Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Journal Club Presentation

Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies‡

Presenter: Natthapong Hongku, M.D.

Commentator: Napaphat Poprom

Academic Coordinator: Pawin Numthavaj, M.D., Ph.D.

19 January 2018

Wisdom of the Land

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Special Issue Paper

Research Synthesis Methods

Received 22 June 2011, Revised 11 May 2012, Accepted 15 May 2012 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1044

Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies‡

J. P. T. Higgins,[a,b*†] D Jackson,[a] J. K. Barrett,[a] G Lu,[c] A. E. Ades[c] and I. R. White[a]

Outline of talks

- **Introduction**
- Concepts of heterogeneity and inconsistency
- Models for inconsistency
- Discussion and Conclusion

Introduction

- Compare multiple treatments
  -> More useful than pairwise comparisons alone
  -> Ranking of treatment interventions

- Network meta-analyses (NMA)
  -> Simultaneous analysis of both direct and indirect comparisons among multiple treatments across multiple studies

GleserandOlkin,1994;HigginsandWhitehead,1996;Lumley,2002; Lu and Ades, 2004; Caldwell et al., 2005; Lu and Ades, 2006; Salanti et al., 2008; Ioannidis, 2009)

- Component of NMA

-> Assessment of the different of comparable sources of evidence both substantively and statistically

- Simple indirect comparison will be confounded

-> If studies involving one of interested treatments are fundamentally different from the studies involving the other treatment

"Termed incoherence or inconsistency in the literature"

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Outline of talks

- Introduction
- **Concepts of heterogeneity and inconsistency**
- Models for inconsistency
- Discussion and Conclusion

Concepts of heterogeneity and inconsistency

- **Indirect comparisons**
- Heterogeneity
- Consistency
- Loop inconsistency
- Multi-arm trials
- Design inconsistency
- Design-by-treatment interaction

Indirect comparisons

Indirect comparison

- **Maintains** benefits of randomization within each trial

- Provided that the differences affect **only** prognosis of participants and not their response to treatment

Indirect comparisons

- Trial 1: Two-arm trial of 'B–A' comparison

  [$\hat{\delta}_1^{AB}$ = Estimated effect sizes of trial 1]

- Trial 2: Two-arm trial of 'C–B' comparison

  [$\hat{\delta}_2^{BC}$ = Estimated effect sizes of trial 2]

- Then, indirect comparison of 'C–A' as….

$$\hat{\delta}_{(\text{indirect})}^{AC} = \hat{\delta}_1^{AB} + \hat{\delta}_2^{BC}$$

Indirect comparisons

On assumption….

- B treatment as same in both trials
- When ‘B–A’ and ‘C–B’ are added together
  -> B effects **are cancelled out**
- Indirect comparison
  -> reflective of difference between A and C is not testable in absence of further information

Indirect comparisons

- Third trial of 'C–A' (yielding result $\hat{\delta}_3^{AC}$)

  -> Compare the indirect comparison with a direct comparison

- Network of 3 trials is consistent

  If underlying treatment effects are related as follows:

$$\delta_3^{AC} = \delta_1^{AB} + \delta_2^{BC}, \qquad (1)$$

δ_1^{AB}, δ_2^{BC} and δ_3^{AC} -> represent true effects underlying 3 studies

Indirect comparisons

Equation (1) -> very unlikely to hold for particular set of 3 trials because….

- In terms of heterogeneity:
  Within each treatment comparison, each individual study is **not fully representative** of all studies of that particular comparison

- In terms of inconsistency:
  Across treatment comparisons, there are important differences in types of studies contributing to comparisons

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Concepts of heterogeneity and inconsistency

- Indirect comparisons
- **Heterogeneity**
- Consistency
- Loop inconsistency
- Multi-arm trials
- Design inconsistency
- Design-by-treatment interaction

Heterogeneity

- Multiple studies of same research question have different values of estimated effect measurement.
- By holding treatment comparison constant and varying study index.

- Heterogeneity may be present for comparison 'B–A'

  if $\delta_i^{AB} \neq \delta_j^{AB}$ for some pair of studies i and j

Heterogeneity

- Heterogeneity has been **argued** to be inevitable in meta-analysis…

: 2 trials of same pairwise comparison **are unlikely** to have equal underlying treatment effects

-> Equality in equation 1 is questionable

[Particular instance of ‘C–A’ investigated in trial 3
is unlikely to represent all instances of ‘C–A’ comparisons]

Heterogeneity

- Random-effects model

  -> Common way to allow for heterogeneity

- Assumes that underlying effects in multiple studies of same comparison come from normal distribution

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Concepts of heterogeneity and inconsistency

- Indirect comparisons
- Heterogeneity
- **Consistency**
- Loop inconsistency
- Multi-arm trials
- Design inconsistency
- Design-by-treatment interaction

Consistency

Consistency equation

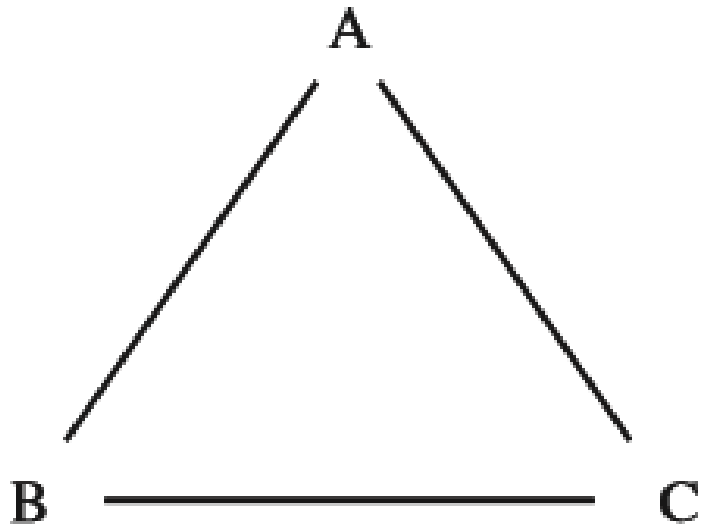
$$\delta^{AC} = \delta^{AB} + \delta^{BC}$$

- Show **desirable relationship** between direct and indirect sources of evidence for **single comparison**

Consistency

- Network with only two-arm trials

  Triangle of relationships with 3 (nontouching) solid edges



Consistency in 2 arm trials

- Each edge represents 1 or >2 arm trials comparing 2 treatments identified at either end of edge
- All 3 edges used same line style (solid line)
  -> There is no conflict (inconsistency)

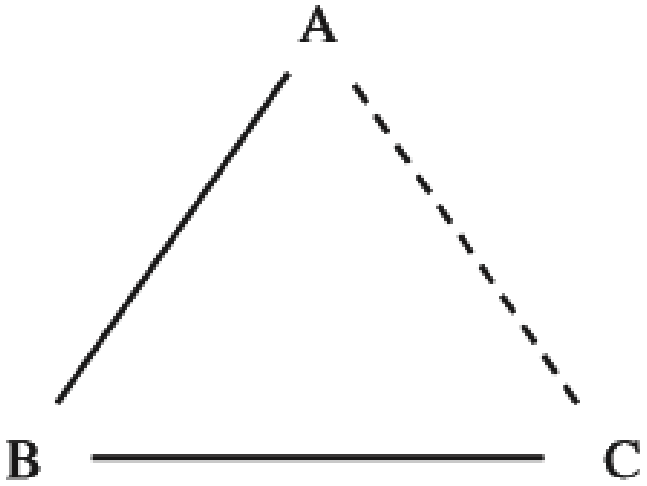
Concepts of heterogeneity and inconsistency

- Indirect comparisons
- Heterogeneity
- Consistency
- **Loop inconsistency**
- Multi-arm trials
- Design inconsistency
- Design-by-treatment interaction

Loop Consistency

Studies of different treatment comparisons
-> Different in ways that affect effect sizes

Drawing edges in different line styles



Evidence on direct comparison AC (dashed line) **conflicts with** evidence drawn via indirect comparison AB, BC (solid lines)

Loop inconsistency in 2 arm trials

Loop Consistency

- Occur only at least 3 separate sets of studies making different comparisons

  (For example, 'B–A', 'C–A' and 'C–B' studies)

Or

- When both indirect and direct estimates of effect size are available

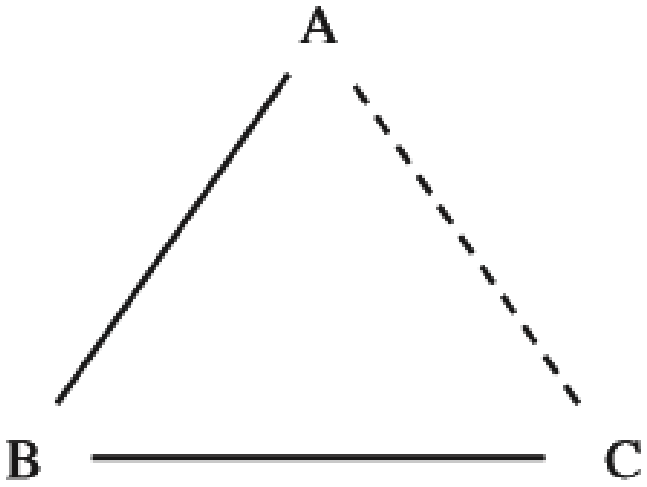
  (For example, when 'C–B' is measured both directly and via 'A' as indirect estimate)

Causes of Loop Consistency

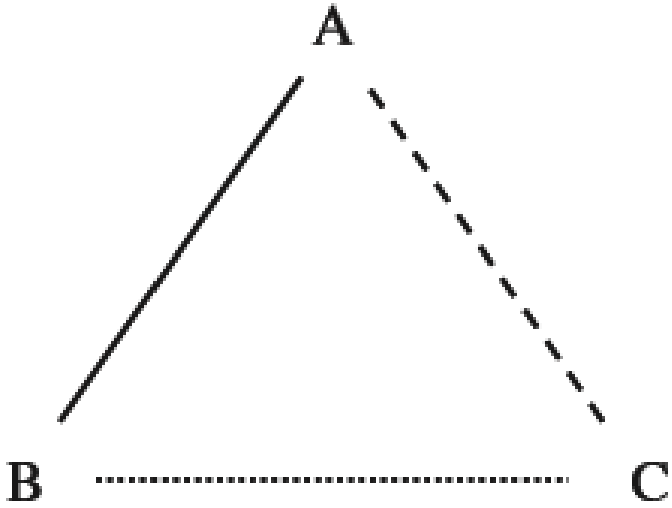
: Participants in head-to-head studies ‘C–A’ are **different** from studies ‘B–A’ and ‘C–B’ (Because they are contraindicated for treatment B)

: Versions of treatment B are different in studies ‘B–A’ and ‘C–B’ (Because of different doses) that are associated with magnitude of treatment effect -> Sum of ‘B–A’ and ‘C–B’ not equal to ‘C–A’

: Studies of different comparisons were different periods, different settings or different contexts (studies ‘C–B’ = recent but studies A = old historical standard)



Loop inconsistency in
2 arm trials



Loop inconsistency in
2 arm trials (Alternative)

- 3 edges could be drawn in 3 different line styles to indicate that different effect modifiers are associated with each edge in loop
- Difference cannot be tested statistically & have to be informed by expert judgement

Concepts of heterogeneity and inconsistency

- Indirect comparisons
- Heterogeneity
- Consistency
- Loop inconsistency
- **Multi-arm trials**
- Design inconsistency
- Design-by-treatment interaction

Multi-arm trials

- Studies with > 2 treatment arms
- Network of multi-arms trial may be **consistent**

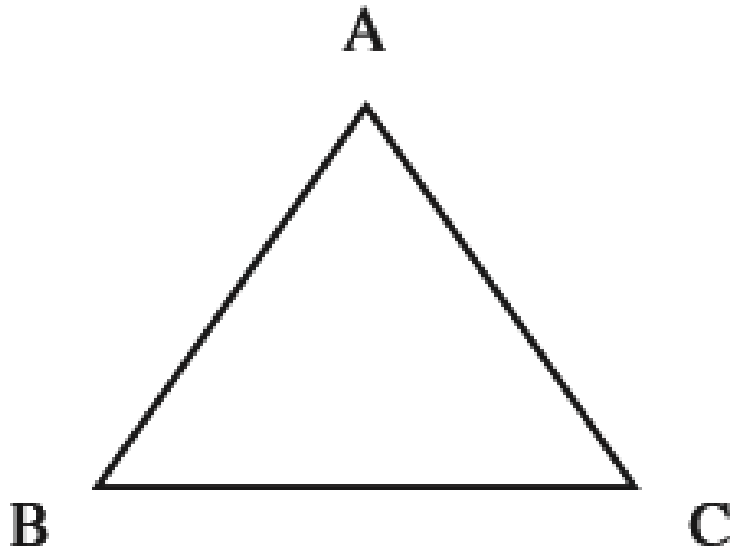
: By structurally (because all studies include all treatments) or…

: By observation (assumptions around equality of direct & various indirect comparisons hold across studies) or…

: Combination of both

Multi-arm trials

- Loop inconsistency cannot occur within multi-arm trial



Consistency is considered in 3 arms trials using closed (joined-up) polygon

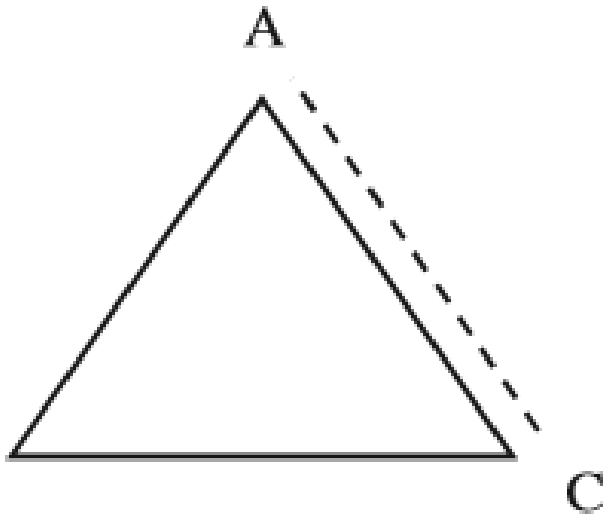
Concepts of heterogeneity and inconsistency

- Indirect comparisons
- Heterogeneity
- Consistency
- Loop inconsistency
- Multi-arm trials
- **Design inconsistency**
- Design-by-treatment interaction

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Design consistency

- 'Design' of a study = Set of treatments compared within study
- Design inconsistency = Differences in effect sizes between studies involving different sets of treatments
- Potential conflicts between study designs represented by different line styles



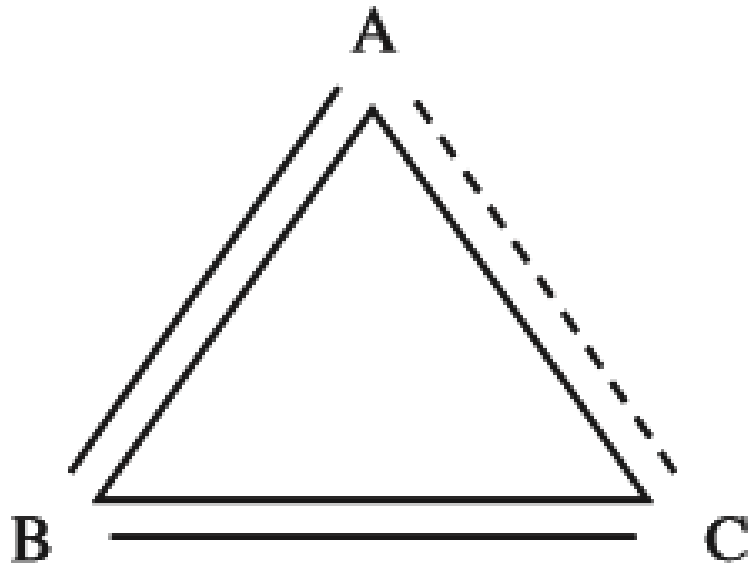
AC effect size in 3-arm trials (drawn as solid line) differs from 2-arm trials (drawn as dashed line)

Design consistency

When multi-arm trial is involved
: loop inconsistency in two-arm trials indicates design inconsistency

Because multi-arm trial must be internally consistent
-> difference between effect sizes from multi-arm trial

Design consistency

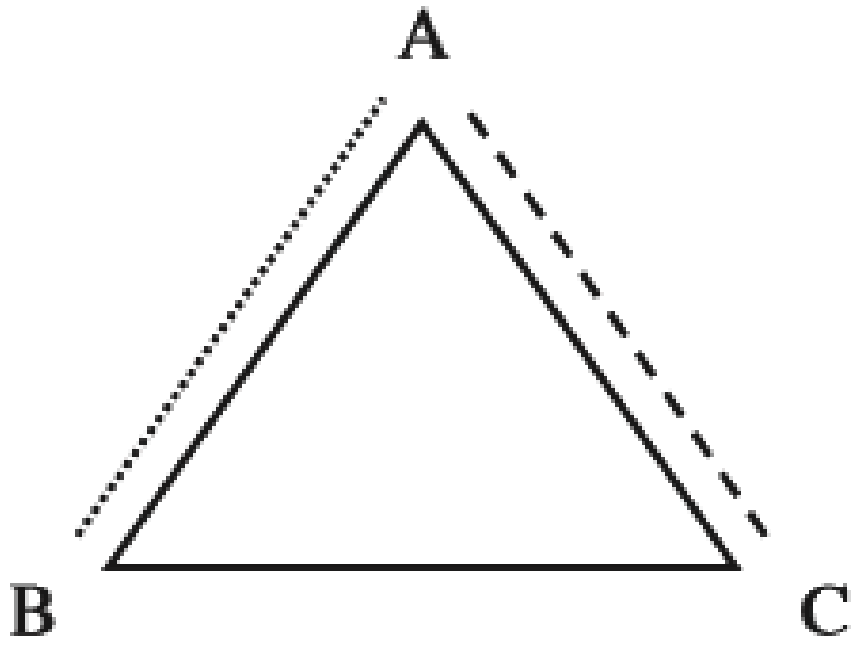


Design inconsistency and loop inconsistency

: Pairwise trials display loop inconsistency

whereas 3-arm trials conflicts with at least 1 pairwise trial reflecting design inconsistency

Design consistency

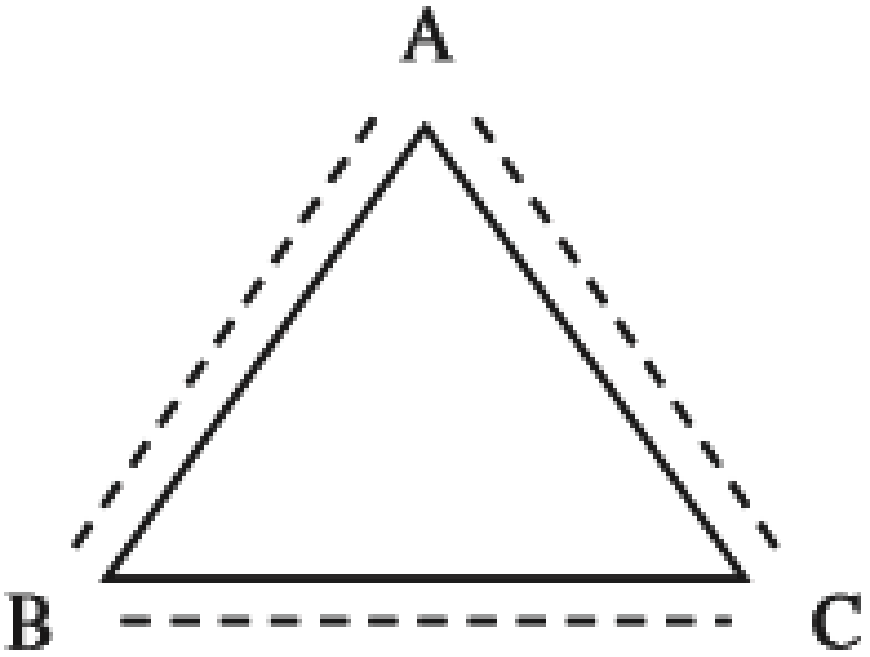


Design inconsistency with one 3-arm trial & two 2-arm trials

- Loop can be constructed by isolating BC comparison from 3-arm trial and comparing it with two-arm trials
- However, this ignores presence of consistent loop within 3-arm trials

-> unclear we should describe this network as displaying loop inconsistency

Design consistency



- Two-arm trials
  -> consistent among themselves, but effect sizes differ from those of multi-arm trial
- Does this display design inconsistency without loop inconsistency?

Design inconsistency without loop inconsistency?

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Concepts of heterogeneity and inconsistency

- Indirect comparisons
- Heterogeneity
- Consistency
- Loop inconsistency
- Multi-arm trials
- Design inconsistency
- **Design-by-treatment interaction**

Design-by-treatment interaction

- Loop inconsistency reflects important question of whether direct and indirect evidence agree with each other
- Design inconsistency reflects less substantive interest question; particular choice of treatments in study is associated with different effect sizes for particular contrasts

Presence of multi-arm trials, distinction between two types is difficult

Design-by-treatment interaction

- Meta-regression approach for design inconsistency
- Method describes by Lu and Ades for loop inconsistency
- We argue case for statistical model that encompasses both types of inconsistency
- This is a model that includes the full set of design-by-treatment interaction terms

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Outline of talks

- Introduction
- Concepts of heterogeneity and inconsistency
- **Models for inconsistency**
- Discussion and Conclusion

Models for inconsistency

- **Design-by-treatment interaction model**
- Lumley model
- The Lu–Ades model

Design-by-treatment interaction model

- Full model for design-by-treatment interaction

  J=A, B, ... index treatments

  d=1, ... index designs

  i=1, ... index studies

Consider parameter μ_{di}^{AJ} reflect treatment effect

compare treatment J with reference treatment A in study i

Model for treatment effect parameters: not necessary for treatment A to have been included in every study

$$\mu_{di}^{AJ} = \delta^{AJ} + \beta_{di}^{AJ} + \omega_{d}^{AJ},$$

$$\mu_{di}^{AJ} = \delta^{AJ} + \beta_{di}^{AJ} + \omega_{d}^{AJ},$$

- Where δ^{AJ} is fixed effect of treatment J relative to A

- β_{di}^{AJ} is study-by-treatment interaction term to reflect standard heterogeneity
  (variability in treatment effects for comparison AJ, within studies of design d)

- ω_{d}^{AJ} is design-by-treatment interaction term to reflect inconsistency (variability between designs)

$$\mu_{di}^{AJ} = \delta^{AJ} + \beta_{di}^{AJ} + \omega_{d}^{AJ},$$

- This model has largest number of degrees of freedom (d.f.) among models allowing for both loop and design inconsistency

- Inconsistency parameter ω_{d}^{AJ} describes perturbation in AJ comparison in specific study design D

Design-by-treatment interaction model

- Prefer fixed effects because…..

: Common distribution assumption in random-effects formulation is impossible

: Facilitates straightforward test of null hypothesis of consistency throughout network of comparisons

: Easier and less sensitive to reparameterization

: There are often too few inconsistency parameters for random-effects model to produce reliable inferences on random-effects variance parameter

Design-by-treatment interaction model

Disadvantage of fixed-effects approach

: Constraints are required on ω_d^{AJ} terms in order to avoid over-parameterization

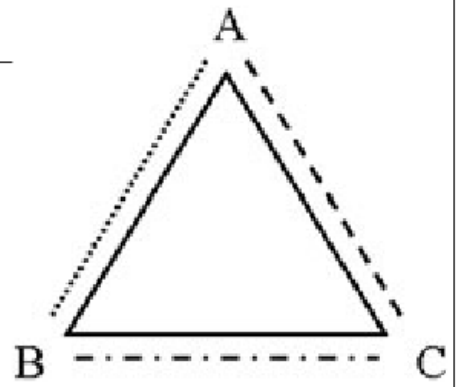
-> Interpretation of δ^{AJ} parameters is not straightforward depends on parameterization of model

Design-by-treatment interaction model

- Example, Network with 4 designs AB, AC, BC and ABC

Table 1. Design-by-treatment interaction model for three treatments (all possible designs). Heterogeneity terms have been omitted.

Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	$\delta^{AB} + \omega_2^{AB}$	—
AC	Ref	—	$\delta^{AC} + \omega_3^{AC}$
BC	Ref	δ^{AB}	$\delta^{AC} + \omega_4^{AC}$



: This network has potential for 3 conflicts & has 3 d.f. for inconsistency

: Reflected 4 different line styles

Design-by-treatment interaction model

Three potential conflicts may be parameterized as follows:

- Difference in 'B–A' effects between AB and ABC studies ω_2^{AB}
- Difference in 'C–A' effects between AC and ABC studies ω_3^{AC}
- Difference in 'C–B' effects between BC and ABC studies which could be placed on either treatment B or C

(we adopt convention of placing it on last of such possibilities ω_4^{AC})

For example, following separates out parameter to represent loop inconsistency

: Loop inconsistency in two-arm trials, by contrasting direct evidence 'B–A' in AB studies with indirect evidence involving 'C–A' from AC studies and 'C–B' from BC studies

$$\omega^{(1)} = \left(\delta^{AB} + \omega_2^{AB}\right) - \left(\left[\delta^{AC} + \omega_3^{AC}\right] - \left[\left(\delta^{AC} + \omega_4^{AC}\right) - \delta^{AB}\right]\right) = \omega_2^{AB} - \omega_3^{AC} + \omega_4^{AC}$$

: Design inconsistency, by contrasting 'B–A' effects between AB and ABC studies

$$\omega^{(2)} = \omega_2^{AB}$$

: Design inconsistency, by contrasting 'C–A' effects between AC and ABC studies

$$\omega^{(3)} = \omega_3^{AC}$$

Models for inconsistency

- Design-by-treatment interaction model
- **Lumley model**
- The Lu–Ades model

Lumley model

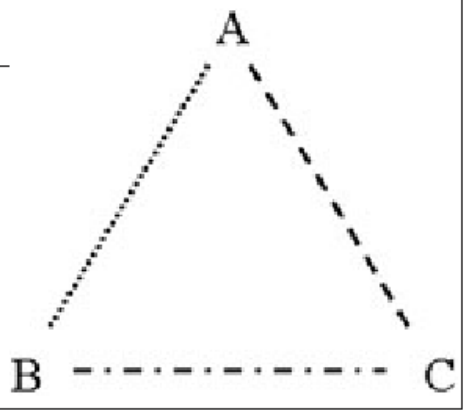
- Full design-by-treatment interaction model proposed by Lumley in 2002
- Inconsistency factors follow random-effects distribution
- **However,** model was constructed **only for two-arm trials**, & not extension to multi-arm trials

Lumley's model for simple network of 3 pairwise comparisons

Table 2. Lumley model for three treatments (applicable only to two-arm trials).

Design	A	B	C
AB	Ref	$\delta^{AB} + \omega_1$	—
AC	Ref	—	$\delta^{AC} + \omega_2$
BC	Ref	δ^{AB}	$\delta^{AC} + \omega_3$

Assumption for inconsistency factors: $\omega_d \sim N(0, \sigma_\omega^2)$



- Network of only pairwise studies

-> Inconsistency parameters viewed as being clearly to specific pairwise comparisons

- Design-by-treatment interaction model present only 1 inconsistency factor for data set

Models for inconsistency

- Design-by-treatment interaction model
- Lumley model
- **The Lu–Ades model**

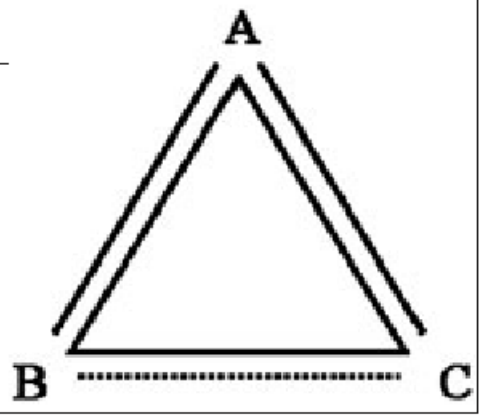
The Lu–Ades model

Described by Lu and Ades (2006)

- Motivated primarily by loop inconsistency
- One inconsistency parameter is added for each independent closed loop in evidence network
  (not including loops created only by multi-arm trials)

Table 3. Lu and Ades model for three treatments with order *A*, *B*, *C* (all possible designs).

Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	δ^{AB}	—
AC	Ref	—	δ^{AC}
BC	Ref	δ^{AB}	$\delta^{AC}+\omega^{AC}$



The Lu–Ades model

- Not guarantee that all possible independent closed loops are identified
- Lu and Ades ensured that they include all closed loops

: by careful selection of modelled treatment contrasts from multi-arm trial

- For example

: There are AB, AC and ABC studies

: Ensured that BC contrast is among two modelled contrasts from ABC study

[because it forms closed loop with two-arm studies]

The Lu–Ades model

Contains subset of inconsistency parameters from design-by-treatment interaction model

Table 1. Design-by-treatment interaction model for three treatments (all possible designs). Heterogeneity terms have been omitted.

Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	$\delta^{AB} + \omega_2^{AB}$	—
AC	Ref	—	$\delta^{AC} + \omega_3^{AC}$
BC	Ref	δ^{AB}	$\delta^{AC} + \omega_4^{AC}$

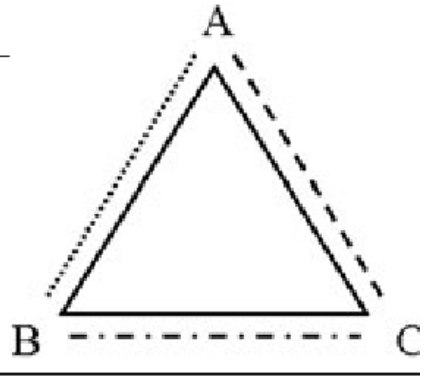
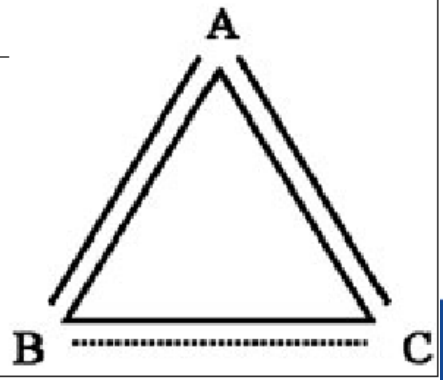
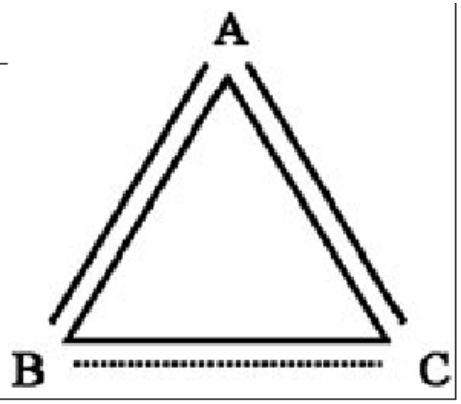


Table 3. Lu and Ades model for three treatments with order *A*, *B*, *C* (all possible designs).

Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	δ^{AB}	—
AC	Ref	—	δ^{AC}
BC	Ref	δ^{AB}	$\delta^{AC} + \omega^{AC}$



Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	δ^{AB}	—
AC	Ref	—	δ^{AC}
BC	Ref	δ^{AB}	$\delta^{AC}+\omega^{AC}$



- Model assumptions in choice of treatment as follows:

: All studies containing treatment A are assumed to estimate same treatment effects

: All studies containing treatment B but not treatment A are assumed to estimate same treatment effects

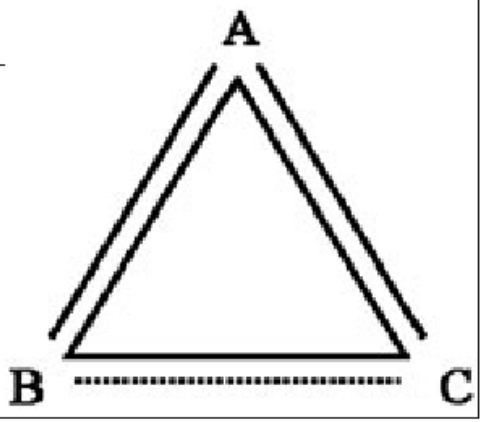
: All studies containing treatment C but not treatment A or treatment B are assumed to estimate same treatment effects

Although basing on loop inconsistency model, but

- Lu and Ades model depends on choice of baseline treatment to which all other treatments are compared (treatment A)

Table 3. Lu and Ades model for three treatments with order *A*, *B*, *C* (all possible designs).

Design	A	B	C
ABC	Ref	δ^{AB}	δ^{AC}
AB	Ref	δ^{AB}	—
AC	Ref	—	δ^{AC}
BC	Ref	δ^{AB}	$\delta^{AC}+\omega^{AC}$



Outline of talks

- Introduction
- Concepts of heterogeneity and inconsistency
- Models for inconsistency
- **Discussion**

Discussion and Conclusion

- Propose design-by-treatment interaction models to identify inconsistency in NMA
- Propose more complex methods for illustrating evidence networks in multi-arm studies
- Evidence inconsistency is impossible within multi-arm study

-> Loop inconsistency is difficult when there is mixture of two-arm and multi-arm studies

Discussion and Conclusion

Design-by-treatment approach

- Integrates idea of loop inconsistency with possibility of design inconsistency
- Only way to avoid arbitrary modelling constraints in NMA

- Design-by-treatment interaction models requires careful evaluation in practice

Discussion and Conclusion

Major concern about validity of NMA

: Possibility of loop inconsistency

: Attention to differences due to design
(choice of treatments included in studies)

: Presence or absence of other treatment arms in individual trials
: Complexity and loss in power in statistical tests involved
(differences across designs -> low likelihood of finding important conflicts)

Discussion and Conclusion

If inconsistency is identified in network

Q: What is the best strategy to proceed? -> **Remain unclear**

<u>Strategies for addressing inconsistency</u>

: Removing portions of evidence network

: Splitting nodes in network
(> 2 different treatments replace what was previously included as single treatment)

: Using study-level or individual-level covariates

: Seeking relevant inferences to presence of inconsistency

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Section for Clinical Epidemiology and Biostatistics

Thank you for your comments