Statistical considerations in indirect comparisons and network meta-analysis

Said Business School, Oxford, UK

March 18-19, 2013

Handout E4-L

Key assumptions in indirect comparisons and network meta-analysis

Georgia Salanti

www.mtm.uoi.gr

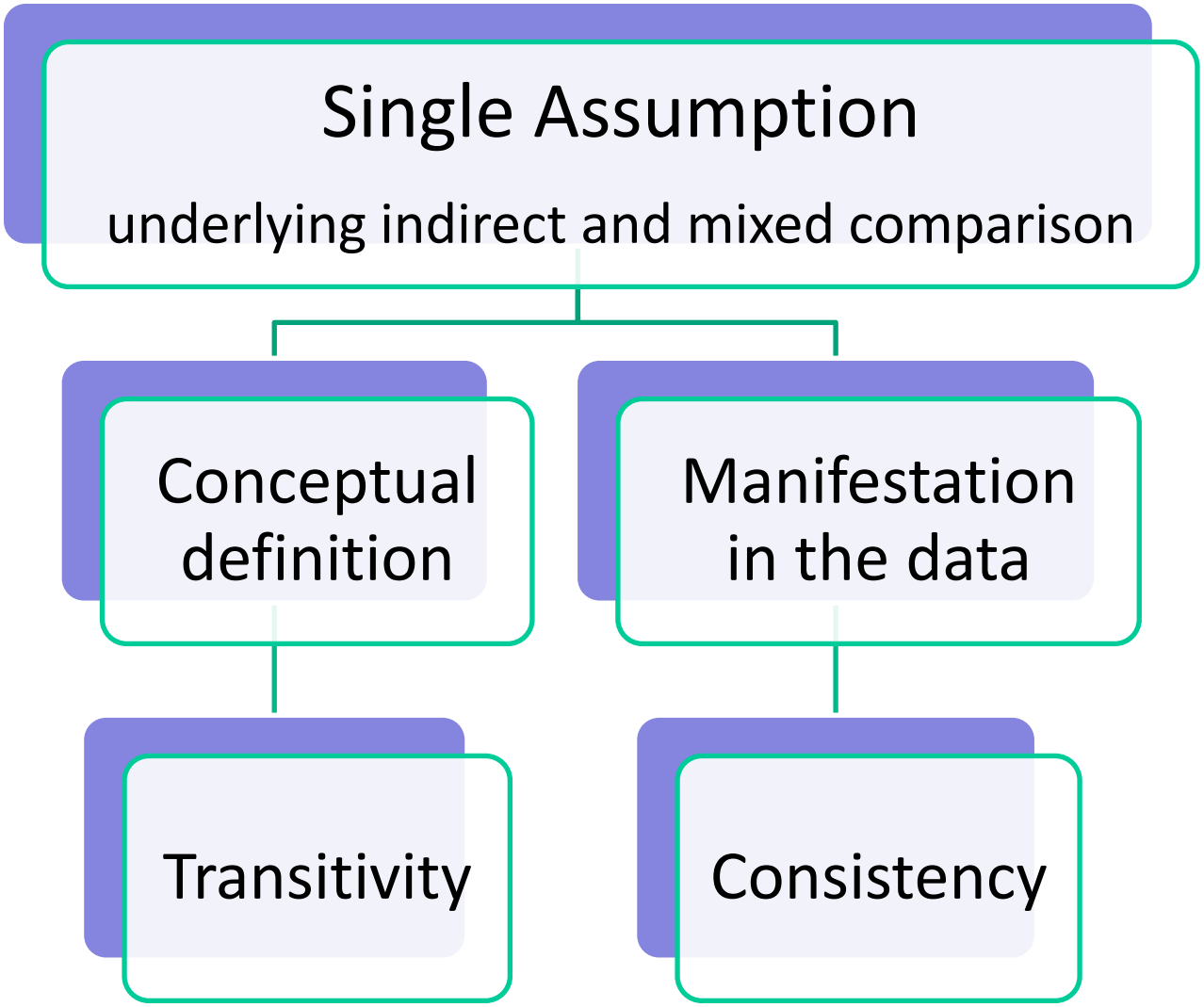
Department of Hygiene and Epidemiology

University of Ioannina School of Medicine

Criticism of indirect comparison

- Indirect comparison respects randomisation but **it is not randomized evidence**
  - The treatment comparisons have not been randomized across studies
  - Meta-regression and subgroup analysis provide **observational evidence** as the characteristic they regress on hasn't been randomized across studies
  - Indirect comparison is a special type of regression (using the comparison as explanatory variable)
- Is direct evidence preferable to indirect evidence?
- Shall we use indirect comparison only in the absence of direct evidence?

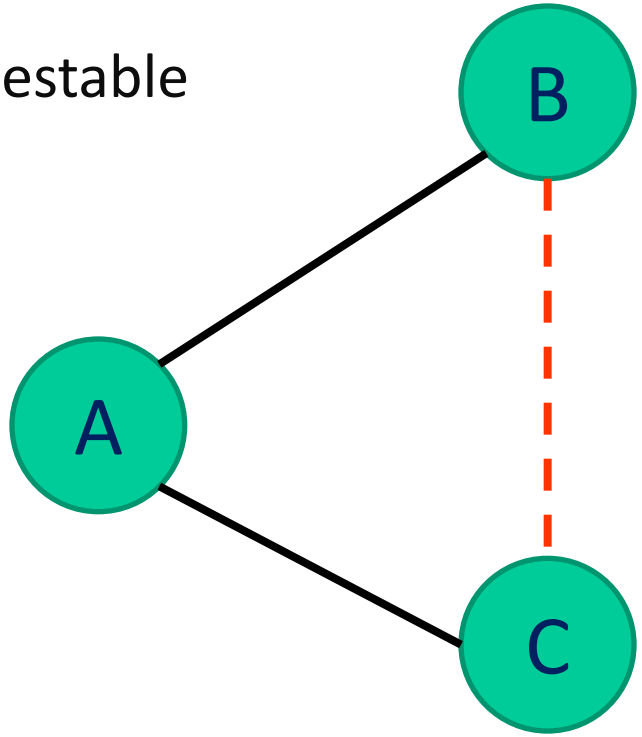
Assumption underlying indirect/mixed comparison



Transitivity

An underlying assumption when μ^{I}_{BC} is calculated is that one can learn about B versus C via A.

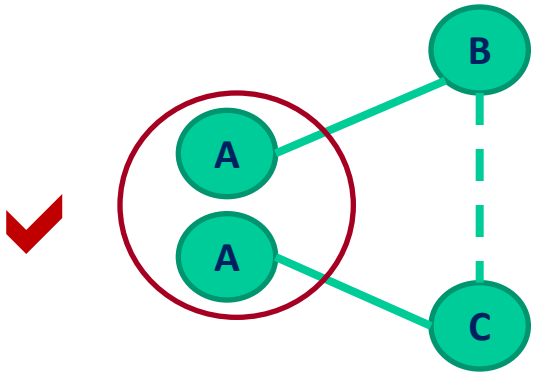
Sometime it is an untestable assumption



The anchor treatment A is 'transitive'

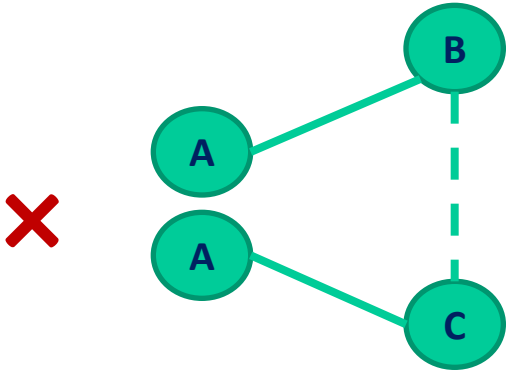
....but you can evaluate clinically and epidemiologically its plausibility

Transitivity means... (1)



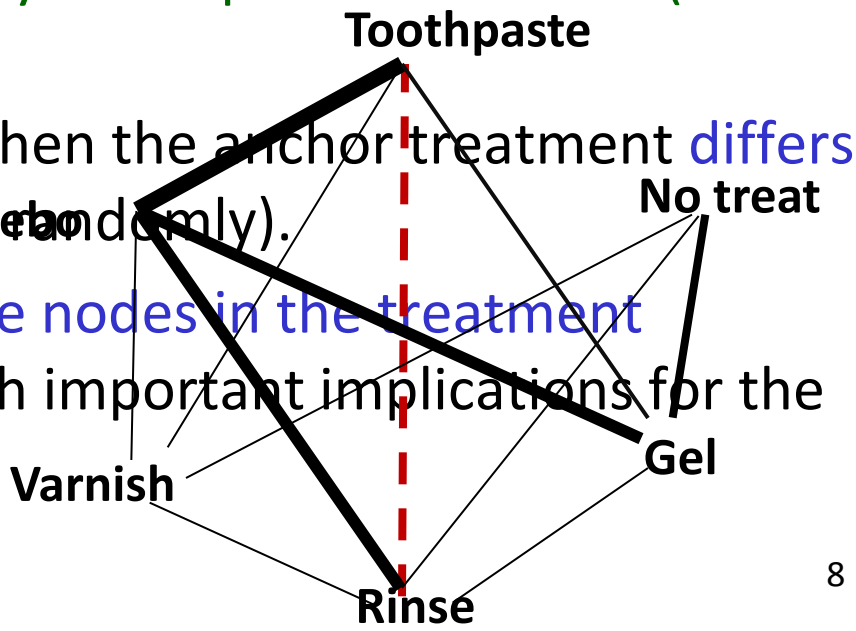
Treatment A is similar when it appears in AB and AC trials

Plausible when A is placebo given in different forms
(e.g. injection versus pill)?



Transitivity means... (1)

- **Example:** When comparing different fluoride treatments, comparison between fluoride toothpaste and fluoride rinse can be made via placebo.
  - 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 et al. JCE 2009).
- Note that transitivity is violated when the anchor treatment differs systematically between trials (not randomly).
- Consequently, the definition of the nodes in the treatment network is a **challenging issue** with important implications for the joint analysis

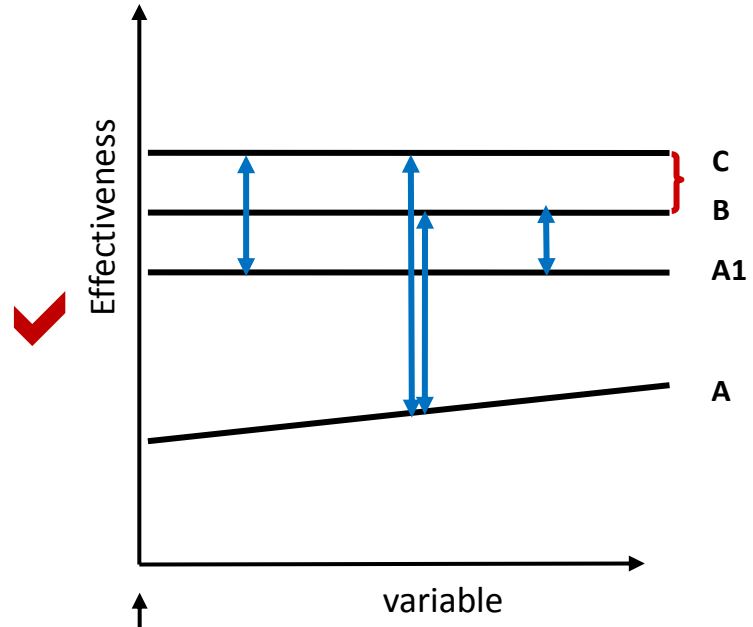


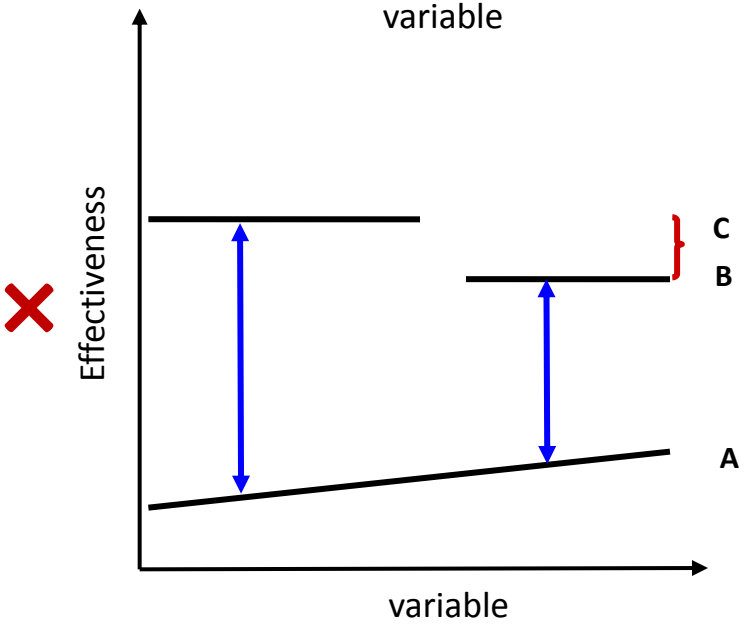
Transitivity means... (2)

- AC trials do not have B arms and AB trials do not have treatment C
- Another way to define the transitivity assumption is to consider these 'missing' arms are missing at random (Lu and Ades 2006).
- Evidence in many medical areas shows that the choice of comparator is often placebo or a suboptimal intervention rather than a realistic alternative such as an established effective treatment.

  (Heres et al. Am J Ps 2006; Rizos et al. JCE 2011;Salanti et al. Ann Int Med 2008).
- If the choice of the comparison is associated, directly or indirectly, with the relative effectiveness of the interventions then the assumption of transitivity is violated

Transitivity means... (3)





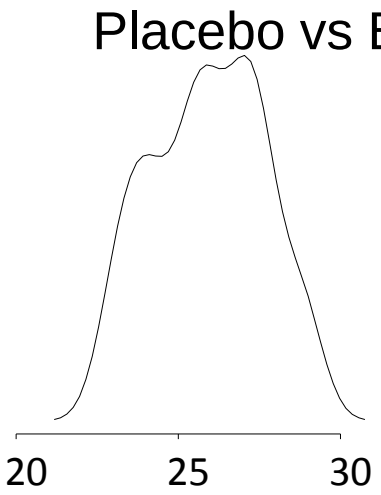
...that AC and AB trials do not differ with respect to the distribution of effect modifiers

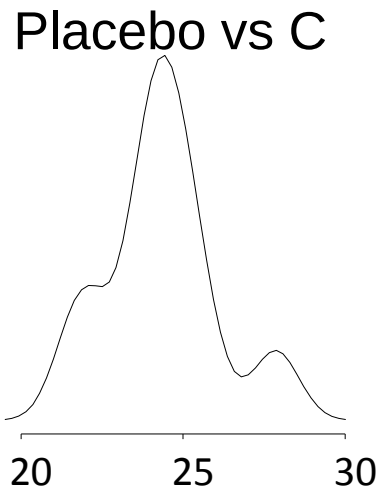
- Difficult to defend when you have older and newer treatments
- Variables are often unobserved

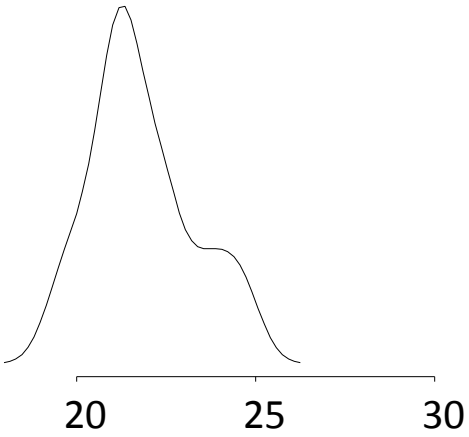
Transitivity means... (3)

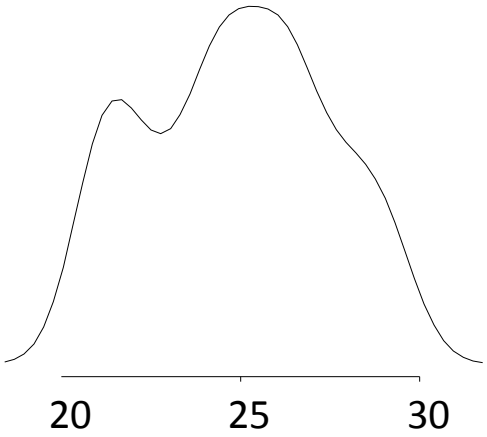
- This formulation facilitates evaluation of the transitivity assumption.
  - Distribution of effect modifiers of the relative treatment effects for similarity in AC and AB trials
- Clinicians and methodologists that aim to synthesize evidence from many comparisons should identify a priori possible effect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

Transitivity means... (3)





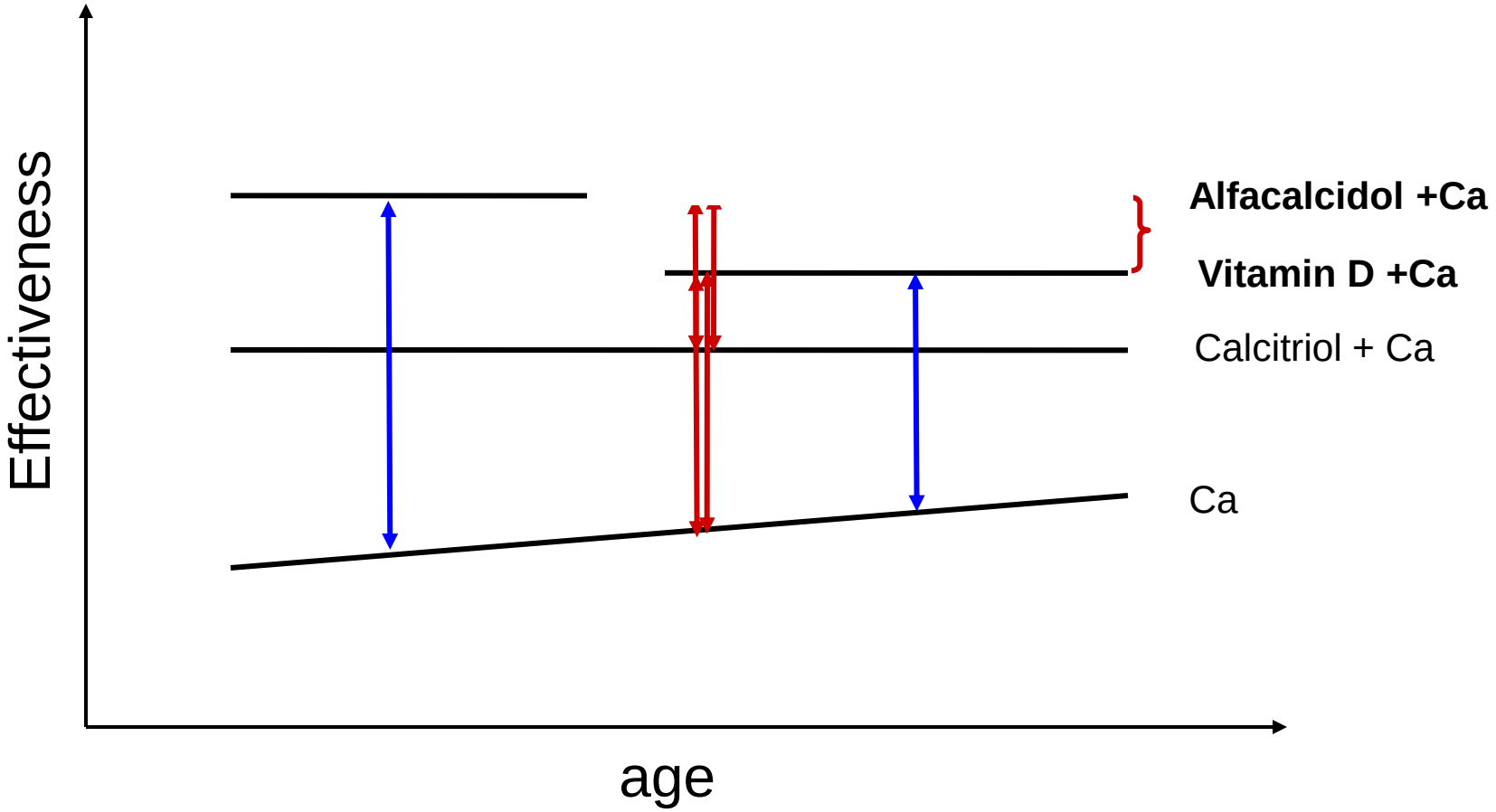




Transitivity means... (3)

- 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.
- 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 violate the transitivity assumption.
  - For example, if age is an effect modifier and AC trials differ in terms of mean age of participants (which will be presented as heterogeneity in AC studies) but the same variability is observed in the set of BC trials then transitivity may hold even if age is an effect modifier.

Transitivity means... (3)



Calcif Tissue 2005 Richy et al

Transitivity means... (4)

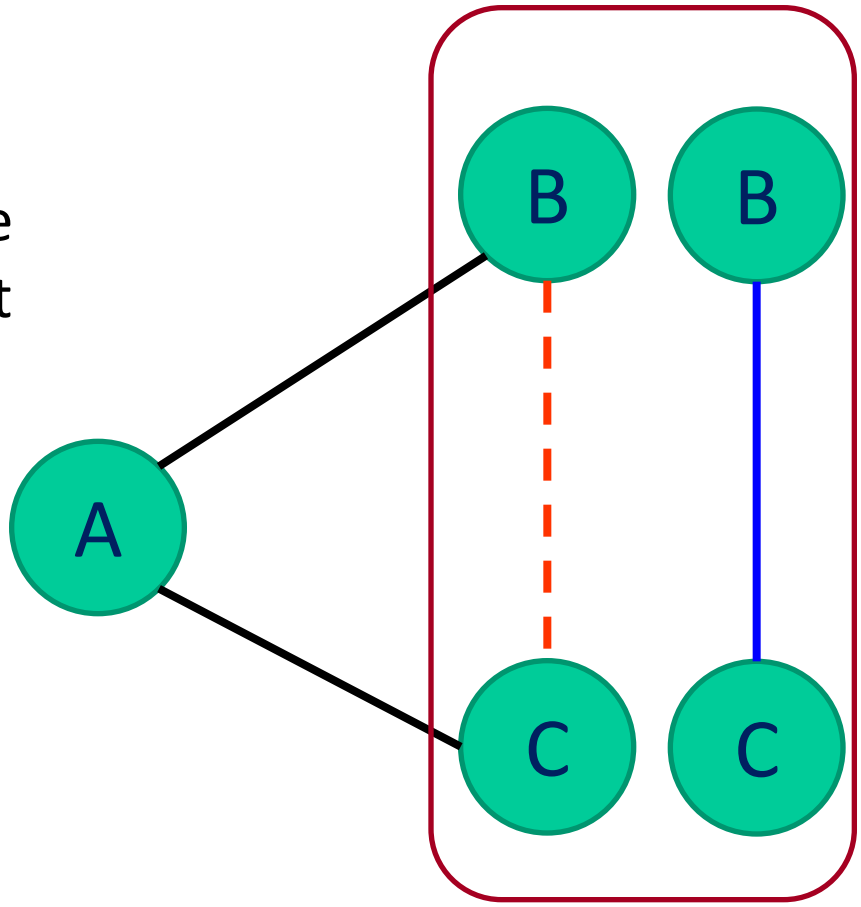
- … that all treatments are "**jointly randomizable**"
- **This consideration is a fundamental one** and should be addressed when building the evidence network
- The assumption of transitivity could be violated if interventions have different indications.
  - Ex: treatment A is a chemotherapy regimen administered as a second line treatment, whereas treatments B and C can be either as first or second line
  - we cannot assume that participants in a BC trial could have been randomized in an AC trial!
- Treatments can be comparable in theory but not in practice!
  - Ex: interferon and natalizumab are used for relapsing-remitting MS patients mitoxantrone for patients with a progressive disease.
  - However, evidence to support this clinical 'tradition' is not solid and it would be appealing to compare the three treatments.

Transitivity assumption

- 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.

Consistency

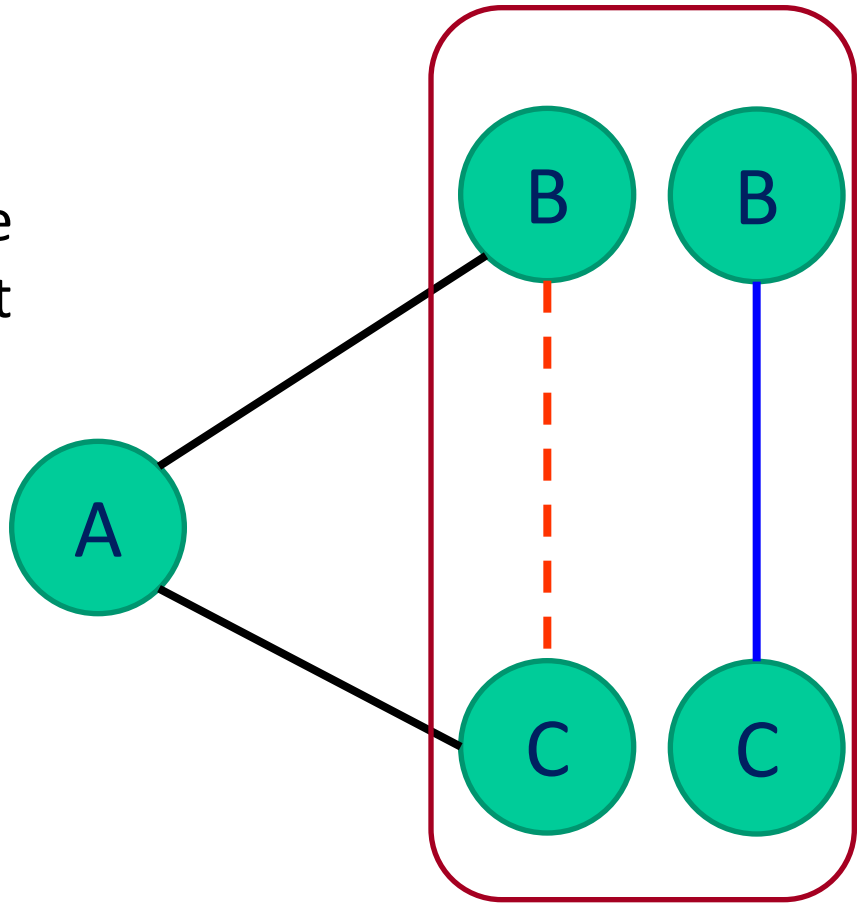
Direct and indirect evidence are in agreement



μ^{I}_{BC} μ^{D}_{BC} → μ^{M}_{BC}

Consistency

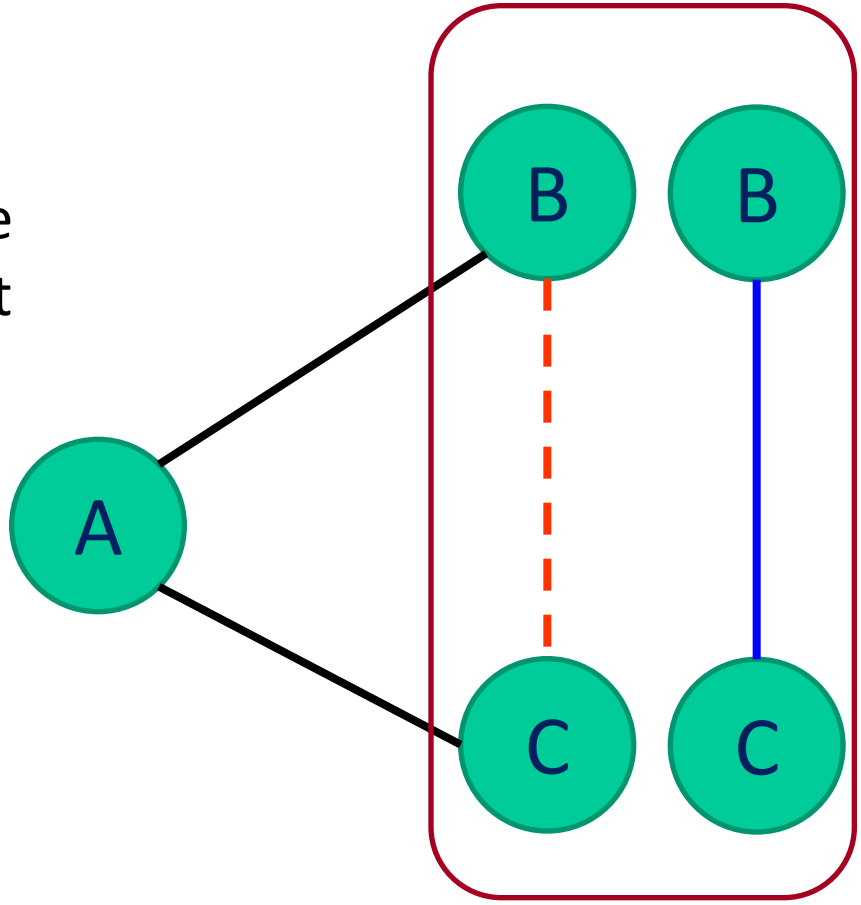
Direct and indirect evidence are in agreement



$\mu^{I}_{BC} = \mu^{D}_{BC}$

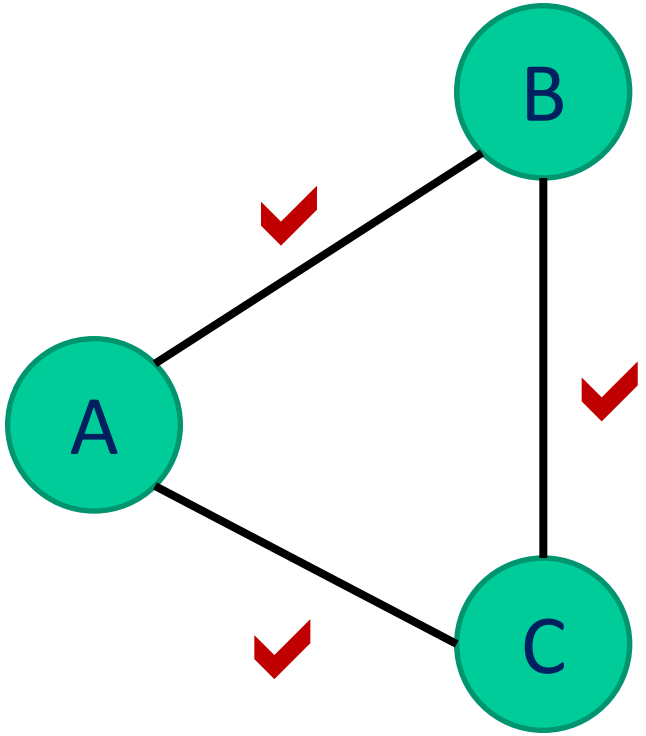
Consistency

Direct and indirect evidence are in agreement



$$\mu^{D}_{AC} - \mu^{D}_{AB} = \mu^{I}_{BC} = \mu^{D}_{BC}$$

Consistency=transitivity across a loop



In a simple triangular loop consistency holds when transitivity can be assumed for at least two out of the three nodes A, B and C (as if A and B are transitive then C is transitive as well).

Consistency means...

- That each treatment in the loop pertains to a ‘fixed’ definition independently of its comparator.
- That the ‘missing’ treatments in each trial in the loop are missing at random
- All sets of trials grouped by comparison are similar with respect to the distribution of effect modifiers

Statistical consistency

- Consistency is a property of a 'closed loop' (a path that starts and ends at the same node) or 'cycle' (as in graph theory).
- A statistically significant difference between μ^{D}_{BC} and μ^{I}_{BC} typically defines statistical inconsistency.
- Consistency can be evaluated statistically by comparing μ^{D}_{BC} and μ^{I}_{BC} in a simple z-test (often called the Bucher method).
- Alternatively, one could estimate the inconsistency as

  IF= $|\mu^{D}_{BC} - \mu^{I}_{BC}|$ (often called 'inconsistency factors')

  and its confidence interval
- If consistency holds, it may be reasonable to pool μ^{D}_{BC} and μ^{I}_{BC}

Estimating inconsistency

- In a ABC loop of evidence:

$$IF = |\mu^{I}_{BC} - \mu^{D}_{BC}| = |\mu^{I}_{AC} - \mu^{D}_{AC}| = |\mu^{I}_{AB} - \mu^{D}_{AB}|$$

$$var(IF) = var(\mu^{I}_{BC}) + var(\mu^{D}_{BC})$$

$$95\%CI : IF \pm 1.96\sqrt{var(IF)}$$

- If the 95% CI excludes zero, then there is statistically significant inconsistency
- A test for H_0: IF=0

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

Example: Gel versus Toothpaste

- Indirect SMD_{GvsT} = - 0.15 (-0.27, -0.03)
- Direct SMD_{GvsT} = 0.04 (-0.17, 0.25)
- Inconsistency factor = 0.19 (-0.05, 0.43)

- Is it important?
- You can a apply a z-test

  Z=1.55, p=0.12

Consistency in practice

- There are examples of indirect comparisons in the literature where although the key assumption has not been met, authors have formed conclusions that indirect comparisons may be valid (e.g. Chou 2006).
- Fears persist that indirect comparisons may systematically over- or under-estimate treatment effects when compared to direct (Bucher 1997, Mills 2011) .
- However, such concerns may be misplaced.
  - Given that inconsistency is a property of a 'loop' of evidence it follows that a seeming 'over-estimation' of treatment efficacy on one side of the triangle network (e.g. μ^{I}_{BC}) may represent an 'under-estimation' on another (μ^{I}_{AC})

Empirical evidence

- Song (2011) examined 112 independent 3-treatment networks and detected 16 cases of statistically significant discrepancies.
- Veroniki et al (2013) examined 315 loops and up to 10% were inconsistent
  - Depends on the estimator of heterogeneity
  - Inconsistency more probable in loops with comparisons informed by a single study
- Veroniki et al (2013) examined 40 networks and one in eight was found statistically inconsistent

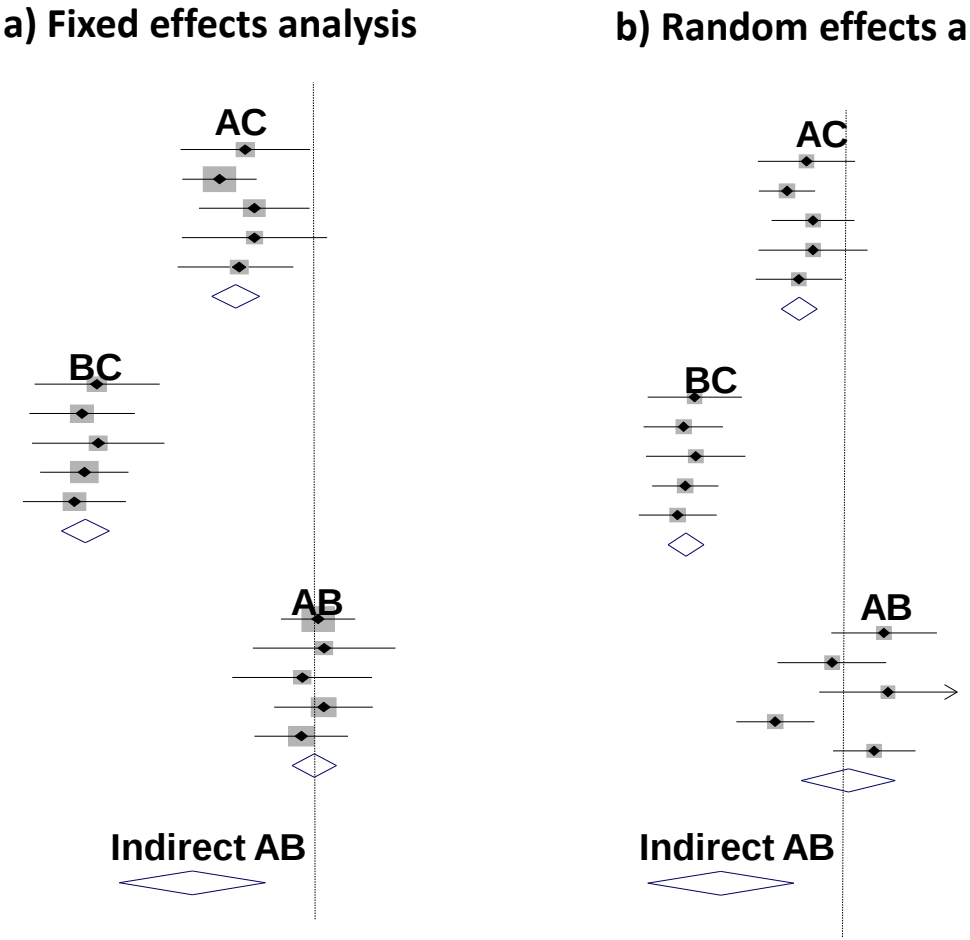
Issues with statistical estimation of consistency (1)

- In a traditional meta-analysis a statistically non-significant Q test should not be interpreted as evidence of homogeneity
- Consequently an non-significant inconsistency test result should not be taken as proof for the absence of inconsistency
  - the methodological and clinical plausibility of the consistency assumption should be further considered
- The test for inconsistency may have low power. The analyst must therefore be extremely cautious when interpreting non-significant IFs.
- 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 (2)

- 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 (MM, REML, SJ etc)
  - Whether the same or different heterogeneity parameters are used for the three comparisons AB, AC, BC

Statistical consistency and heterogeneity



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.

Summary

- The assumption of consistency underlies the indirect and mixed comparison process
- Transitivity refers to the validity of the indirect comparison and can be evaluated conceptually
- Statistical evaluation of the consistency can take place in a closed loop
- Care is needed when interpreting the results of a consistency test as issues of heterogeneity and power may limit its usefulness
- Conceptual evaluation of the consistency assumption should include
  - Checking for effect modifiers that differ across comparisons
  - Checking the definition of each node/treatment
  - Checking the 'random' choice of comparators

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.
- 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
- Mills, E.J., Ghement, I., O'Regan, C., & Thorlund, K. 2011. Estimating the power of indirect comparisons: a simulation study. *PLoS.One.*, 6, (1) e16237
- 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