FUNDED BY
NIHR | National Institute for Health Research

Not a giant leap, but a small step from pairwise to network meta-analysis

NIHR Complex Reviews Support Unit

March 2019

University of Glasgow
UNIVERSITY OF LEICESTER
LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Overview

13.30-14.30	Basic principles and underlying assumptions of pairwise and network meta-analysis Presentation of results in network meta-analysis
14.30-15.00	Practical exercise I
15.00-15.30	Coffee break
15.30-16.00	Implementation of network meta-analysis
16.00-16.30	Practical exercise II
16.30-17.00	Dealing with heterogeneity in network meta-analysis

Evidence synthesis should help us to make decisions

Fundamental Questions

- What is the most appropriate treatment or strategy for the patient?
- What further studies should be commissioned?
- Why? How? What? When?

Four principles for synthesizing evidence

Reward the creation of analyses for policymakers that are inclusive, rigorous, transparent and accessible, urge **Christl A. Donnelly** and colleagues.

FOUR PRINCIPLES

These features help researchers, policymakers and others to commission, do, share, appraise and use evidence syntheses.

INCLUSIVE

- Involves policymakers and is relevant and useful to them.
- Considers many types and sources of evidence.
- Uses a range of skills and people.

RIGOROUS

- Uses the most comprehensive feasible body of evidence.
- Recognizes and minimizes bias.
- Is independently reviewed as part of a quality-assurance process.

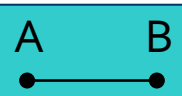
TRANSPARENT

- Clearly describes the research question, methods, sources of evidence and quality-assurance process.
- Communicates complexities and areas of contention.
- Acknowledges assumptions, limitations and uncertainties, including any evidence gaps.
- Declares personal, political and organizational interests and manages any conflicts.

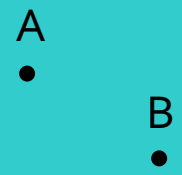
ACCESSIBLE

- Is written in plain language.
- Is available in a suitable time frame.
- Is freely available online.

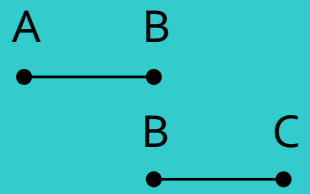
A Taxonomy of Comparisons



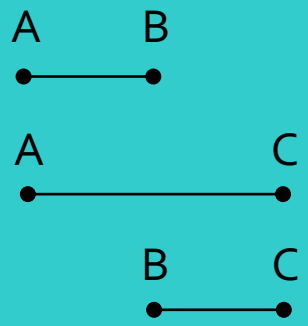
Direct Comparison:
head to head evidence



'Naïve' or 'Unadjusted' Indirect Comparison:
Absolute effect estimates from individual trial arms



'Adjusted' Indirect Comparison:
Relative effect estimates between treatments



Mixed Treatment Comparison/'Network' Meta-Analysis:
'Adjusted' indirect comparison extended to more complex networks of trial evidence
(head to head and indirect evidence)

Direct Comparison: Meta-analysis of RCTs of the effect of aspirin in preventing death after myocardial infarction

	Aspirin		Placebo	
Study	Deaths	Total	Deaths	Total
MRC-1	49	615	67	624
CDP	44	758	64	771
MRC-2	102	832	126	850
GASP	32	317	38	309
PARIS	85	810	52	406
AMIS	246	2267	219	2257
ISIS-2	1570	8587	1720	8600

Fleiss. The statistical basis of meta-analysis. Stat Methods Med.Res.1993, 2: 121-145

Data from MRC-1

	Death	**Alive**
Aspirin	49	566
Placebo	67	557

The Odds Ratio – MRC-1

	Death	Alive
Aspirin	49	566
Placebo	67	557

- **OR** = (49x557)/(566x67) = 0.72
- **Ln(OR)** = ln(0.72) = -0.33
- **Var Ln(OR)** = 1/49 + 1/566 + 1/67 + 1/557 = 0.04
- **SE Ln(OR)** = SQRT(0.04) = 0.20
- **95% CI for Ln(OR)** = -0.33 ± 1.96*0.20 = 0.06, -0.72
- **95% CI for OR** = exp (0.06), exp (-0.72) = 0.49, 1.06

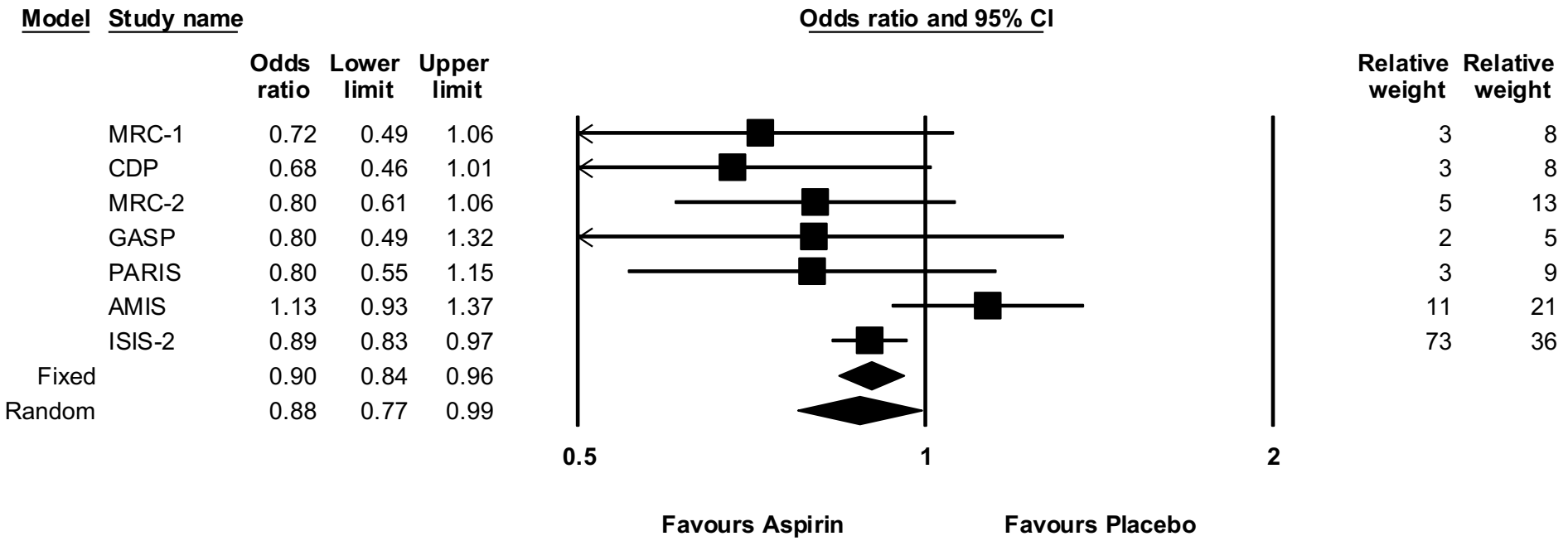
Principles of Meta-analysis

- Summary statistic for individual studies
    - Risk ratios (binary variables)
    - Difference between means (continuous variables)
- Pooled treatment effect estimates
    - Weighted average of treatment effects

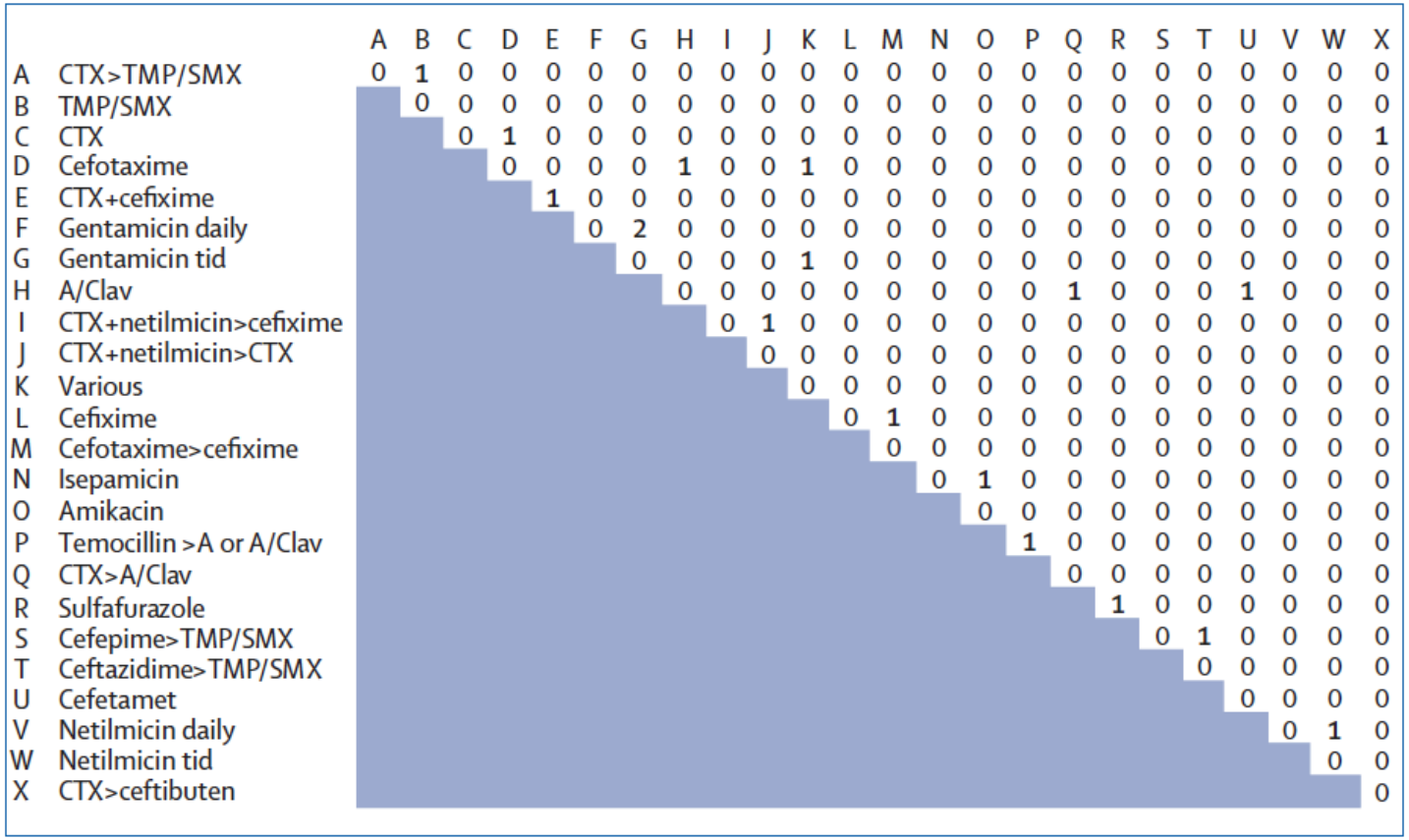
$$\frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

$$\frac{\text{sum of (weight x estimate)}}{\text{sum of weights}}$$

Comparing fixed and random effects results



The fundamental problem with pairwise meta-analysis



		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
A	CTX>TMP/SMX	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	TMP/SMX		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	CTX			0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
D	Cefotaxime				0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
E	CTX+cefixime					1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	Gentamicin daily						0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	Gentamicin tid							0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
H	A/Clav								0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
I	CTX+netilmicin>cefixime									0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	CTX+netilmicin>CTX										0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	Various											0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	Cefixime												0	1	0	0	0	0	0	0	0	0	0	0	0
M	Cefotaxime>cefixime													0	0	0	0	0	0	0	0	0	0	0	0
N	Isepamicin														0	1	0	0	0	0	0	0	0	0	0
O	Amikacin															0	0	0	0	0	0	0	0	0	0
P	Temocillin >A or A/Clav																1	0	0	0	0	0	0	0	0
Q	CTX>A/Clav																	0	0	0	0	0	0	0	0
R	Sulfafurazole																		1	0	0	0	0	0	0
S	Cefepime>TMP/SMX																			0	1	0	0	0	0
T	Ceftazidime>TMP/SMX																				0	0	0	0	0
U	Cefetamet																					0	0	0	0
V	Netilmicin daily																						0	1	0
W	Netilmicin tid																							0	0
X	CTX>ceftibuten																								0

***Figure*: Comparisons of antibiotic regimens for acute pyelonephritis in children**
Matrix shows number of available direct comparisons (0=no comparison, 1=one comparison, 2=two comparisons). Drugs with different doses are grouped together except for daily versus multiple daily doses of aminoglycosides. Sequential regimens are shown by ">". A=amoxicillin, A/Clav=co-amoxiclav, CTX=ceftriaxone, tid=three times daily, TMP/SMX=co-trimoxazole. From randomised trials in reference 4.

Ioannidis et al. Indirect comparisons: the mesh and mess of clinical trials. Lancet 2006;368:1470–5

One solution is to “lump” treatments together to allow pairwise analysis

The trials assessed **11 growth factors in 30 comparisons**: platelet-derived wound healing formula, autologous growth factor, allogeneic platelet-derived growth factor, transforming growth factor β2, arginine-glycine-aspartic acid peptide matrix, recombinant human platelet-derived growth factor (becaplermin), recombinant human epidermal growth factor, recombinant human basic fibroblast growth factor, recombinant human vascular endothelial growth factor, recombinant human lactoferrin, and recombinant human acidic fibroblast growth factor.

Martí-Carvajal et al. Growth factors for treating diabetic foot ulcers. Cochrane database Syst Rev. 2015;10:CD008548.

The results for any growth factor versus placebo or no growth factor

Comparison 1. Any growth factor versus placebo or no growth factor

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Complete wound closure	12	1139	Risk Ratio (M-H, Fixed, 95% CI)	1.51 [1.31, 1.73]
2 Lower limb amputation (minimum of one toe)	2	219	Risk Ratio (M-H, Fixed, 95% CI)	0.74 [0.39, 1.39]
3 Ulcer-free days following treatment for diabetic foot ulcers (free from any recurrence)	1		Hazard Ratio (Fixed, 95% CI)	0.64 [0.14, 2.94]
4 Adverse events (non-serious and serious)	4	385	Risk Ratio (M-H, Fixed, 95% CI)	0.98 [0.79, 1.22]

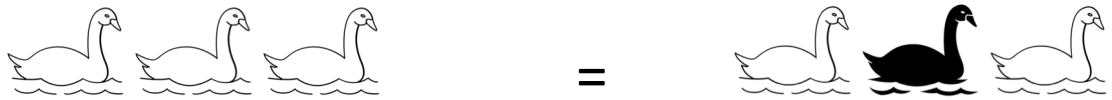
A summary of results

- **Any growth** factor compared with **placebo or no growth factor** increased the number of participants with complete wound healing (345/657 (52.51%) versus 167/482 (34.64%); RR 1.51, 95% CI 1.31 to 1.73; I^2 = 51%, 12 trials)

- The result is mainly based on **platelet-derived wound healing formula** (36/56 (64.28%) versus 7/27 (25.92%); RR 2.45, 95% 1.27 to 4.74; I2 = 0%, two trials), and **recombinant human platelet-derived growth factor** (becaplermin) (205/428 (47.89%) versus 109/335 (32.53%); RR 1.47, 95% CI 1.23 to 1.76, I^2= 74%, five trials)

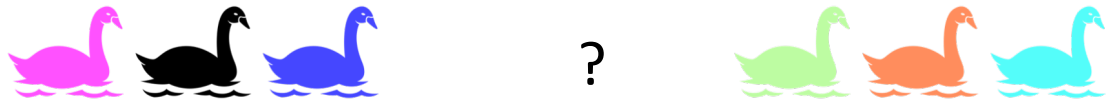
Meta-analysis can be difficult to interpret if treatments are “lumped together”

- A (statistically significant) difference is observed between the groups being compared



At least one treatment in one group is different from at least one treatment in the other group

- A (statistically significant) difference is not observed between the groups being compared



Independent Pairwise Comparisons

Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation

A Boland
Y Dundar
A Bagust
A Haycox
R Hill
R Mujica Mota
T Walley
R Dickson*

Liverpool Reviews and Implementation Group, New Medical School, Liverpool, UK

Executive summary

Health Technology Assessment 2003; Vol. 7: No. 15

Health Technology Assessment
NHS R&D HTA Programme

Multiple treatments and trial comparisons

TABLE 2 *Summary of included clinical studies*

Alteplase/streptokinase	Alteplase/ tenecteplase	Alteplase/ reteplase	Streptokinase/ reteplase	Dose-ranging and mixed regimes
GUSTO I[18] Central Illinois[43] Cherng *et al.*[44] ECSG[45] GISSI-2/ISG[46,47] ISIS-3[48] KAMIT[49] PAIMS[50] TIMI 1[51] White *et al.*[59]	ASSENT-2[20] *	GUSTO III[19] * RAPID-2[17] *	INJECT[52]	COBALT[57] (t-PA)* Xu *et al.*[58] (SK) Six *et al.*[53] (SK) ASSENT-1[54] (TNK) TIMI 10B[55] (TNK)* RAPID-1[56] (r-PA)

* *Involved accelerated alteplase*

Analyses of Mortality up to 35 Days

Outcome	Study	Alteplase	Streptokinase	OR random effect (95% CI)
Mortality up to 35 days	Central Illinois[43]	6/123	9/130	0.69 (0.24 to 2.00)
	Cherng *et al.*[44]	2/59	5/63	0.41 (0.08 to 2.18)
	ECSG[45]	3/64	3/65	1.02 (0.20 to 5.23)
	GISSI-2/ISG[46,47]	929/10,372	887/10,396	1.05 (0.96 to 1.16)
	GUSTO I[18]	652/10,344	1,472/20,173	0.85 (0.78 to 0.94)
	ISIS-3[48]	1,418/13,746	1,455/13,780	0.97 (0.90 to 1.05)
	PAIMS[50]	4/86	7/85	0.54 (0.15 to 1.93)
	TIMI 1[51]	7/143	12/147	0.58 (0.22 to 1.52)
	White *et al.*[59]	5/135	10/135	0.48 (0.16 to 1.45)
	Total	**3,026/35,072**	**3,860/44,974**	**0.94 (0.85 to 1.04)**
				Test for heterogeneity $\chi^2 = 13.96$, df = 8, $p = 0.083$

Outcome	Study	Accelerated alteplase	Reteplase	OR random effect (95% CI)
Mortality up to 35 days	GUSTO III[19]	356/4,921	757/10,138	0.97 (0.85 to 1.10)
	RAPID-2[17]	13/155	7/169	2.12 (0.82 to 5.46)
	Total	**369/5,076**	**764/10,307**	**1.24 (0.61 to 2.53)**
				Test for heterogeneity $\chi^2 = 2.60$, df = 1, $p = 0.11$

Outcome	Study	Accelerated alteplase	Tenecteplase	OR random effect (95% CI)
Mortality up to 35 days	ASSENT-2[20]	522/8,488	523/8,461	0.99 (0.88 to 1.13)

Outcome	Study	Reteplase	Streptokinase	OR random effect (95% CI)
Mortality up to 35 days	INJECT[52]	270/2,994	285/2,992	0.94 (0.79 to 1.12)

This is difficult to summarise…

"Definitive conclusions on efficacy are that streptokinase is as effective as non-accelerated alteplase, that tenecteplase is as effective as accelerated alteplase, and that reteplase is at least as effective as streptokinase.

Some conclusions require interpretation of data, i.e. whether streptokinase is as effective as, or inferior to accelerated alteplase; and whether reteplase is as effective as accelerated alteplase or not.

Depending on these, two further conclusions on indirect comparisons arise, whether tenecteplase is superior to streptokinase or not, and whether reteplase is as effective as tenecteplase or not."

*From Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica Mota R, et al. Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. Health Technol Assess 2003;**7**(15).*

Acknowledgements to Julian Higgins

Simultaneous comparison of multiple treatments: combining direct and indirect evidence

Deborah M Caldwell, A E Ades, J P T Higgins

How can policy makers decide which of five treatments is the best? Standard meta-analysis provides little help but evidence based decisions are possible

Several possible treatments are often available to treat patients with the same condition. Decisions about optimal care, and the clinical practice guidelines that inform these decisions, rely on evidence based evaluation of the different treatment options.[1 2] Systematic reviews and meta-analyses of randomised controlled trials are the main sources of evidence. However, most systematic reviews focus on pair-wise, direct comparisons of treatments (often with the comparator being a placebo or control group), which can make it difficult to determine the best treatment. In the absence of a collection of large, high quality, randomised trials comparing all eligible treatments (which is invariably the situation), we have to rely on indirect comparisons of multiple treatments. For example, an indirect estimate of the benefit of A over B can be obtained by comparing trials of A *v* C with trials of B *v* C,[3–5] even though indirect comparisons produce relatively imprecise estimates.[6] We describe comparisons of three or more treatments, based on pair-wise or multi-arm comparative studies, as a multiple treatment comparison evidence structure.

- AND DENI McINTYRE/SPL

Medical Research Council Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol BS8 2PR

Deborah M Caldwell
research associate

A E Ades
professor of public health science

Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge CB2 2SR

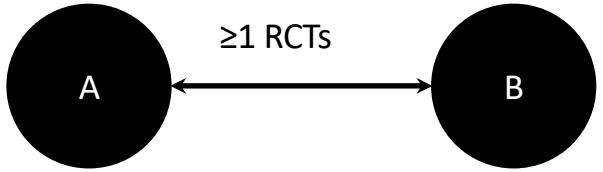
J P T Higgins
senior research associate

Correspondence to:
D M Caldwell
d.m.caldwell@bristol.ac.uk

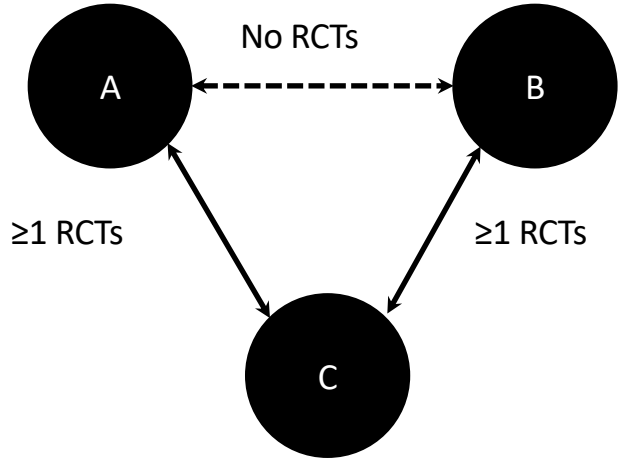
BMJ 2005;331:897–900

Network meta-analysis:
an extension from pairwise meta-analysis

Direct comparison



Indirect comparison

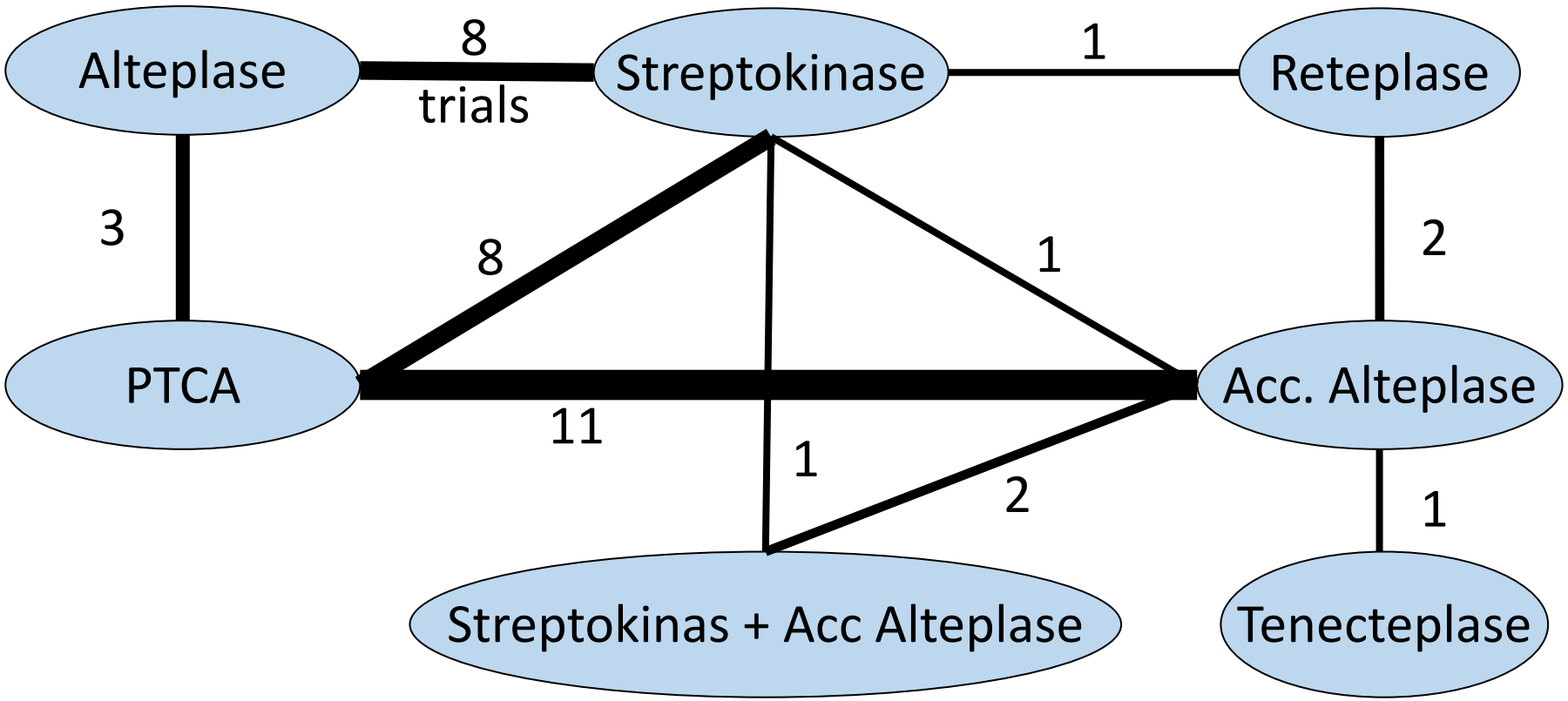


The network meta-analysis was based in the following pairwise comparisons

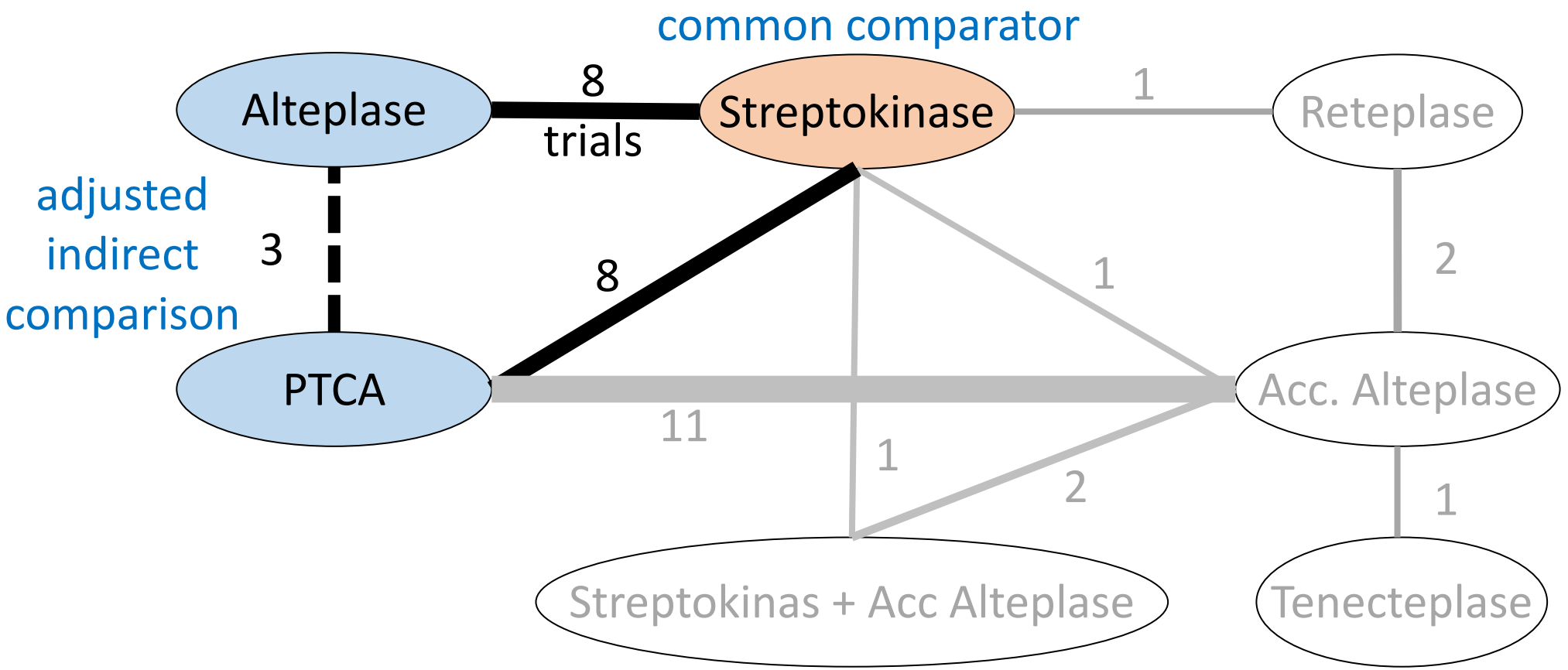
35 Day Mortality Odds Ratios (95% CI)

Treatment:	Compared to: Streptokinase	Alteplase	Acc. Alteplase
Streptokinase	1		
Alteplase	0.89 (0.54 to 1.14)	1	
Acc. Alteplase	0.86 (0.78 to 0.94)		1
Streptokinase+Alteplase	0.96 (0.87 to 1.05)		1.12 (1.00 to 1.25)
Reteplase	0.95 (0.79 to 1.12)		1.02 (0.90 to 1.16)
Tenecteplase	-		1.01 (0.88 to 1.14)
PTCA	0.49 (0.20 to 0.91)	0.63 (0.25 to 1.29)	0.79 (0.55 to 1.05)

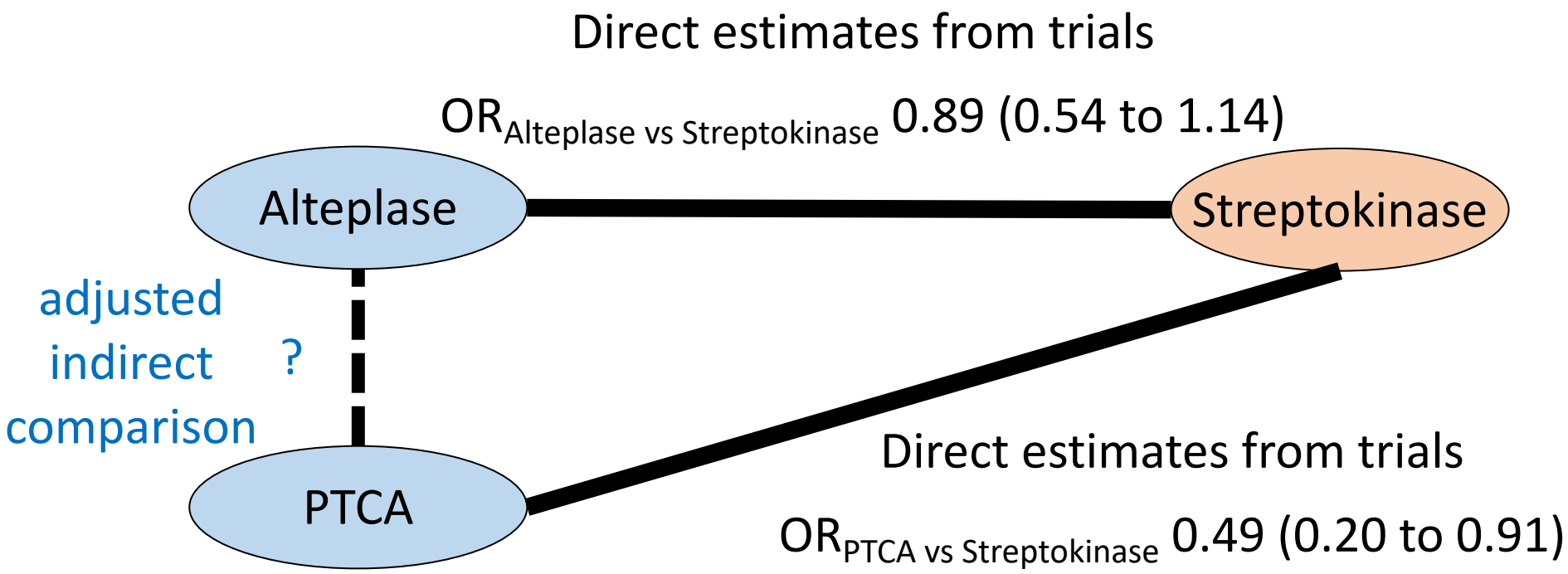
The network of trial evidence is analysed as a ‘whole’



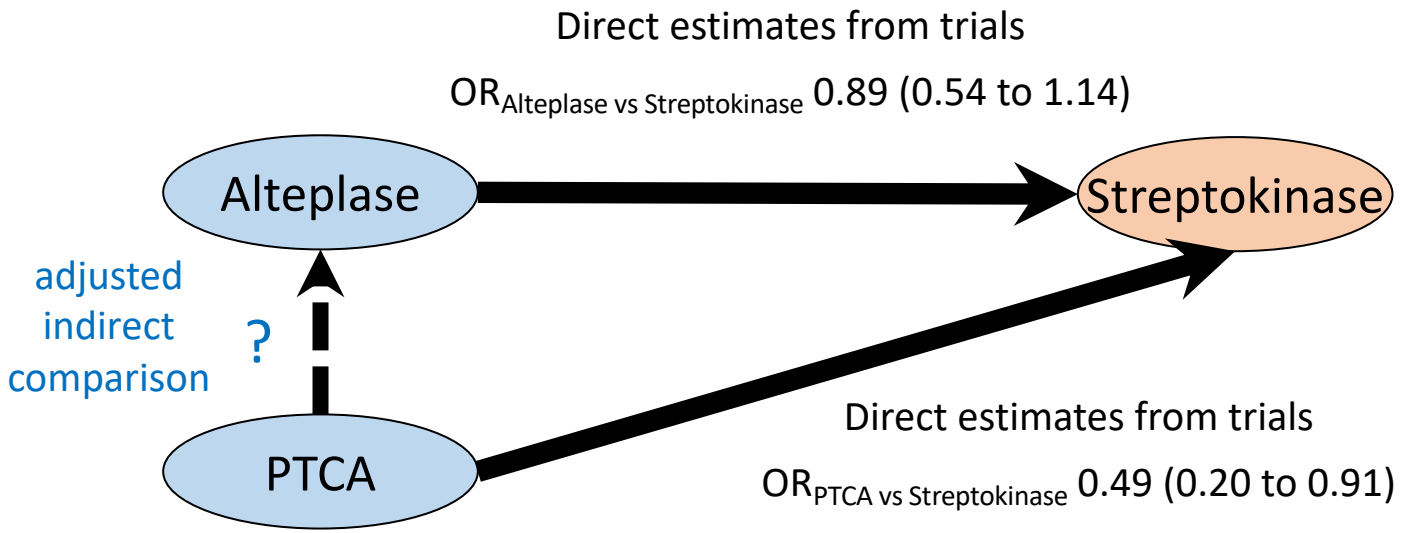
The basic building block: adjusted indirect comparison



Indirect Comparison: PTCA vs Alteplase



Indirect Comparison: PTCA vs Alteplase



Adjusted indirect estimates

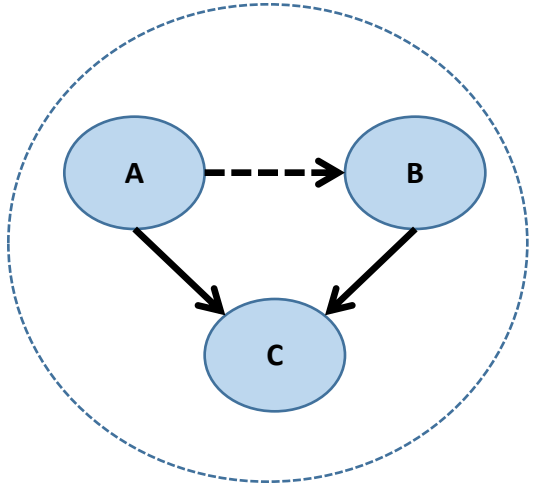
$$OR_{PTCA\ vs\ Alteplace} = OR_{PTCA\ vs\ Streptokinase} / OR_{Alteplase\ vs\ Streptokinase}$$

$$= 0.49 / 0.89$$

$$= 0.55$$

The assumption of consistency

$$OR_{AB} = {OR_{AC}}/{OR_{BC}}$$



$$\log(OR_{AB}) = \log(OR_{AC}) - \log(OR_{BC})$$

$$\partial_{AB} = \partial_{AC} - \partial_{BC}$$

The assumption of consistency

Uncertainty in Indirect Comparisons

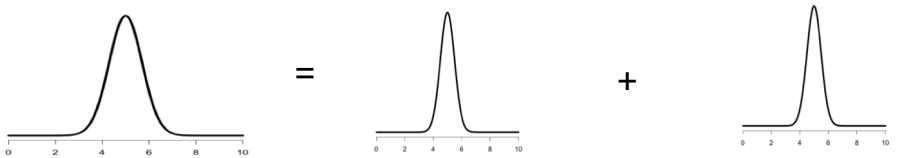
$$OR_{AB} = {}^{OR_{AC}}\!\!\diagup\!\!_{OR_{BC}}$$

$OR_{A\ vs\ C}$ and $OR_{B\ vs\ C}$ are available from pairwise meta-analyses. These are converted to log odds ratios so the indirect estimate becomes the difference between the direct estimates.

$$\ln OR_{AvsB} = \ln OR_{AvsC} - \ln OR_{BvsC}$$

Bucher et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of Clinical Epidemiology 1997;50(6):683-9

Uncertainty in Indirect Comparisons



$$Var\left(OR_{AvsB}\right) = Var\left(OR_{AvsC}\right) + Var\left(OR_{BvsC}\right)$$

This allows the Standard Error (SE) to be calculated on the Log scale:

$$SE\left(\ln OR_{AvsB}\right) = \sqrt{SE\left(\ln OR_{AvsC}\right)^2 + SE\left(\ln OR_{BvsC}\right)^2}$$

Uncertainty in Indirect Comparisons

The mean and 95% CI are calculated on the log scale, and these are then exponentiated to obtain estimates on the odds ratio scale

$$95\%\ CI\ (\ lnOR_{A\ vs\ B}) = lnOR_{A\ vs\ B} \pm 1.96 \times SE(lnOR_{A\ vs\ B})$$

$$Odds\ Ratio = exp(lnOR_{A\ vs\ B})$$

$$95\%\ CI\ (\ OR_{A\ vs\ B}) = Exp(lnOR_{A\ vs\ B} \pm 1.96 \times SE(lnOR_{A\ vs\ B}))$$

An example calculation

1) Calculate mean difference in log relative risk

$$\ln\left(OR_{PTCA.Alt}\right) = \ln\left(OR_{PTCA.Strep}\right) - \ln\left(OR_{Alt.Strep}\right)$$
$$-0.60 = -0.71 - -0.12$$

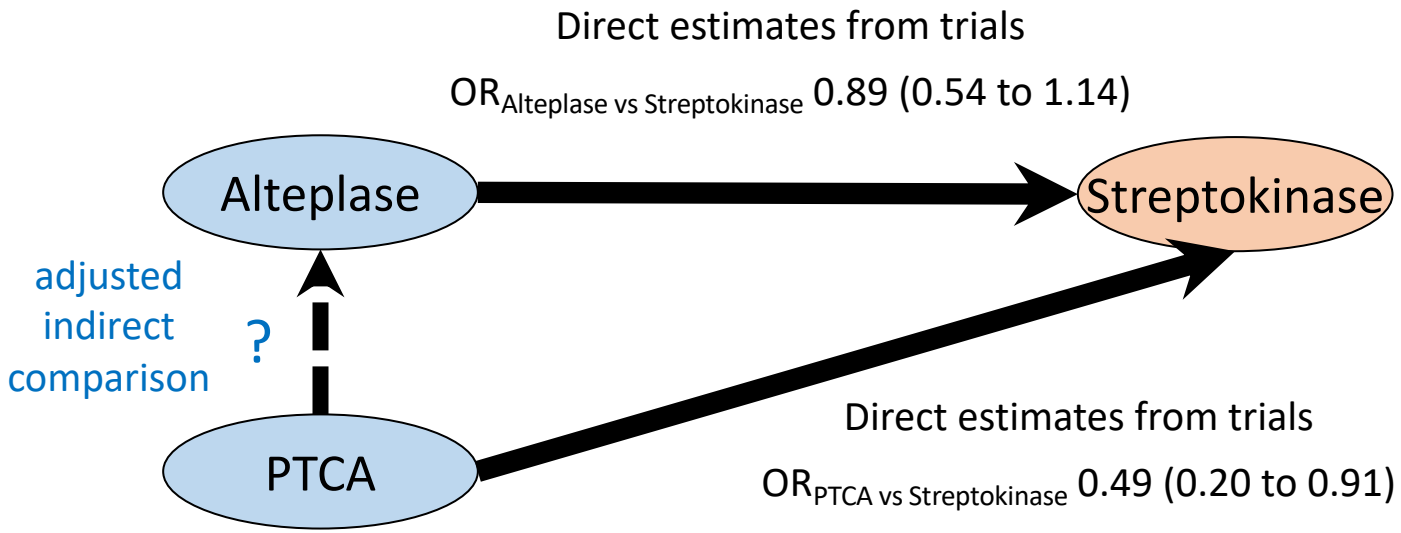
2) Calculate Ses for mean difference in log relative risks

$$SE\left(OR_{alt.ptca}\right)^2 = \sqrt{SE\left(OR_{alt.strep}\right)^2 + SE\left(OR_{ptca.strep}\right)^2}$$
$$0.43 = \sqrt{0.39^2 + 0.19^2}$$

3) Exponentiate to get Relative Risks

$\ln(OR_{PTCA.Alt})$ = -0.60 (SE 0.43) $OR_{PTCA.Alt}$ = 0.55 (95% 0.24 to 1.28)

Indirect Comparison: PTCA vs Alteplase



Adjusted indirect estimates

$$OR_{PTCA\ vs\ Alteplace} = OR_{PTCA\ vs\ Streptokinase} / OR_{Alteplase\ vs\ Streptokinase}$$

$$= 0.49 / 0.89$$

$$= 0.55\ (0.24\ to\ 1.29)$$

Uncertainty in Indirect Estimates

- Only represents uncertainty arising from the sampling error in the contributing trials
- Does not represent uncertainty in the fundamental assumptions
- Does not represent uncertainty in heterogeneity in predictive factors
- Absolute ‘Best Case’ estimate of uncertainty

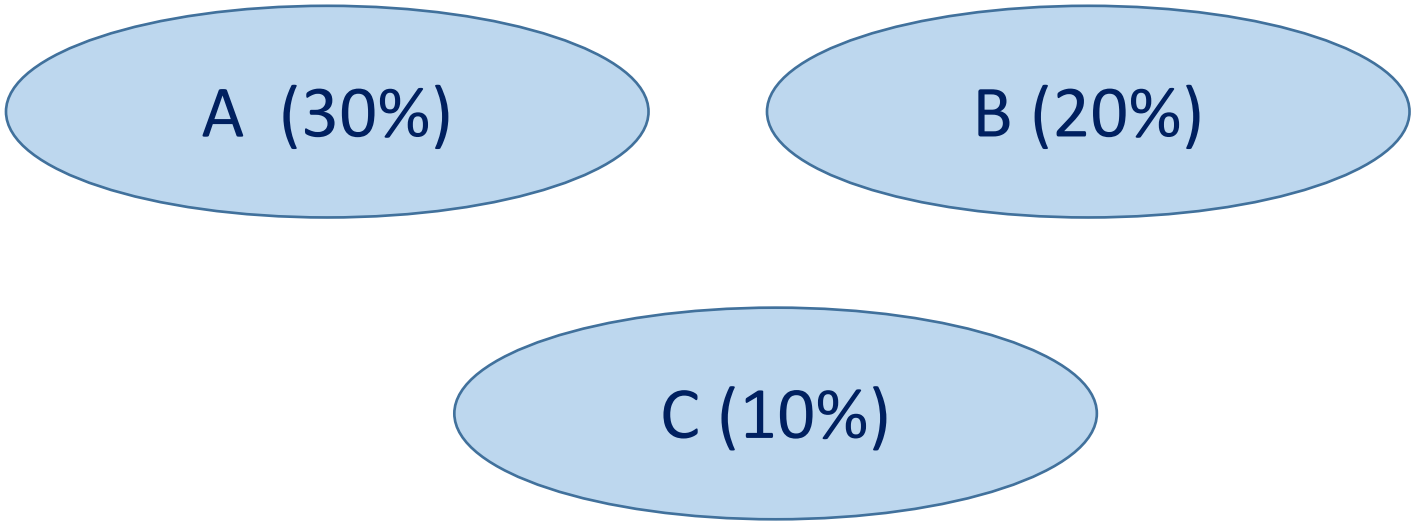
Network Meta-Analysis

- Extension of the basic indirect comparison to more complex networks
- Estimates treatment effects that best ‘fit’ the network of trial comparisons

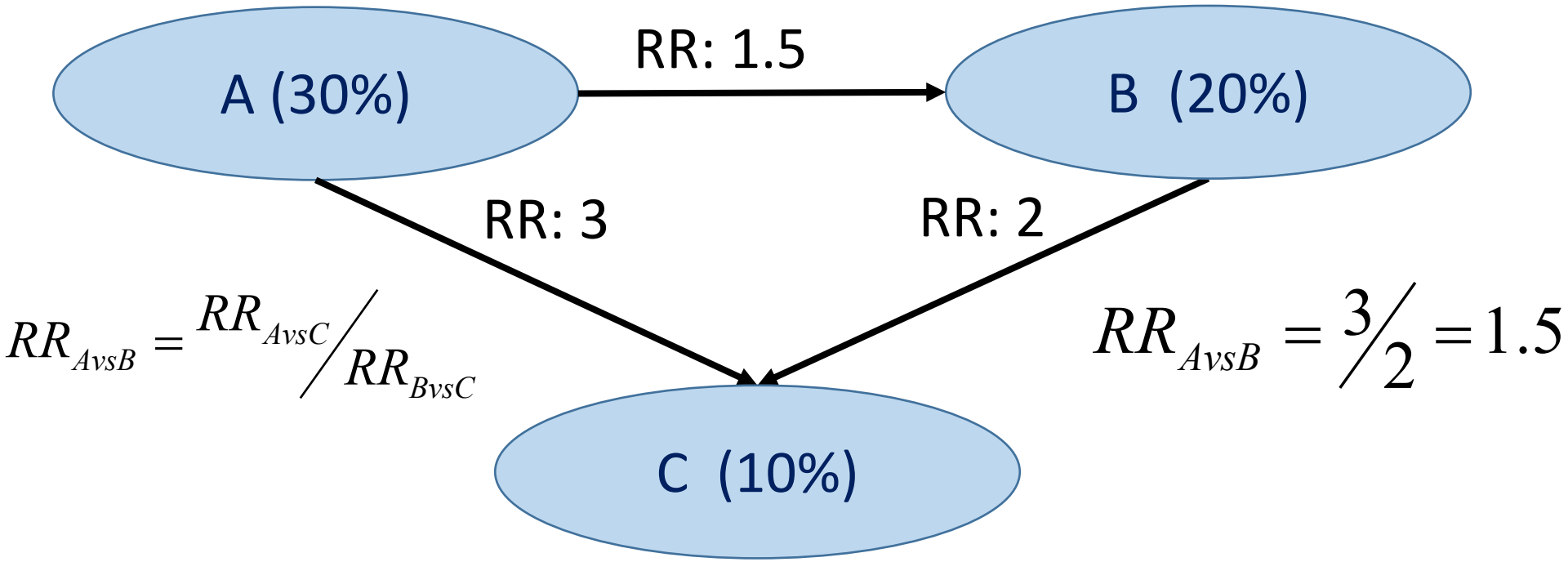
1. $\beta_{Alteplase}$, $\beta_{Reteplease,}$ β_{PTCA} are estimates of the Log Odds Ratio (LOR) of Alteplase, Reteplase and PTCA compared to a reference comparator (e.g. Streptokinase).
2. $LOR_{Alteplase\ vs\ Streptokinase} = \beta_{Alteplase}$
3. $LOR_{Reteplase\ vs\ Streptokinase} = \beta_{Reteplase}$
4. $LOR_{PTCA\ vs\ Streptokinase} = \beta_{PTCA}$
5. $LOR_{Alteplase\ vs\ PTCA} = \beta_{Alteplase} - \beta_{PTCA}$ (consistency assumption)

Where does the assumption of consistency come from?

- Consider a single three arm trial



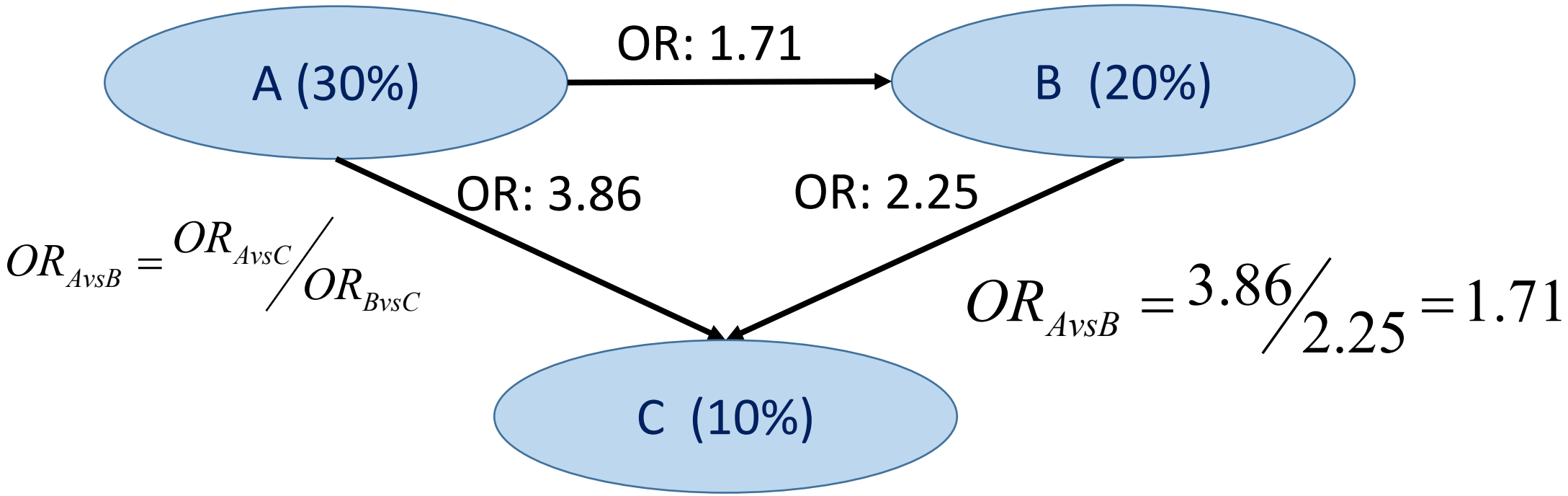
By definition is consistent on the relative risk scale



$$RR_{AvsB} = RR_{AvsC} / RR_{BvsC}$$

$$RR_{AvsB} = 3/2 = 1.5$$

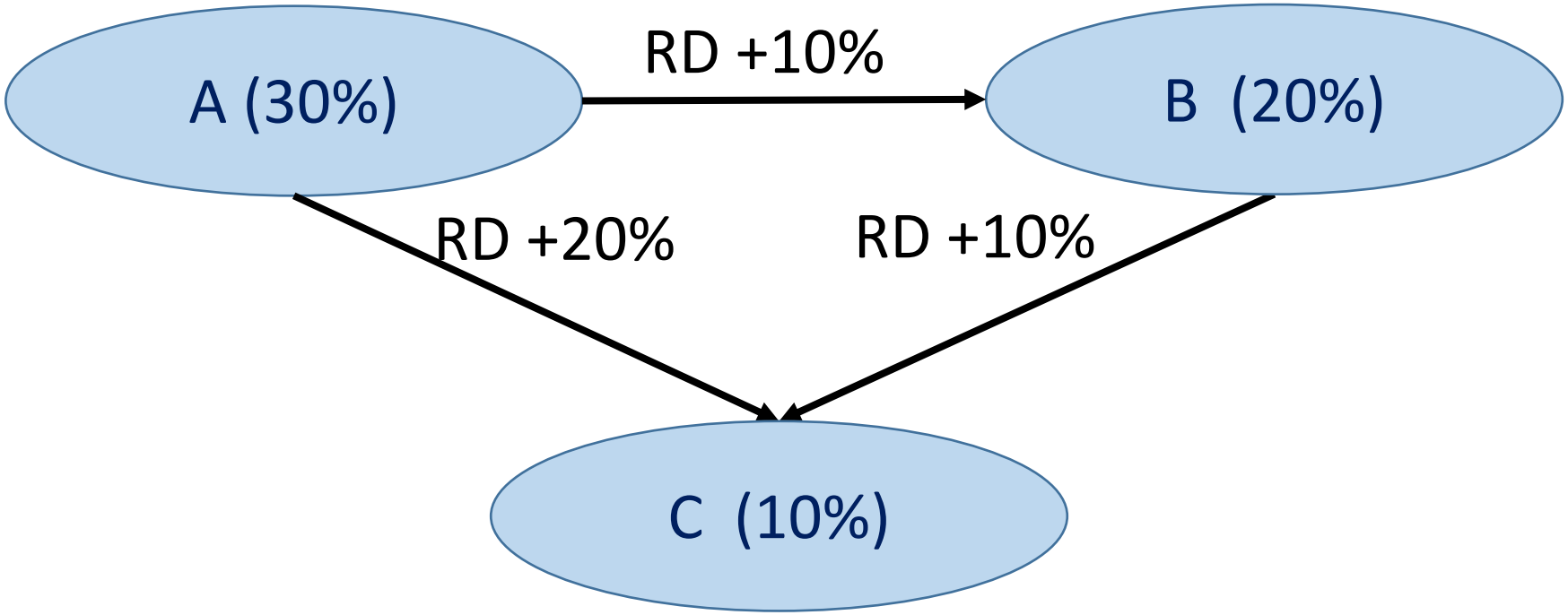
And on odds ratio scale



$$OR_{AvsB} = {OR_{AvsC}}/{OR_{BvsC}}$$

$$OR_{AvsB} = {3.86}/{2.25} = 1.71$$

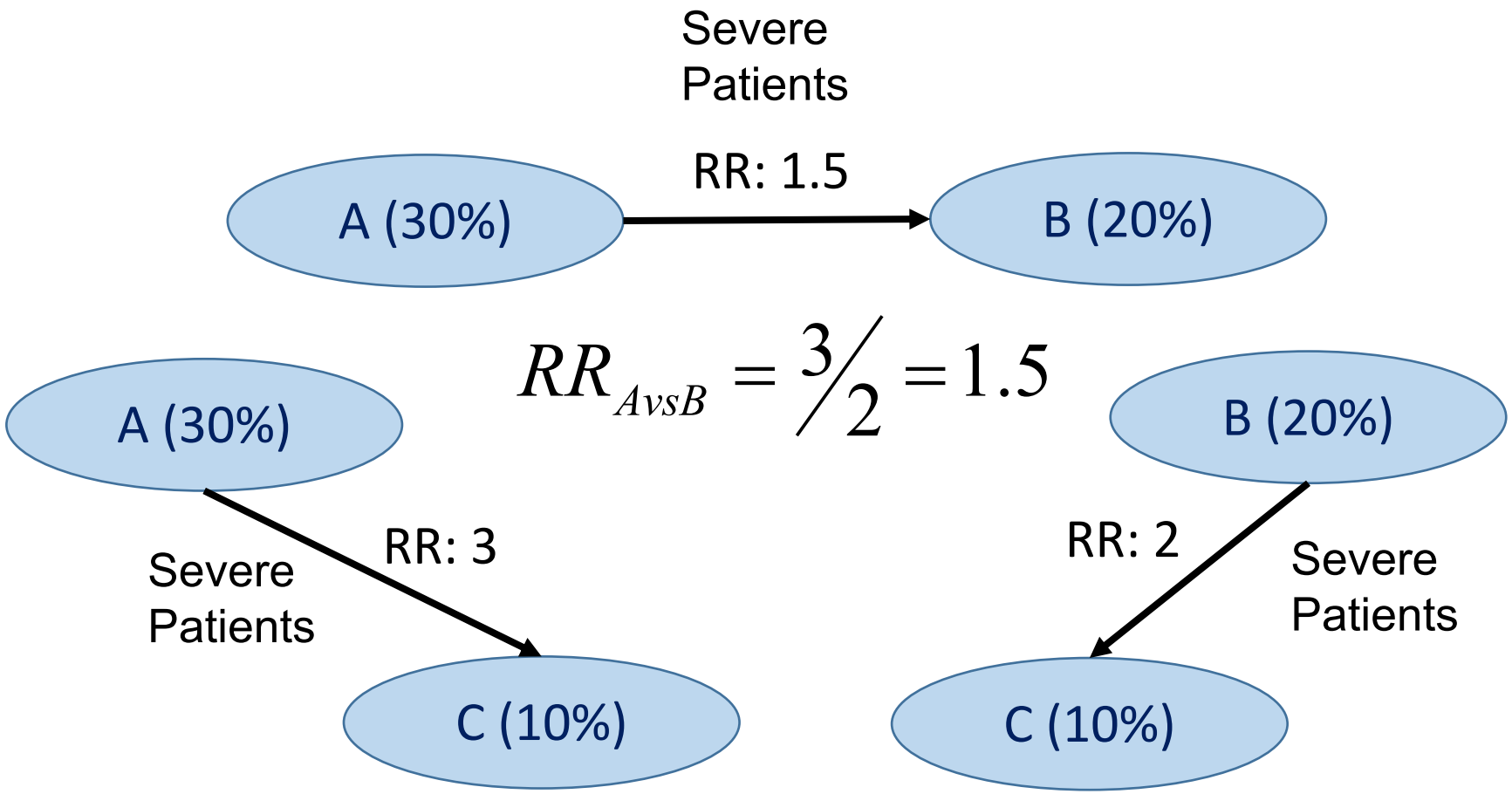
And on risk difference scale



$$RD_{AvsB} = RD_{AvsC} - RD_{BvsC}$$

$$RD_{AvsB} = 20\% - 10\% = +10\%$$

A completely homogeneous set of trials…
will behave like a single multi-arm trial and be consistent



Direct vs Indirect Evidence

Analysis	Comparison	Odds Ratio Mean (95% CI)
Direct (3 Trials)	PTCA vs. Alteplase	0.63 (0.25 to 1.29)
Indirect via Streptokinase (16 Trials)	PTCA vs. Alteplase	0.55 (0.24 to 1.28)
Network Meta-Analysis	PTCA vs. Alteplase	0.65 (0.49 to 0.86)

The basic assumption

- **Similarity**
  - Trials are clinically and methodologically similar and comparable
- **Exchangeability**
  - If patients in one trial were substituted in another, the observed treatment estimates would be the same (allowing for random variation)
- **Transitivity** $\partial_{AB} = \partial_{AC} - \partial_{BC}$ $\partial_{AC} = \partial_{AB} - \partial_{CB}$
- **Consistency**
  - Indirect and direct estimates are consistent

Reasons for preferring direct evidence

- Believed to be less biased than indirect evidence
- Use of indirect evidence may inhibit the conduct of further trials

Motivations for use of indirect comparisons

- No direct trial data exists
- Estimates from a combination of direct and indirect evidence may be more precise
- Indirect evidence may be believed to be less biased
- Estimates from a combination of direct and indirect evidence may more accurately effect uncertainty
- Incoherence between direct and indirect evidence may be informative
  - May facilitate adjustment for heterogeneity between trials
  - More accurate prediction of treatment effects
  - May be useful in selection of appropriate scale for analysis

Some thoughts

- Single source of evidence are rarely sufficient
- Pairwise meta-analysis has limitations
- Network meta-analysis also has limitations
- Trade-offs between methodological limitations and decision-making