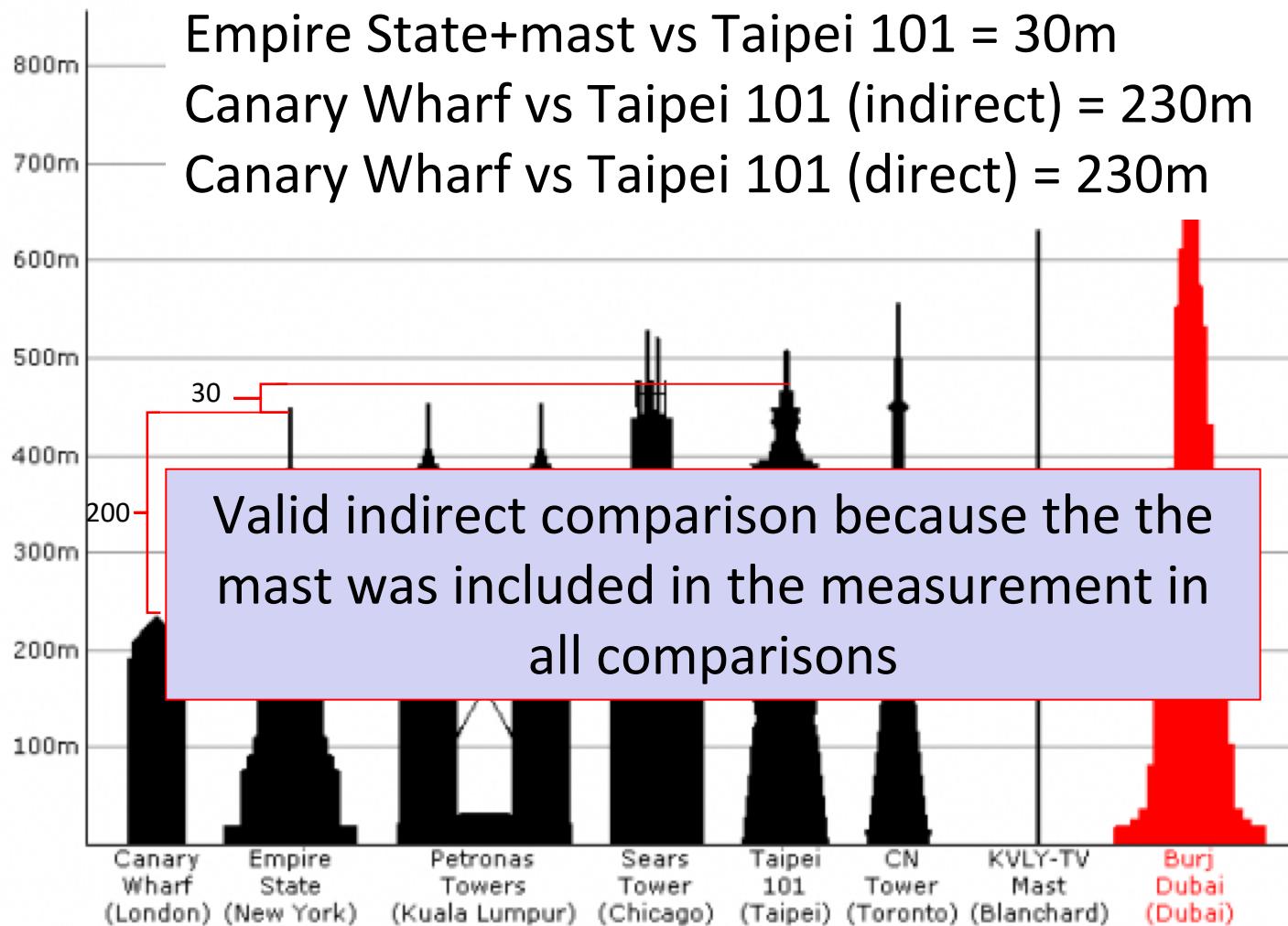
Indirect comparison via *Empire State plus mast*

Canary Wharf vs Empire State+mast = 200m

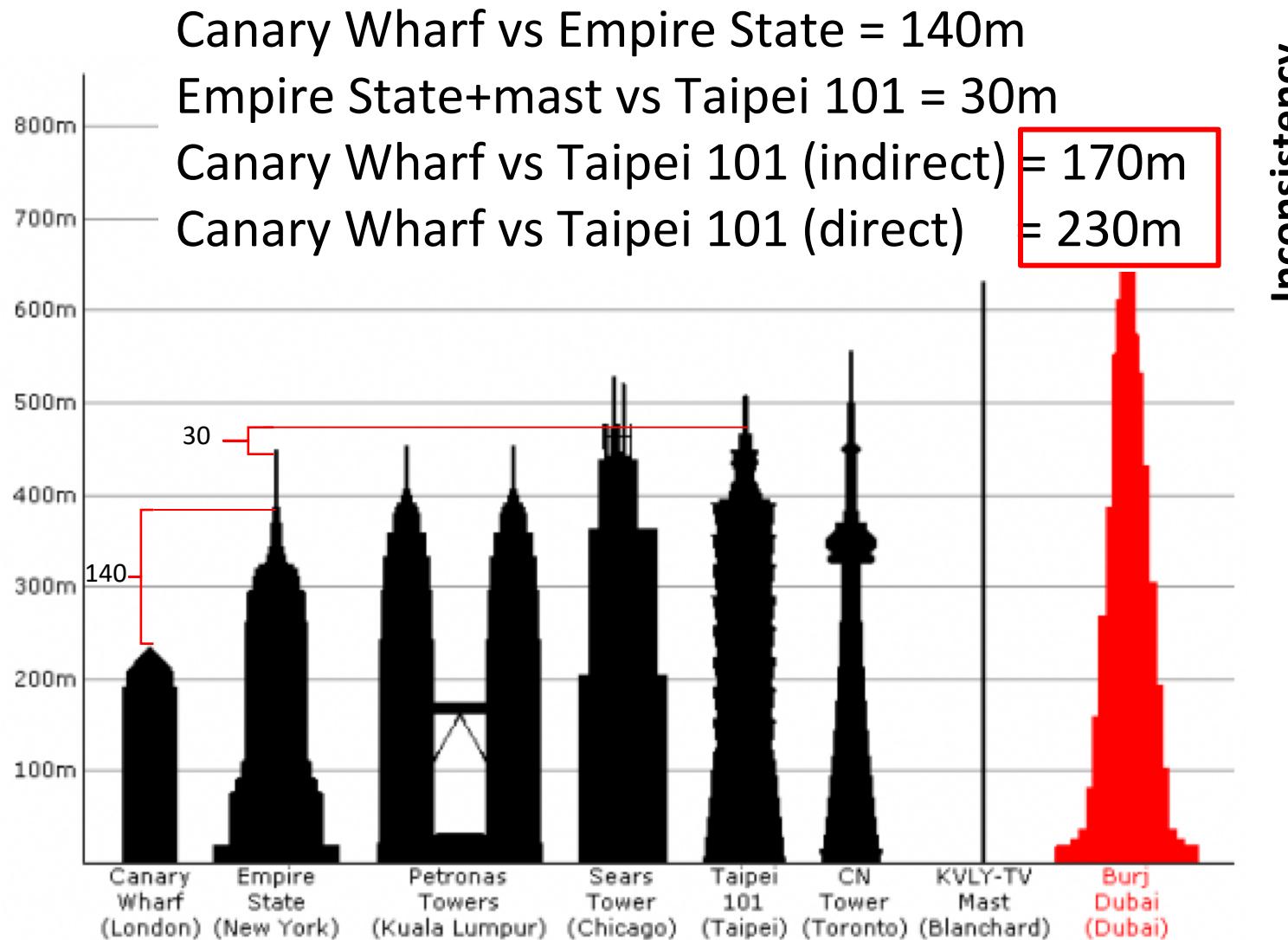
Empire State+mast vs Taipei 101 = 30m

Canary Wharf vs Taipei 101 (indirect) = 230m

Canary Wharf vs Taipei 101 (direct) = 230m



Inconsistent indirect comparison

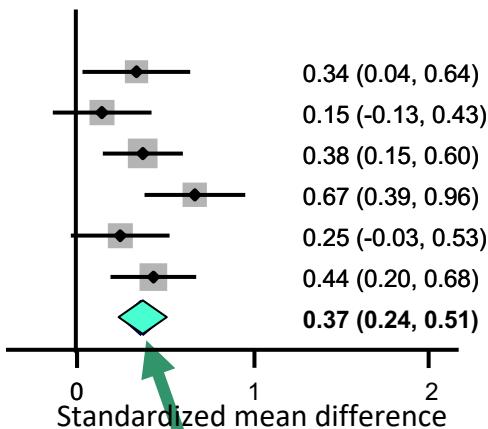


Inconsistency

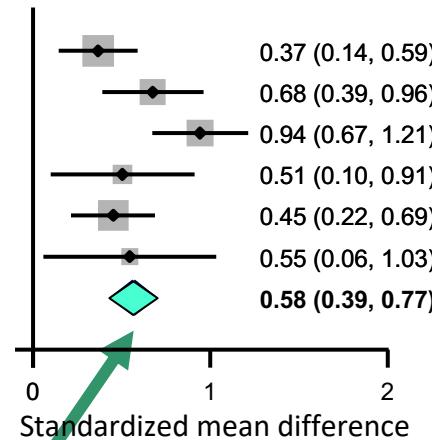
Network meta-analysis is great*

***it can become
a dangerous
weapon if
undertaken
carelessly!**

QTP vs PLA



HAL vs PLA

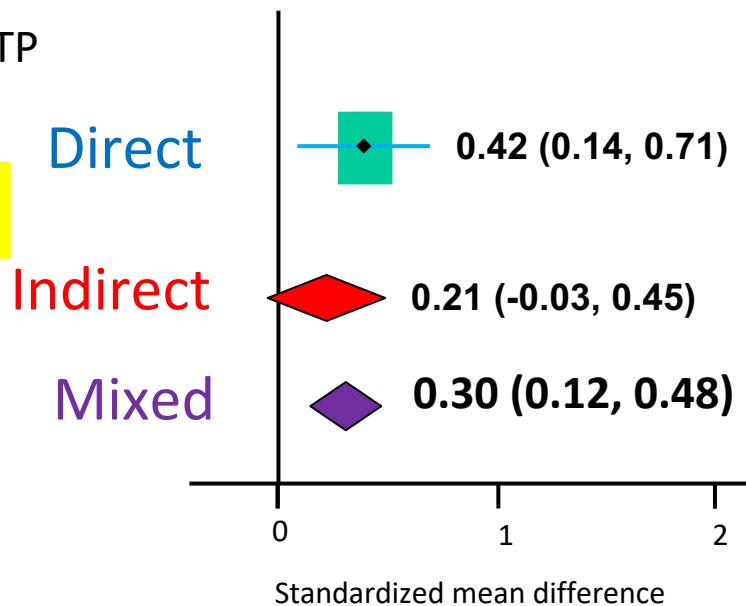


HAL is 0.58 better than P

QTP is 0.37 better than P
...so HAL is 0.21 better than QTP

Are they different?

HAL vs QTP



Indirect comparison

When is it valid?

Validity depends on **transitivity** assumption

advantage of HAL over QTP =
advantage of HAL over PLA

–

advantage of QTP over PLA

Indirect comparison

When is it valid?

Validity depends on **transitivity** assumption

advantage of HAL over QTP =
advantage of HAL over PLA

–

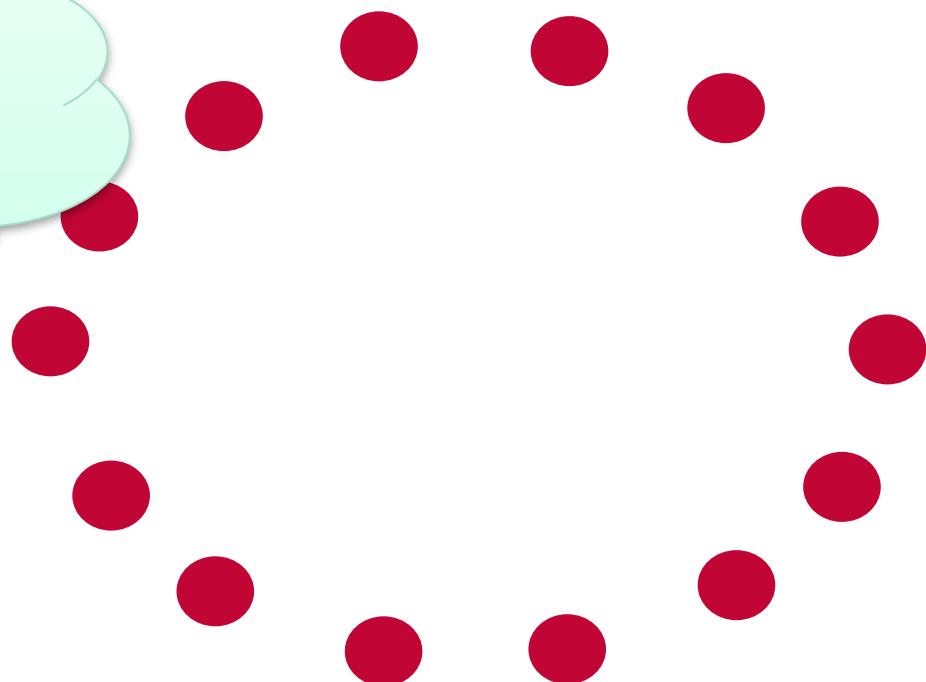
advantage of QTP over PLA

Requires studies to be similar with respect to the
effect modifiers

Ways of looking at transitivity...

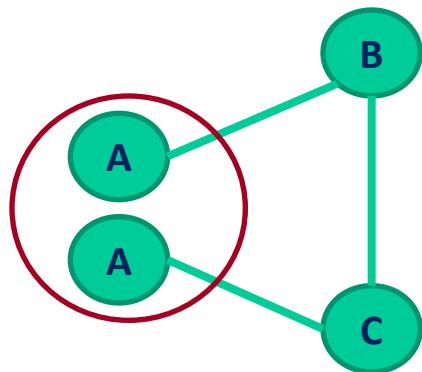
... that all treatments are “jointly randomizable”
for the patients and study settings you consider

think of a multi-arm
trial including all
competing treatments



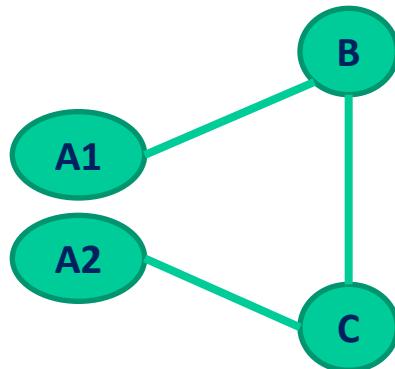
Ways of looking at transitivity...

?



Treatment A must be similar when it appears in AB and AC trials

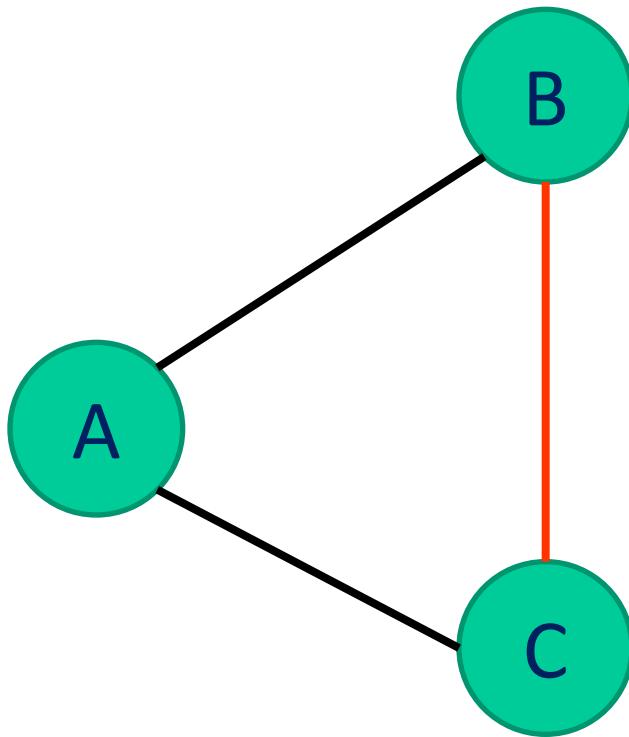
X



Plausible when A can be given in different forms?
(e.g. Compare painkillers and the same drug can be given as an injection or as a pill)?
Placebo or legacy treatments are often problematic

Ways of looking at transitivity...

... A vs B and A vs C trials do not differ with respect to the distribution of effect modifiers

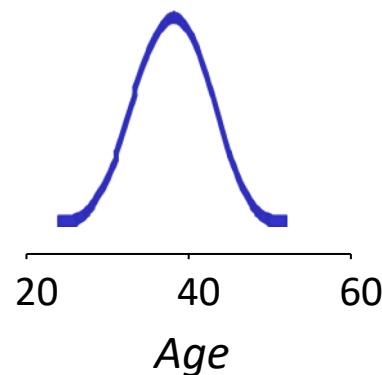


Ways of looking at transitivity...

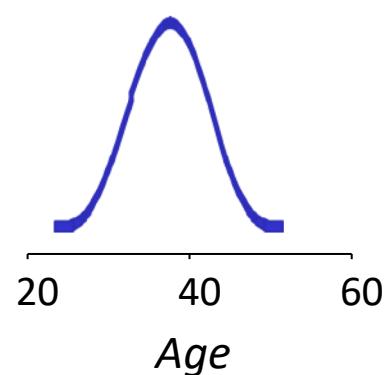
- This formulation facilitates evaluation of the transitivity assumption.
 - Distribution of effect modifiers of the relative treatment effects for similarity in AC and AB trials
- Clinicians and methodologists that aim to synthesize evidence from many comparisons should **identify a priori possible effect** modifiers and compare their distributions across comparisons.

Ways of looking at transitivity... (3)

A vs B

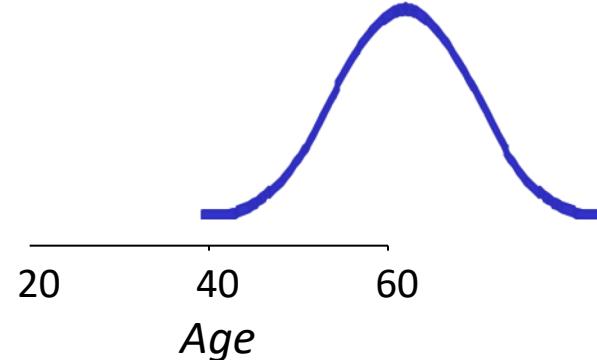
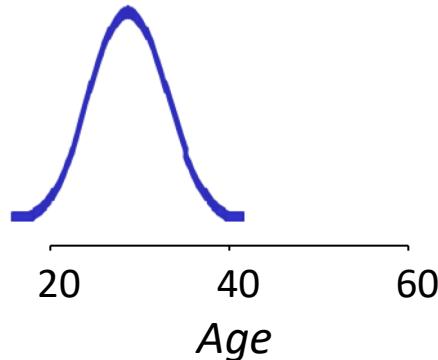


A vs C

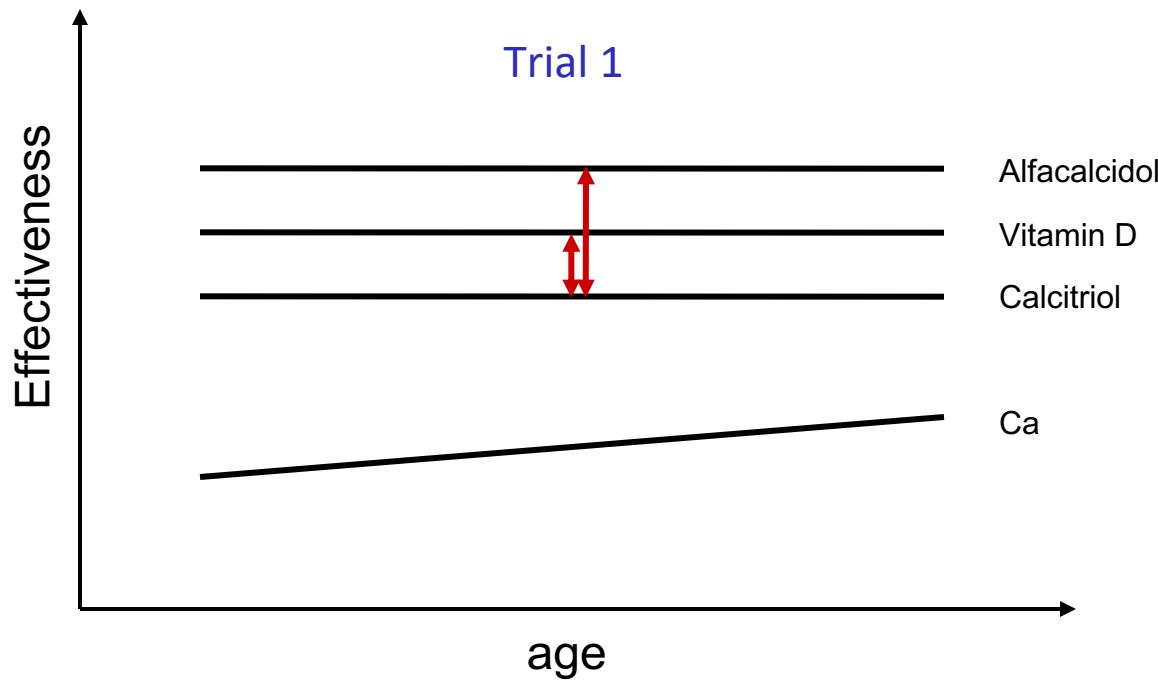


Distribution of an effect modifier across studies

✗

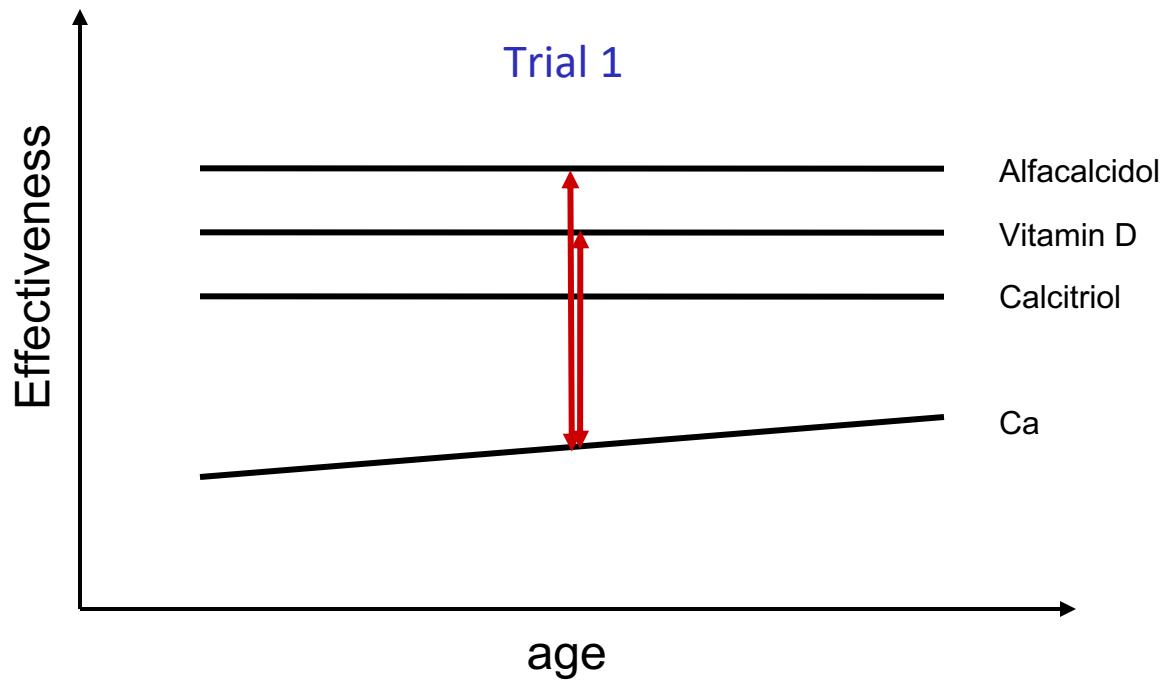


Transitivity refers to the mean effect



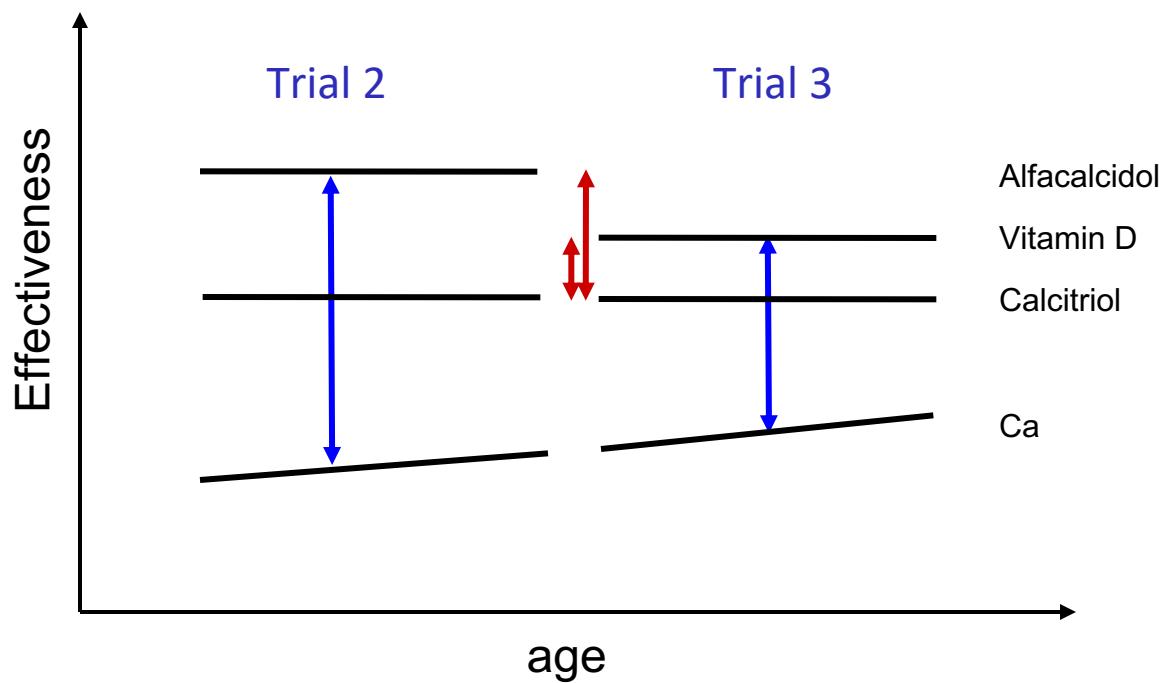
Calcif Tissue 2005 Richy et al

Transitivity refers to the mean effect



Calcif Tissue 2005 Richy et al

Transitivity refers to the mean effect



Calcif Tissue 2005 Richy et al

Ways of looking at transitivity...

- It is important to note however that **the transitivity assumption holds for the mean effect sizes**
 - that is, between the mean summary effects for AC and AB
- Consequently, an effect modifier that differs across studies that belong to the same comparison but has a similar distribution across comparisons will not necessarily violate the transitivity assumption.
 - For example, if age is an effect modifier and AB trials differ in terms of mean age of participants (which will be presented as heterogeneity in AB studies) but the same variability is observed in the set of AC trials then transitivity may hold even if age is an effect modifier.

Transitivity assumption - terminology

- In the literature this assumption has been often referred to as the similarity assumption (e.g. Donegan et al. PLoS 2010)
 - The term ‘transitivity’ describes better the aim of the assumption (to compare two treatments via a third one).
 - ‘similarity’ may wrongly suggest that similarity is required for all characteristics of trials and patients across the evidence base
 - when in reality valid indirect comparison can be obtained even when studies are dissimilar in characteristics that are not effect modifiers
- The violation of the assumption is often referred to in statistical models as ‘treatment-by-trial’ interaction.

Implications for practice

When writing the protocol of a review you need to define

- **Types of interventions**
- Outcomes
- Health condition
- Type of population/settings

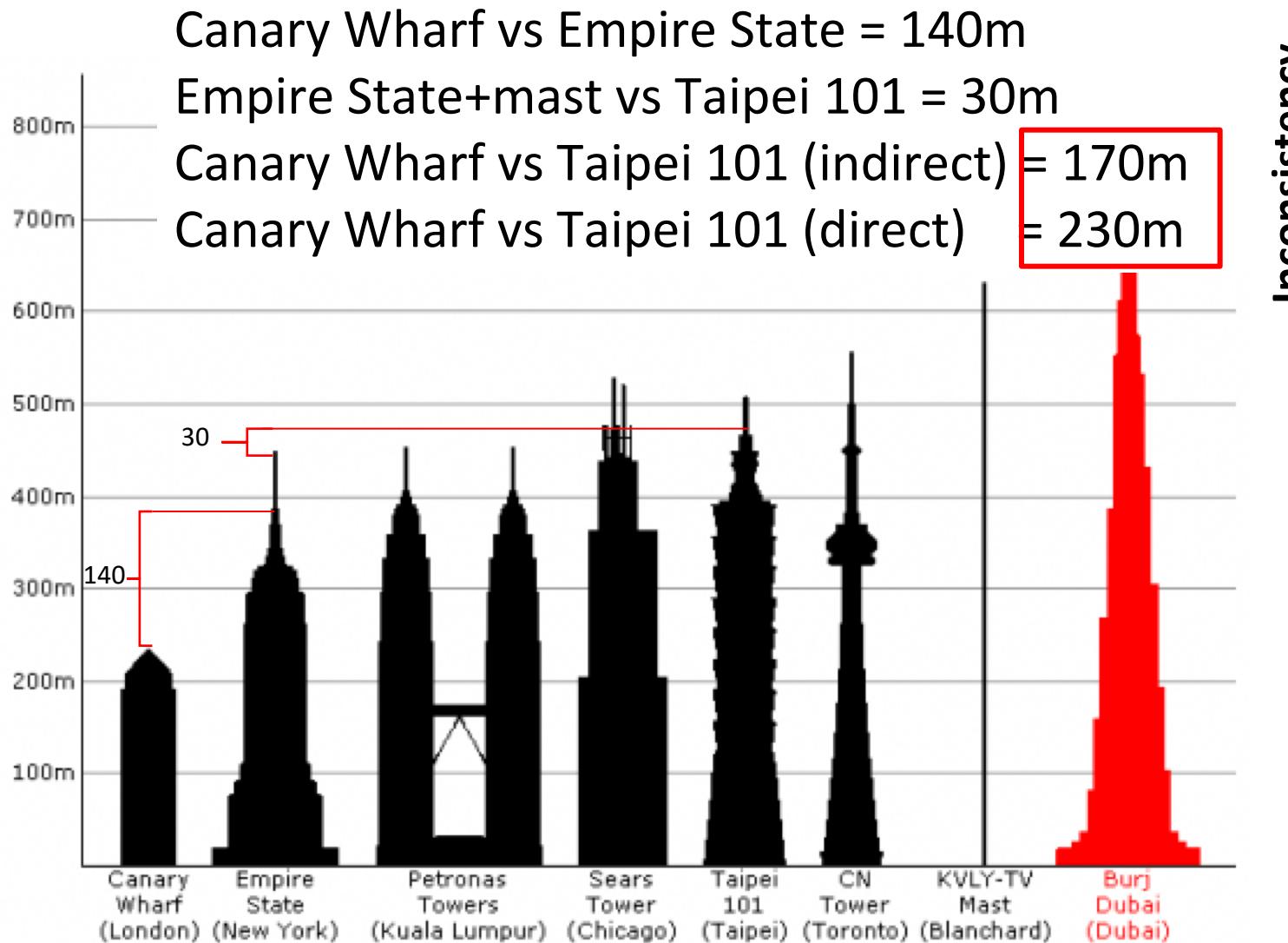
This needs to consider issues of transitivity

You have to make a statement that you assume the treatments to be jointly randomizable.

What to keep in mind when deciding which interventions to include in a network

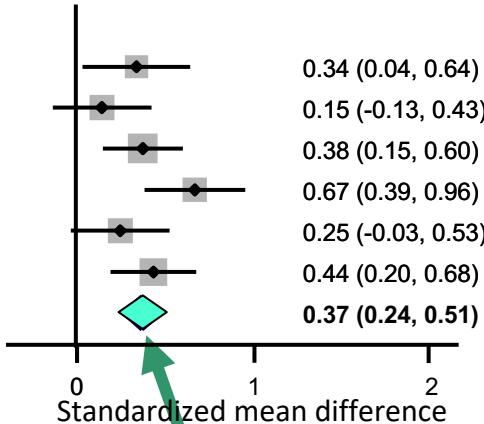
- Restricting your review to compare few interventions
 - limits its usefulness and applicability
 - you must justify your choice
 - risk to have unconnected networks
 - few data, low power (depends on the setting)
- Expanding the database too much to include many treatments
 - Jeopardizes the transitivity assumptions (or at least makes its defense challenging)
 - Renders review process long and data management difficult
- **Watch out for:** old and new treatments, ad-on treatments, intransitive legacy treatments

Inconsistent indirect comparison

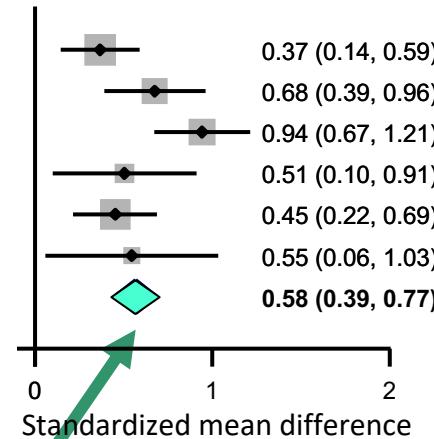


Inconsistency

QTP vs PLA



HAL vs PLA



HAL is 0.58 better than P

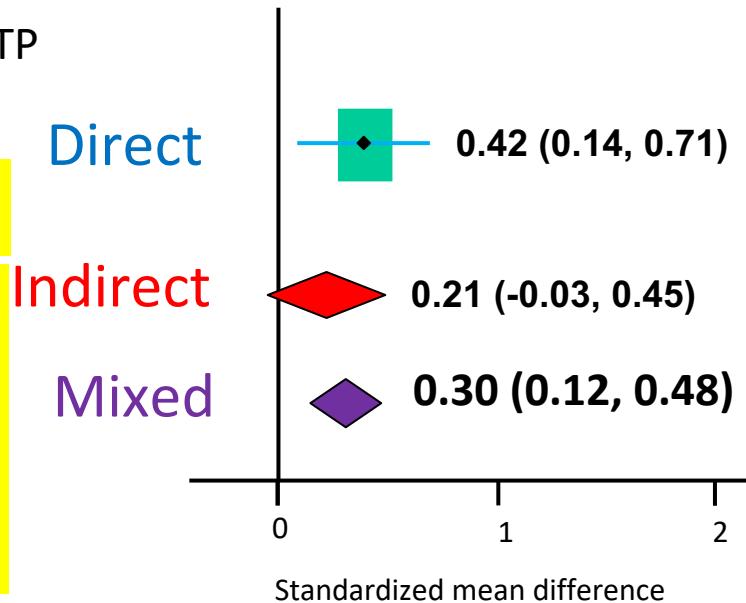
QTP is 0.37 better than P
...so HAL is 0.21 better than QTP

Are they different?

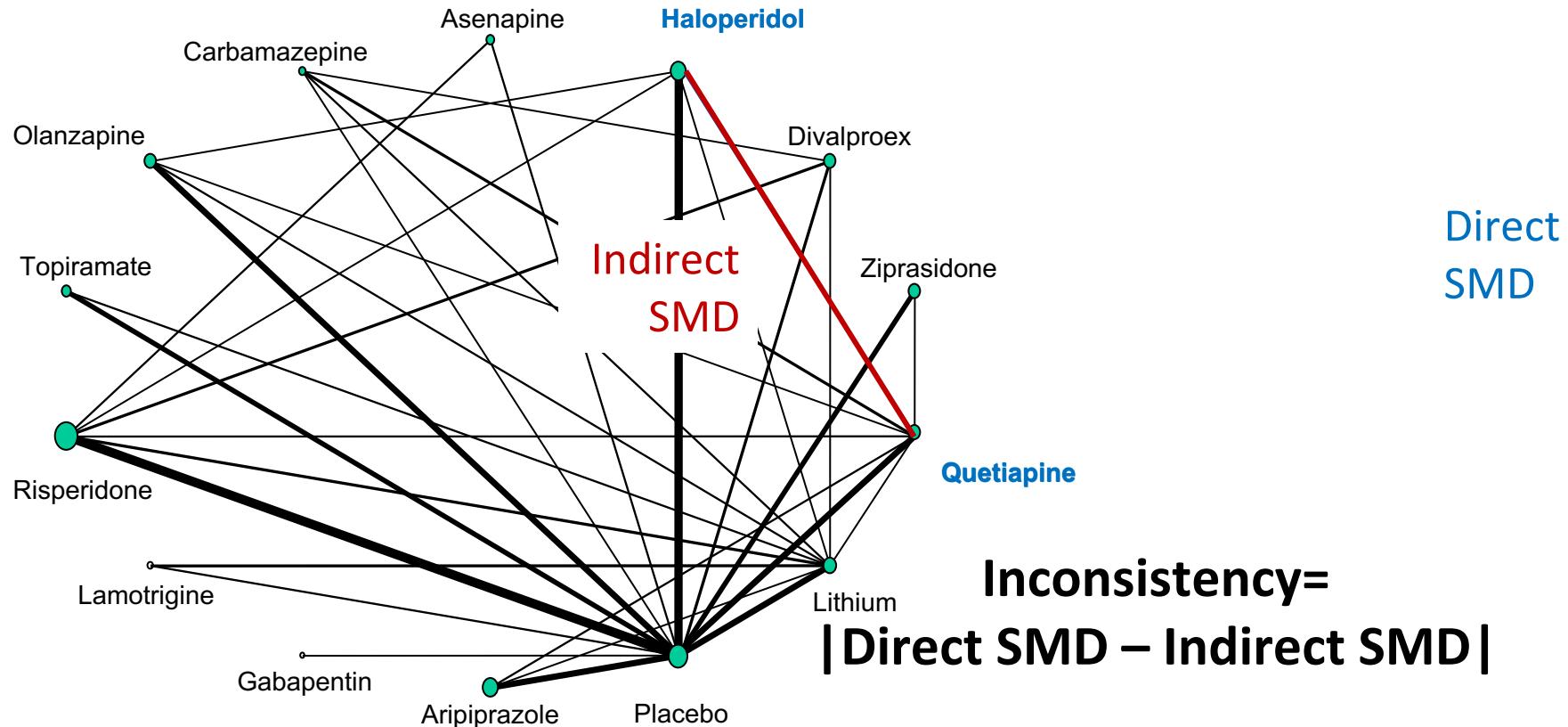
Inconsistency factor = **0.21**
(-0.15, 0.57)

You can apply a z-test: $P=0.25$

HAL vs QTP

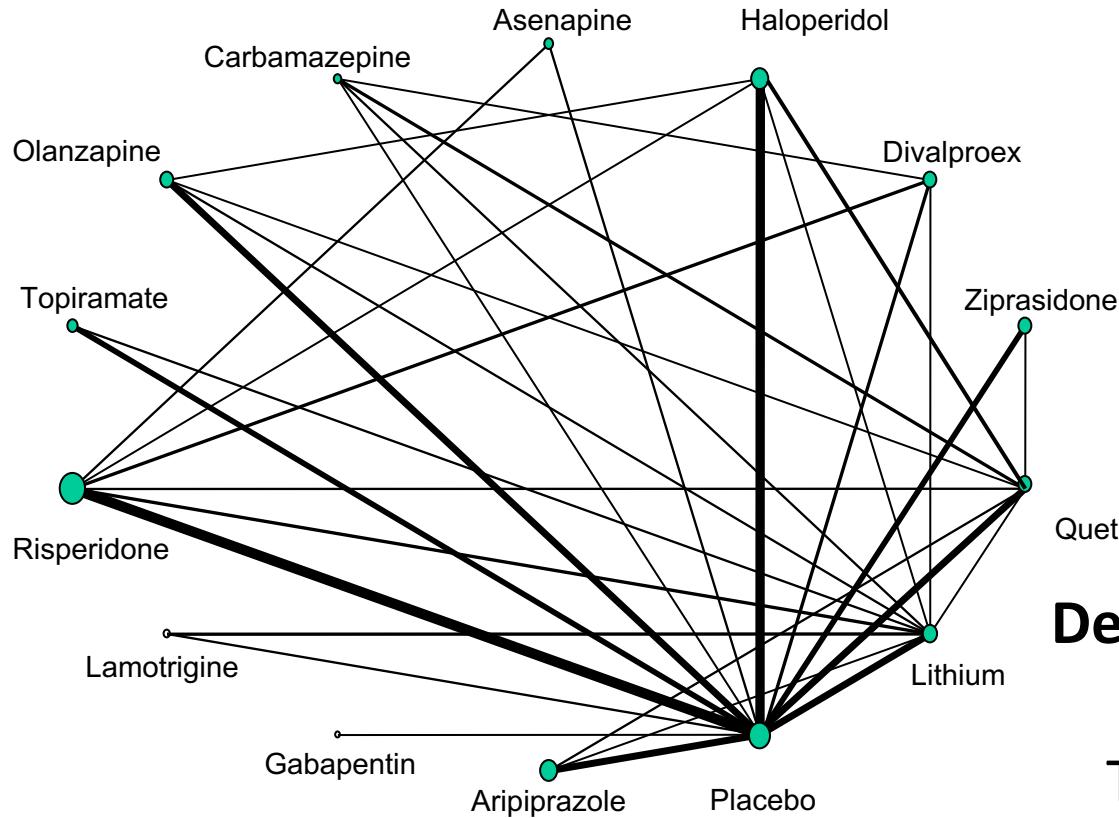


Estimating inconsistency in a network by separating direct and indirect evidence



Estimate inconsistency
for ALL comparisons

Estimating inconsistency in a network using a global test



H_0 : *Direct evidence agrees with indirect in the whole network*

Design-by-treatment test

Extension of the Q-test

The smallest the p-value, the less the support to the hypothesis

Assumption in NMA

Transitivity, congruence, consistency....

In the outset

The treatments we compare are *in principle* jointly randomizable

They have the same indication, I can imagine a mega-trial with all treatments being compared etc

When you find the studies

The groups of studies that compare them do not differ with respect to the distribution of effect modifiers

You can test this assumption if you have enough studies per comparison

When you extract the outcomes

Direct and indirect treatment effects are in statistical agreement

Various statistical tests

Issues with statistical estimation of consistency

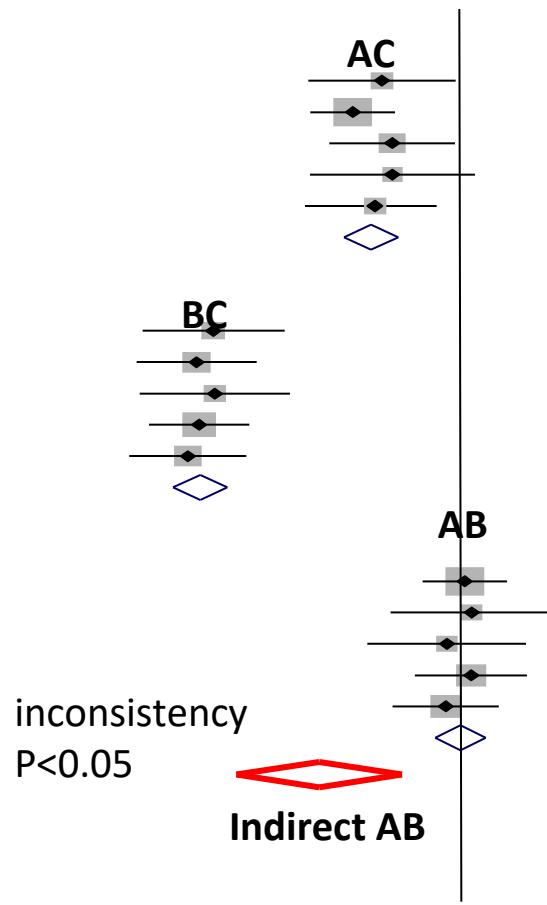
- The test for inconsistency may have very low power.
- The power for a 'typical' loop (comprising of 8 trials and about 2000 participants) to detect **a 35% relative change between direct and indirect estimation of the odds ratio** was **14%** for inverse variance method. (*Veroniki et al. BMC Med Res Methodol. 2014 19;14:106.*)
- An non-significant inconsistency test result should not be taken as proof for the absence of inconsistency
- The lack of direct evidence ('open' triangle) makes the statistical evaluation of consistency impossible
 - but the transitivity assumption is still needed to derive the indirect estimate!

Issues with statistical estimation of consistency

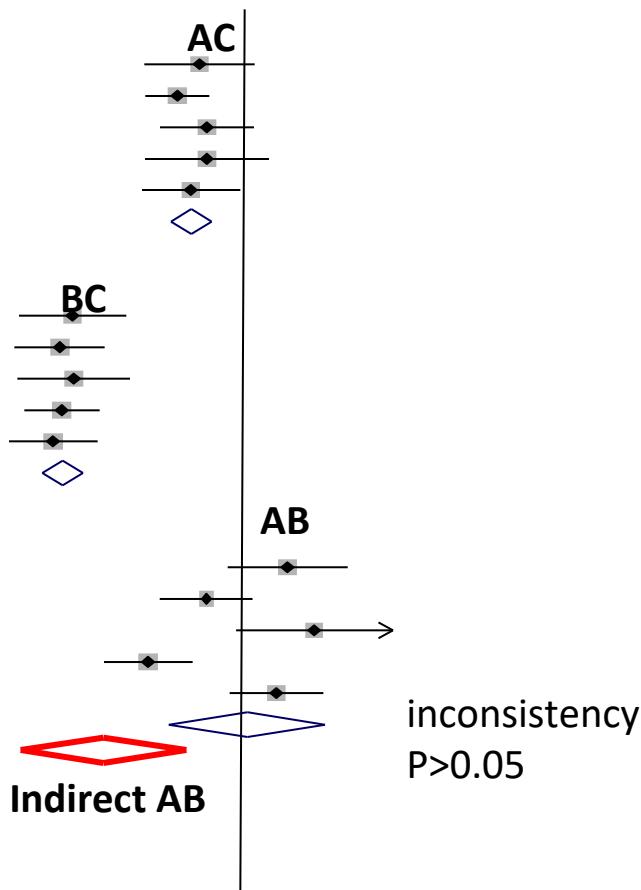
- Inference of the test depends on
 - The amount of heterogeneity
 - Whether random or fixed effects are used to derive direct estimates
 - The estimator of heterogeneity
 - Whether the same or different heterogeneity parameters are used for the three comparisons AB, AC, BC

Statistical consistency and heterogeneity

a) Fixed effects analysis



b) Random effects analysis



Don't just look at P-values: compare direct and indirect

What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Check the data	Studies that 'stand out' in the forest plot are checked for data extraction errors	Using simple loop inconsistency you can identify studies with data extraction errors. Inconsistency in loops where a comparison is informed by a single study is particularly suspicious for data errors.
Try to bypass	There is empirical evidence that some measures are associated with larger heterogeneity than others (Deeks 2002; Friedrich et al. 2011)	Empirical evidence suggests that different effect measures of dichotomous outcomes does not impact on statistical inconsistency (Veroniki et al. 2013)

What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Resign to it	Investigators may decide not to undertake meta-analysis in the presence of excessive heterogeneity	Investigators may decide not to synthesize the network in the presence of excessive inconsistency
Encompass it	Apply random-effects meta-analysis	Apply models that relax the consistency assumption by adding an 'extra' loop-specific random effect (Higgins et al. 2012, Lu & Ades 2006)*.

*However, as random effects are not a remedy for excessive heterogeneity and should be applied only for unexplained heterogeneity, inconsistency models should be employed to reflect inconsistency in the results, not to *adjust* for it.

What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Explore it	Use pre-specified variables in a subgroup analysis or meta-regression	Split the network into subgroups or use network meta-regression to account for differences across studies and comparisons. Specify the variables in the protocol, including bias-related characteristics.

Empirical evidence of inconsistency

Inconsistent loops with the loop-specific approach **303 loops**

	Loop-specific heterogeneities	Network-specific heterogeneity
<i>OR</i>	8%	5%
<i>RRharmful</i>	9%	6%
<i>RD</i>	10%	5%

Empirical evidence from 456 network meta-analyses

Characteristics of NMAs	Total	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	p-value	
<u>Evaluation of clinical and statistical assumptions</u>														
No information or discussion on transitivity	353	100%	100%	78%	92%	85%	87%	87%	78%	70%	69%	33	77%	<0.01
Reported that transitivity is likely to hold	98	0%	0%	11%	8%	15%	13%	13%	22%	28%	29%	10	23%	<0.01
Use appropriate methods to test for inconsistency	150	NA	14%	29%	33%	30%	17%	33%	38%	56%	51%	26	74%	<0.01
Discuss transitivity or inconsistency	285	0%	17%	33%	42%	44%	57%	57%	68%	69%	70%	37	86%	<0.01

Petropoulou M et al Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. *J Clin Epidemiol.* 2016

References

- 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.
- Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med.* 2013 Jul 16;159(2):130-7.
- Baker, S.G. & Kramer, B.S. 2002. The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC.Med.Res.Methodol.*, 2, (1) 13
- Bucher, H.C., Guyatt, G.H., Griffith, L.E., & Walter, S.D. 1997. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin.Epidemiol.*, 50, (6) 683-691
- Deeks, J.J. 2002. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Stat Med*, 21, (11) 1575-1600
- Donegan, S., Williamson, P., Gamble, C., & Tudur-Smith, C. 2010. Indirect comparisons: a review of reporting and methodological quality. *PLoS.One.*, 5, (11) e11054
- Song F, Xiong T, Parekh-Bhurke S et al 2011 Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ.* 2011 16;343
- Veroniki A, Vasiliadis H, Higgins J, Salanti G Evaluation of inconsistency in networks of interventions *IJE* 2013

Outline

- Systematic reviews and meta-analysis
- Approaches to meta-analysis
- Fixed-effect(s) meta-analysis
- Random-effects meta-analysis
- Heterogeneity
- Fixed vs random-effects meta-analysis
- pollev.com/gmhbe