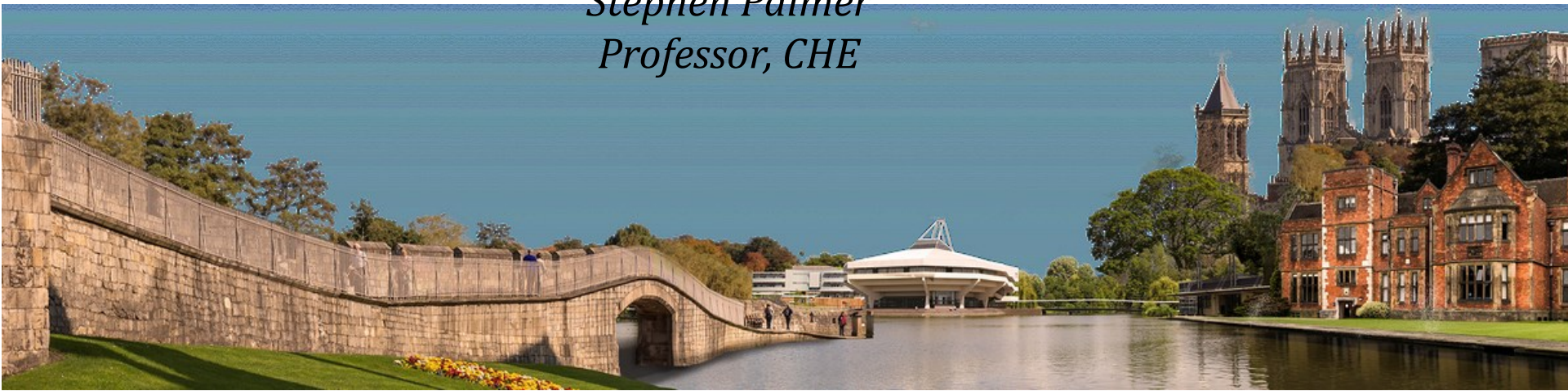
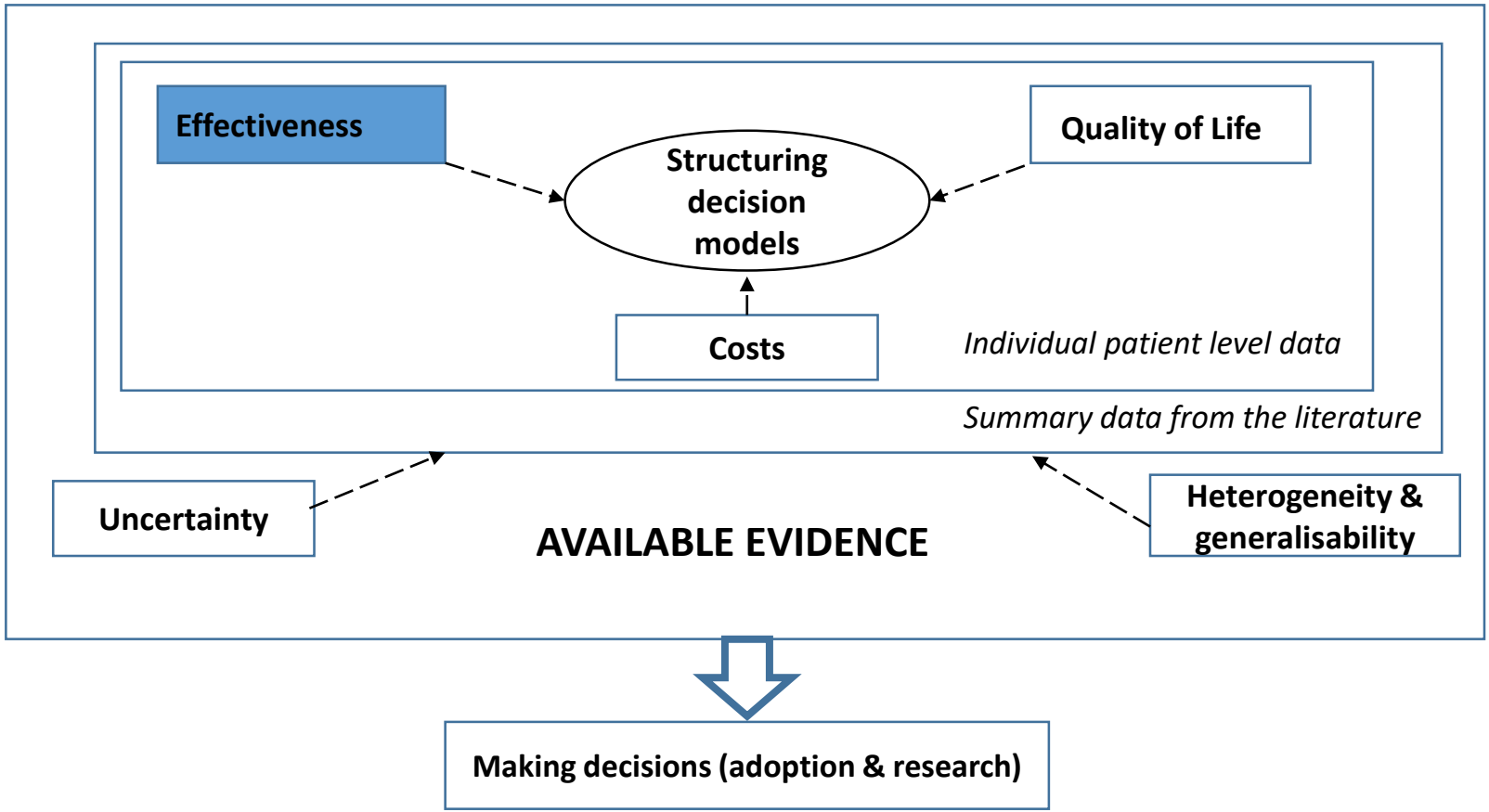
Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating models - effectiveness
3.1: Overview and objectives

Stephen Palmer
Professor, CHE



Course structure – where are we up to?



Overview

- Methods and application of meta-analysis represent a key stage in the population of decision models
- Identifying relevant evidence:
  - Relative treatment effect
  - Other relevant parameters
- Relating evidence to decision problem
  - Use of direct and indirect sources
- Application within decision model to estimate policy relevant outcomes

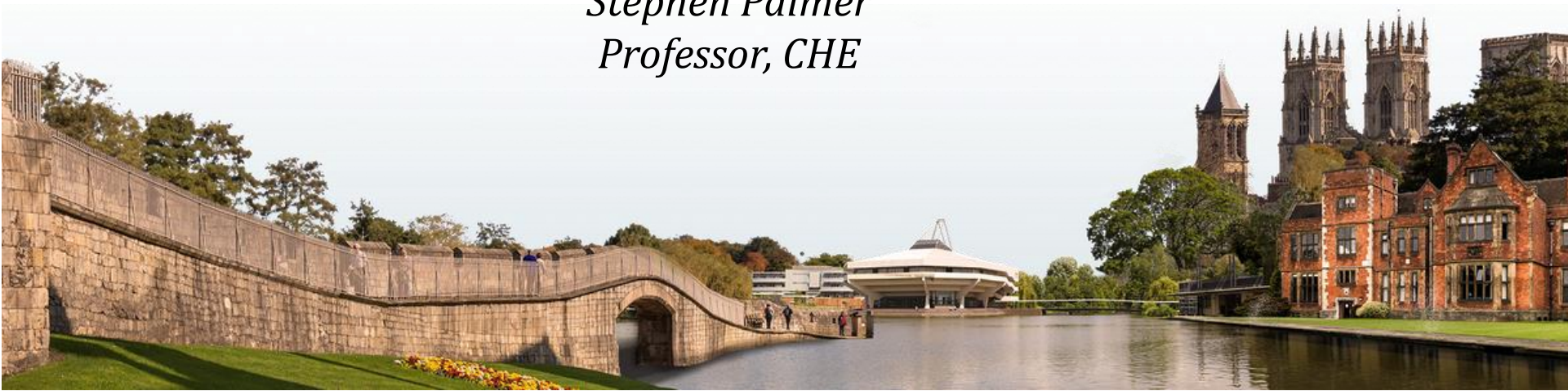
Objectives

- Understand the key concepts relating to the synthesis of effectiveness evidence for use in decision models
- Appreciate the importance of meta-analysis to estimate relative effects
- Understand the key assumptions of meta-analysis and network meta analysis
- Understand how the results of meta-analysis can be incorporated within decision models
- Recognise the importance of heterogeneity for policy decisions

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating models - effectiveness
3.2: Key concepts

Stephen Palmer
Professor, CHE



Objectives

- Appreciate the complementary roles of evidence synthesis and decision modelling.
- Understand the general principles of evidence synthesis including the role of observational data.

Why evidence synthesis and decision modelling?

- A single source is unlikely to provide evidence on final endpoints required for decision making (e.g. lifetime QALYs and costs) for all comparators
- Evidence synthesis and modelling allows:
  - Inclusion of all relevant comparators
  - Combining relevant sources of evidence
  - Linkage of short-term outcomes to longer term impact (extrapolation)
  - Generalisability to population/setting of interest
  - Assessment of how uncertainty in relationships and evidence translates to uncertainty in decision
  - Scientific and value judgments to be made explicit

Evidence synthesis to inform decision models

- Pooling results from multiple studies to inform parameter(s) inputs to decision model
  - Typically used for relative treatment effects
  - Can also be used for baseline outcomes (for reference treatment)
- Why pool?
  - To reduce problems of interpretation due to sampling variation
  - To facilitate synthesis of results from multiple studies
  - To quantify effect sizes and their uncertainty
- Evidence should be identified using systematic literature review
  - e.g. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf

General principles of evidence synthesis

1. Inclusion/exclusion criteria should reflect target population for decision
   - Population (subgroups), Intervention, Comparators, Outcomes
2. Relevance of data sources to decision problem must be discussed/explored
   - Differences in trial populations (severity, previous treatments etc.)
   - Differences in treatment regimens/doses
   - Differences in controls (Best supportive care, placebo, older ineffective treatments)
   - Differences in outcomes (definitions, measures, treatment switching)
   - Any adjustments made

General principles of evidence synthesis (cont)

3. Statistical heterogeneity should be taken into account
   - Assessing heterogeneity – Q statistics, I^2, between studies variance, compare fit of fixed and random effects models
   - Assess inconsistency in network meta-analysis
   - Subgroup-analysis, meta-regression
4. The origin and rationale for the parameter values used in the model should be described and justified
   - Clear description of data sources, rationale for selection
   - Description of synthesis methods (model and code, model fit)
   - Justification of assumptions made (and sensitivity analysis)
   - Results of synthesis
   - Source of baseline used to get absolute effects for decision model

Role of observational data

1. OK for defining the baseline
2. OK for mapping from shorter-term RCT outcomes to longer term clinical outcomes or QALYs (e.g. informing extrapolation of survival curves)
3. Not usually OK for estimating relative treatment effects

- Unless
  - No RCT, or if RCT data sparse
  - Potential bias should be taken into account and integrated into analysis by valid methods, including sensitivity analysis

Summary

- Evidence synthesis can provide important information on relative effectiveness (and potentially baseline outcomes)
- Potentially for all relevant compactors
- Bringing together multiple relevant data sources
- But
  - Is not the sole source of evidence considered important to a decision
  - Will not give estimates of costs or QALYs over an appropriate time horizon
- Decision modelling brings together all forms of evidence
  - Provides outputs required for decision making

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating decision models: effectiveness evidence
3.3: Meta-analysis: introduction

Pedro Saramago, PhD
Research Fellow, CHE

Objectives

- Definition and principles of meta-analysis
- Why conduct a meta-analysis?
- Assumptions of meta-analysis

A definition of meta-analysis…

- A statistical method of quantitatively combining results from multiple studies which address a common scientific question to reveal the nature of relationships that exist among relevant variables

  ultimately, a meta-analysis aims at quantifying effect sizes and their uncertainty

In practice, meta-analysis …

- Pools estimates of *effect measures* across studies with consideration for within/between-study variability

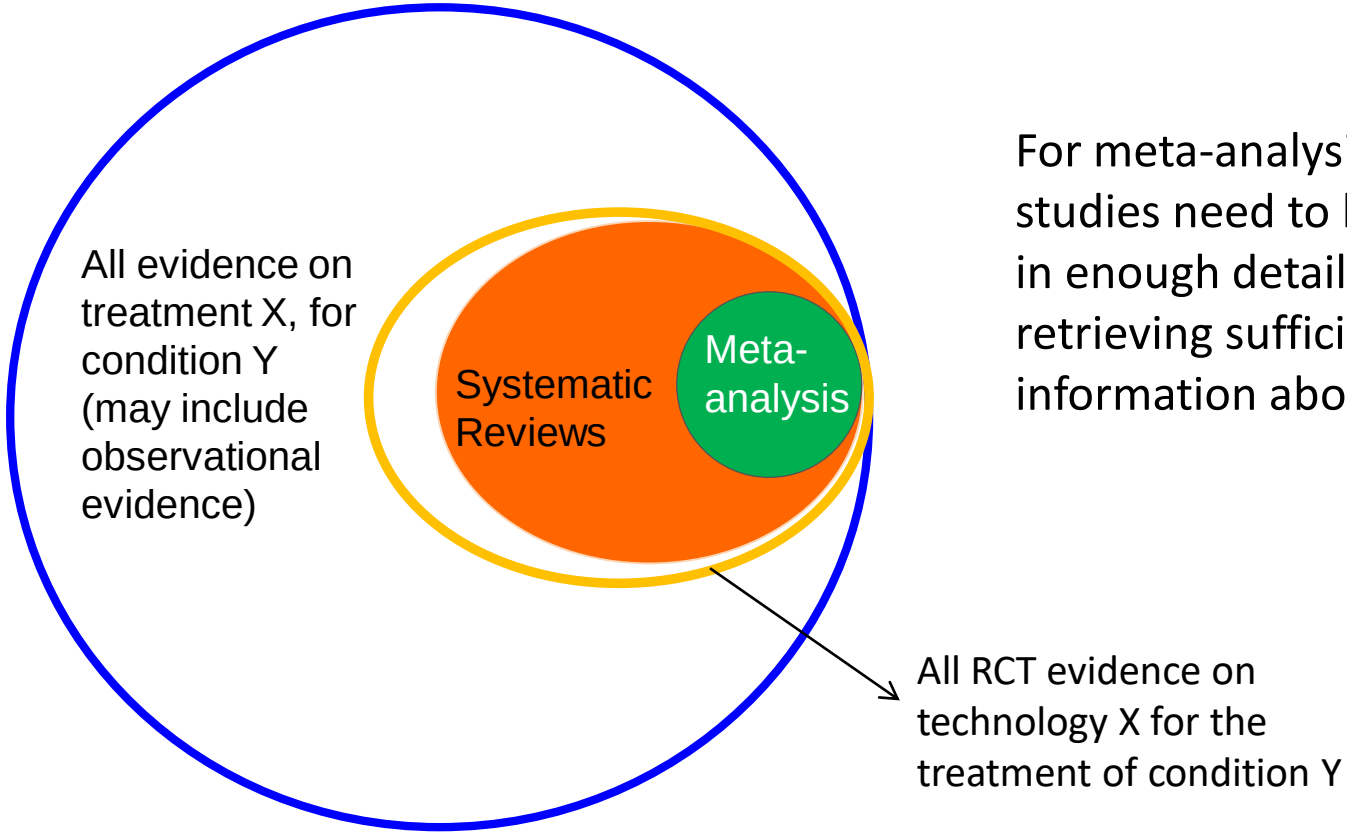
Why conduct meta-analyses?

THE EMERGENCE OF A NEW SPECIES:
THE PROFESSIONAL META-ANALYST

1. *'to obtain increased power'*
2. *'to obtain the best risk estimate from many, often conflicting or even bewildering, studies. In its best form, it is an attempt to clarify some of the heterogeneity between studies'*
3. *'to answer a question which the original studies were not aimed at'*

Source: Rosendaal 1994. *J Clin Epidemiol*; 47(12): 1325-1326

Evidence-base and evidence synthesis?



For meta-analysis, empirical studies need to be reported in enough detail to permit retrieving sufficient information about effects

What info is needed for meta-analysis?

For a particular effect measure, within each study one should extract summary information that can be used to calculate:

- a central/point estimate (e.g. in the format of a RR or OR)
- a measure of uncertainty (typically a standard error or confidence interval)

Note that, given most relative effect measures are ratios, their distribution is expected to be skewed. These are often transformed onto the log-scale (i.e. natural logarithm) for pooling

Important assumptions of meta-analysis …

- **Independence**: It is assumed that study samples are independent
- **Similarity**: It is assumed similarity across studies and some level of homogeneity in their findings

Summary points

- Meta-analysis combines studies estimating a common effect across studies
- Meta-analysis typically requires from studies a point estimate and a measure of uncertainty
- Independence and similarity of studies are key assumptions of meta-analysis

Meta-analysis with head-to-head trial data (standard pairwise MA)

- Focus on binary outcomes (e.g. dead or alive) with OR

- Fixed effect model
- Random effects model
- Subgroups and meta-regression

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating decision models: effectiveness evidence
3.4: Meta-analysis: Fixed- and random-effects pairwise meta-analysis

Pedro Saramago, PhD
Research Fellow, CHE

Objectives

- Understanding forest plots
- Fixed-effect approach and its assumptions
- Concept of heterogeneity, how it can affect a meta-analysis and how can we test for its presence
- Random-effects approach and its assumptions

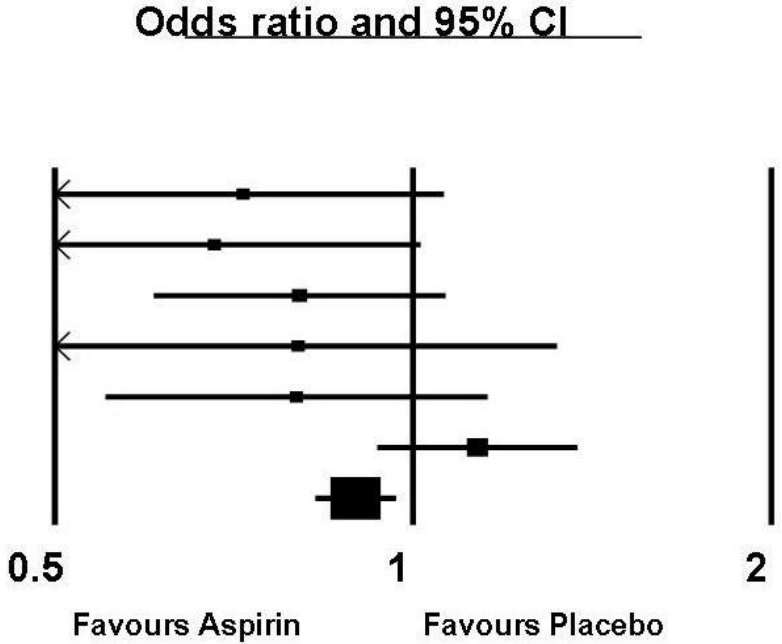
Example: Meta-analysis of RCTs of the effect of aspirin in preventing death after myocardial infarction

Study	Aspirin group		Placebo group	
	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

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

Forest plot of odds ratios – example

Study name	Odds ratio	Lower limit	Upper limit
MRC-1	0.72	0.49	1.06
CDP	0.68	0.46	1.01
MRC-2	0.80	0.61	1.06
GASP	0.80	0.49	1.32
PARIS	0.80	0.55	1.15
AMIS	1.13	0.93	1.37
ISIS-2	0.89	0.83	0.97

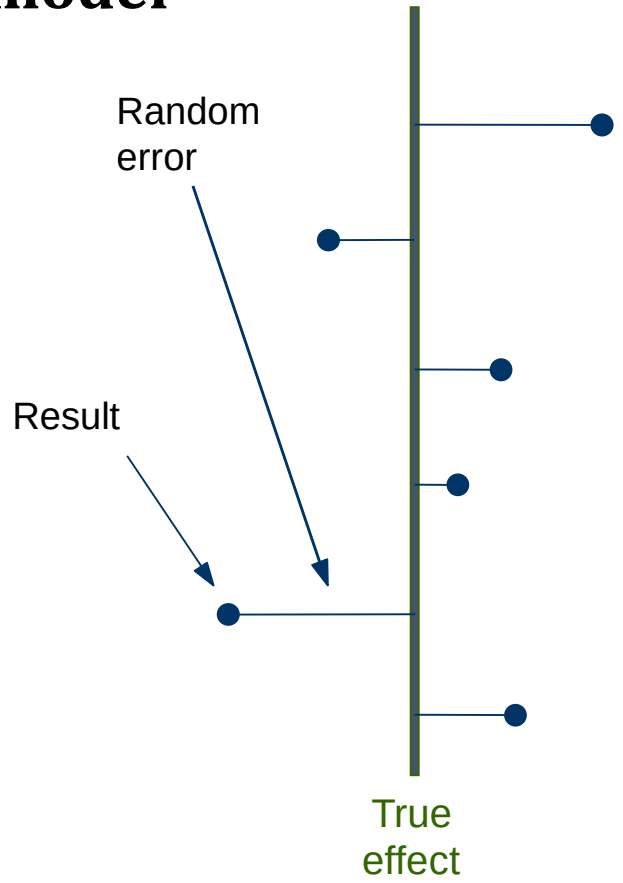


Meta-Analysis, Fixed effects model

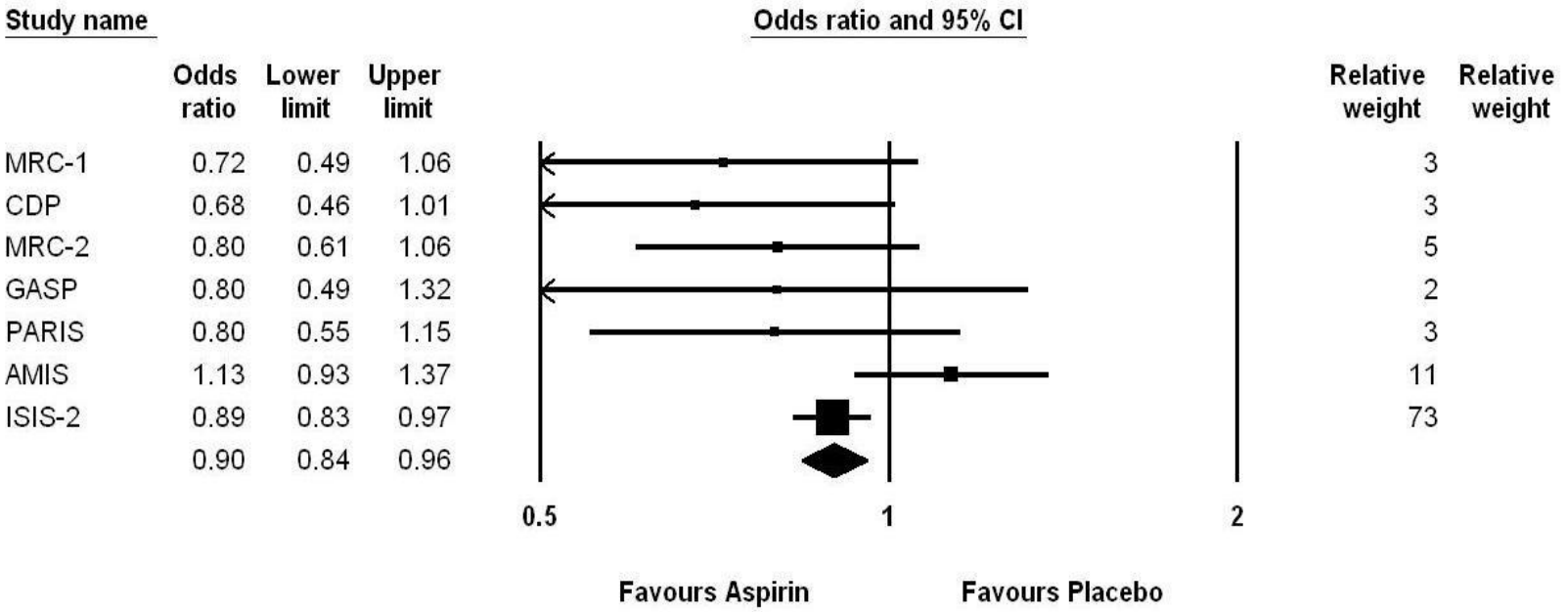
- Statistical homogeneity
- Formally, MA FE assumes:

$$Y_i \sim \text{Normal}(\theta, V_i)$$

- We estimate the common true effect, θ



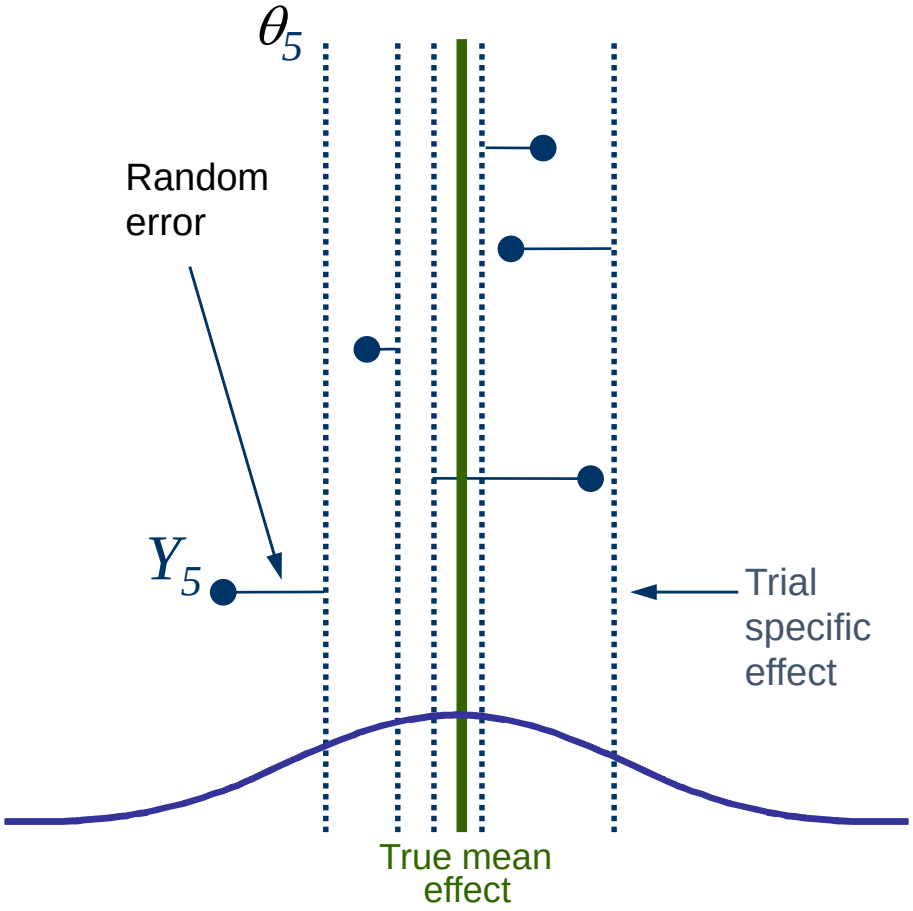
Fixed effect results – example results



Heterogeneity

- Fixed effect model assumes all the trials are estimating exactly the same treatment effect
- Is this reasonable?
  - Trials may differ in design and conduct in many ways including the characteristics of the patients or intervention (e.g. dose of drug)
- Statistical tests exist to assess its existence:
  - Cochran's Q test (aka χ^2): heterogeneity exists if p-value small (<0.1)
  - Higgins' I^2: how much of the total variability is due to heterogeneity? 0% (no) to 100% (much) scale

Random effects model



Model:

- within studies

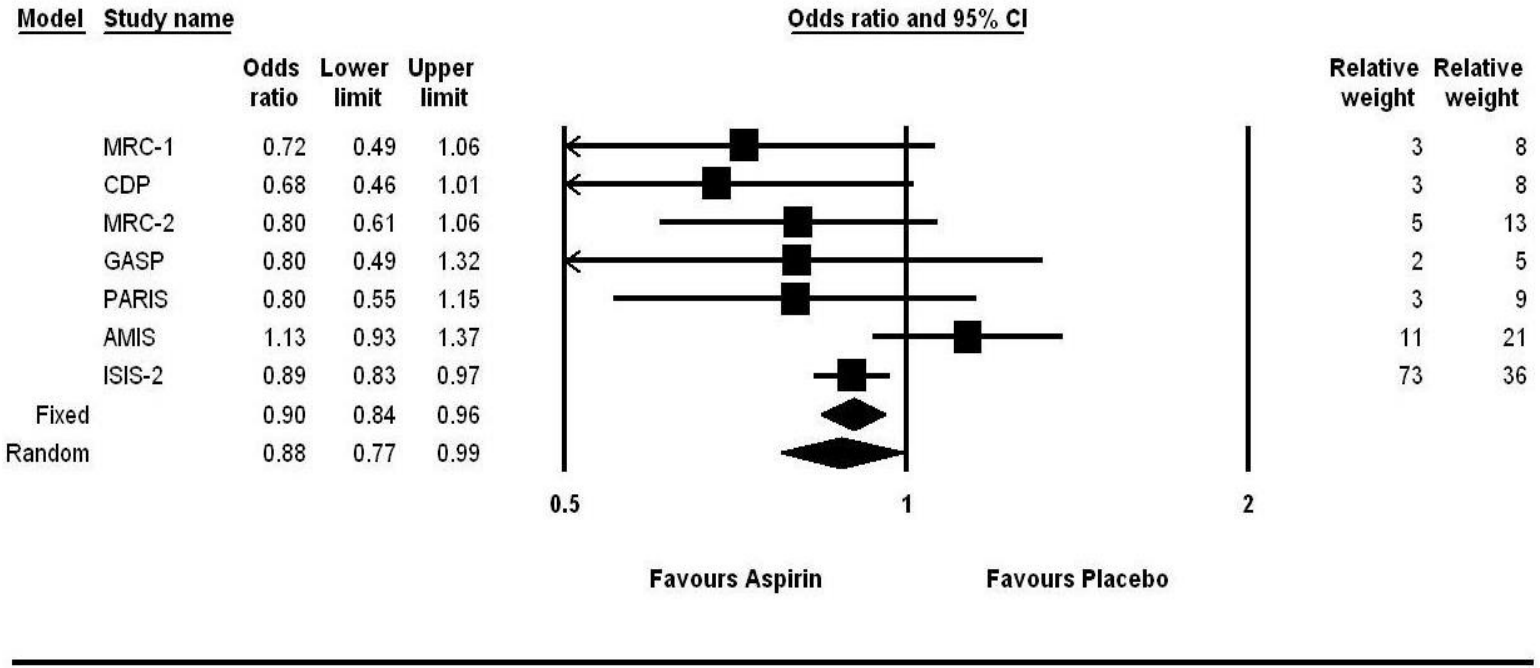
$$Y_i \sim \text{Normal}(\theta_i, V_i)$$

- across studies

$$\theta_i \sim \text{Normal}(\theta, t^2)$$

- There is a distribution of effect θ
- $\text{Weight}_i = 1 / (V_i + t^2)$

Comparison – fixed and random effects – example results



Exploring heterogeneity?

- What causes between study heterogeneity?
  - Clinical heterogeneity, e.g. differences in patients, disease severity, medical history
  - Methodological heterogeneity, e.g. differences in study intervention/conduct, e.g. randomisation, endpoints and time points
  - Chance
- Random effects model only account for it – they do not explain it
- Subgroup analyses/meta-regression methods can help to explain heterogeneity which may provide further insight into the treatment effect

Summary points

- Fixed-effect and random-effects approaches make different assumptions regarding the true effect estimate
- Study diversity may lead to heterogeneity: differences in true effect
- Existence of heterogeneity may be tested and heterogeneity can be quantified

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating decision models: effectiveness evidence
3.5: Meta-analysis: Exploring between-study heterogeneity

Pedro Saramago, PhD
Research Fellow, CHE

Objectives

- Using subgroup analysis and meta-regression to investigate heterogeneity
- Problems with exploring heterogeneity

Subgroup analysis

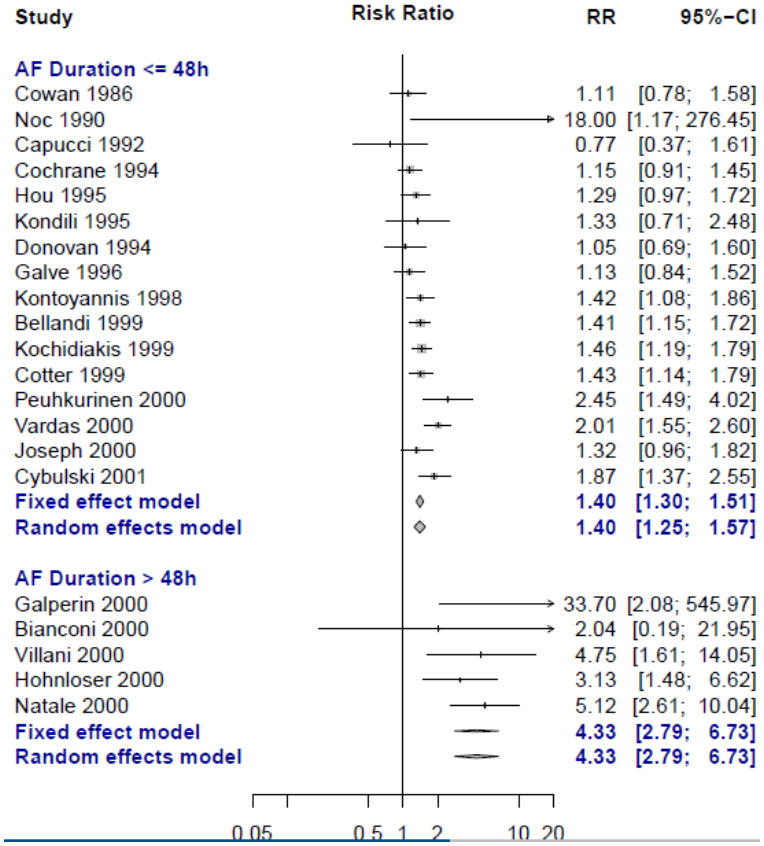
- Split evidence base into subgroups, according to some relevant factor (binary, categorical or continuous), e.g. high risk *vs* low risk (patient-level), high quality *vs* low quality studies (study-level)
- Perform a meta-analysis within each subgroup – a subgroup analysis
- Assess evidence within and across subgroups, comparing meta-analysis subgroup results

Example: Amiodarone for Conversion of Atrial Fibrillation (AF)

- 21 randomised or quasi-randomised controlled trials with AF patients
- Amiodarone compared with placebo, digoxin, CCB, or no treatment
- Primary outcome: conversion to sinus rhythm within 4 weeks (binary outcome)
- A priori specified subgroup analysis: mean duration of the current AF episode (<=48 vs > 48 hours)

Source: Letelier *et al*. (2003), *Arch Intern Med*; 163: 777-85

Example: Amiodarone for Conversion of AF – forest plot with subgroups



Source: Letelier *et al*. (2003), *Arch Intern Med*; 163: 777-85

Meta-regression model

$$Y_i = \theta_i + \beta x_i + e_i \qquad \theta_i \sim N(\theta, \tau^2),$$

- Study level covariates may be included in meta-analysis models to explore and adjust for systematic differences between studies
- β is the regression coefficient for the covariate included, x_i is the value of the covariate for the *i*th study
- This is a random effects model with a study level covariate added
- Similar in principle to standard linear regression, but regression is weighted to take into account the different sizes of studies
- Extends naturally to multiple covariates

Example: Primary angioplasty for AMI

- Previous meta-analyses that compare thrombolysis and primary angioplasty (PCI) following acute MI have shown significant clinical benefits from angioplasty in terms of reducing major adverse clinical events
- Thrombolytic treatment remains the default treatment option in many countries (including the UK).
- Possible reasons include additional delay in initiating reperfusion treatment through angioplasty
- Meta-regression to evaluate the relationship between treatment effect and the time delay involving initiation of angioplasty

Source: Asseburg *et al. Heart* 2007; 93: 1244-1254

Example: Meta-Regression Results

	Alternative time delays		
	30 minutes	**60 minutes**	**90 minutes**
Endpoint	***Odds ratio***	***Odds ratio***	***Odds ratio***
Death	**0.54 (0.29, 0.92)**	**0.77 (0.44, 1.29)**	**1.15 (0.49, 2.36)**
Non-fatal re-infarction	**0.30 (0.14, 0.59)**	**0.39 (0.21, 0.72)**	**0.55 (0.2,9 1.27)**
Non-fatal stroke	**0.47 (0.05, 0.69)**	**0.56 (0.09, 0.75)**	**0.79 (0.08, 1.43)**

Example: Cost-effectiveness results (ICERs)

Base case: 'AVERAGE' DELAY

Time delay	Treatment	Mean costs	Mean QALYs	ICER
Average trials (54.3 minutes)	Primary PCI	£12,760	7.12	£9,241
	Thrombolytics	£10,080	6.83	NA

Scenario analysis: ALTERNATIVE DELAYS

Time delay	Treatment	Mean costs	Mean QALYs	ICER
30 minutes	Primary PCI	£12,820	7.23	£6,850
	Thrombolytics	£10,080	6.83	NA
60 minutes	Primary PCI	£12,750	7.09	£10,269
	Thrombolytics	£10,080	6.83	NA
90 minutes	Primary PCI	£12,670	6.87	£64,750
	Thrombolytics	£10,080	6.83	NA

Problems with subgroup-analysis / meta-regression

- Often too few studies: insufficient data to detect relationships
- *Post hoc* conclusions, inflation of type I error, low power and spurious findings
- Statistically non-significant relationships should not be equated to absence of true relationships.
- Observational relationships – aggregation or ecological bias
- Only subset of trials may have covariate information – potentially biasing results.

Summary points

- Subgroup analysis and meta-regression may be used to investigate heterogeneity
- Applying these approaches may present several problems and results should be interpreted with caution

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Population decision models: effectiveness evidence
3.6: Network meta-analysis: introduction

Pedro Saramago, PhD
Research Fellow, CHE

Objectives

- Limitations of pairwise meta-analysis
- Taxonomy of comparisons
- Indirect comparisons, the building blocks

Recognition of limitations of conventional approaches to synthesis for decision making

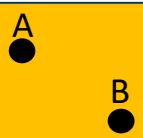
- Increased recognition of potential limitations for decision making purposes
  - The technologies of interest may not have been compared in head-to-head trials
  - Even when head-to-head RCTs exist, additional evidence maybe considered relevant
  - Decisions often relate to multiple technologies (n>2)

Implications

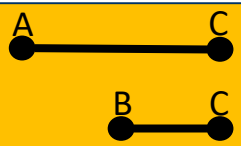
- May need to consider direct and indirect evidence
- Conventional methods of meta-analysis need extending to accommodate more complicated evidence structures
- Recognition that indirect and mixed treatment comparisons potentially offer a solution

A taxonomy of comparisons

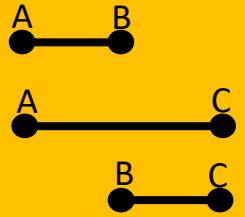
Direct Comparison



Naive or Unadjusted Indirect Comparison:
Absolute effect estimates from individual trial arms.



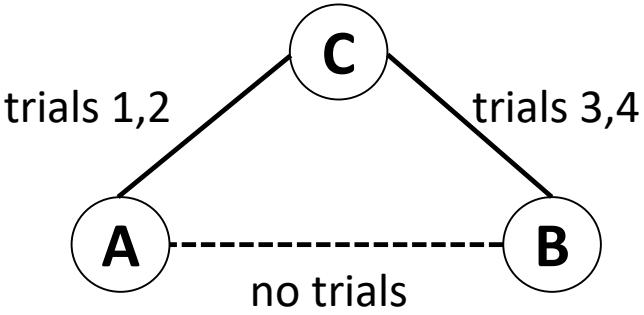
'Adjusted' indirect comparison:
Difference between relative treatment

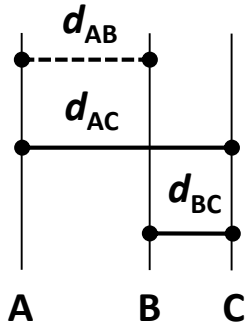


Network Meta-Analysis:
Adjusted Indirect comparison extended to more complex networks of trial evidence

Indirect and Mixed treatment Comparisons

- IC and MTC (or NMA) methods (generalisation of meta-analysis methods) allow comparisons of strategies not directly assessed within individual primary studies, ***without breaking within-study randomisation***
- No head to head evidence on *A* v *B* (comparison of interest)
- Treatment *C* is a common comparator of *A* & *B*
- Then $d_{AB\_ind} = d_{AC\_dir} - d_{BC\_dir}$, this is an *indirect comparison* of *A* vs *B*





$d_{AB} = d_{AC} - d_{BC}$

Summary points

- Conventional approaches to synthesis, like meta-analysis, present limitations for decision making
- IC and NMA are extensions of standard pairwise MA
- Indirect comparisons and network meta-analysis are valid approaches to enable the synthesis of all of the evidence

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Population decision models: effectiveness evidence
3.7: Network meta-analysis: assumptions

Pedro Saramago, PhD
Research Fellow, CHE

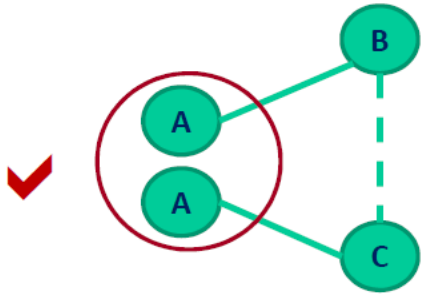
Objectives

- Key assumptions of network meta-analysis:
  - Similarity and homogeneity
  - Exchangeability and transitivity
  - Consistency

The basic assumptions underlying network meta-analysis

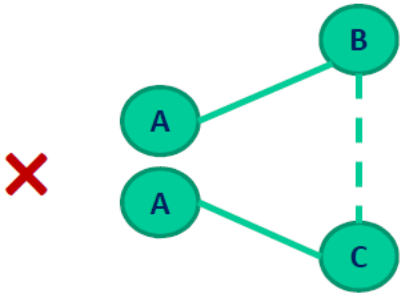
- **Similarity and homogeneity** (= MA)
- **Exchangeability/ Transitivity**
- **Consistency**

Similarity is…



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)?



Homogeneity is…

- The effect estimates across trials do not differ beyond what would be expected by chance

Exchangeability is…

- The ‘missing’ arm is missing at random
- A *vs* B do not have a ‘C’ arm and the A *vs* C studies do not have a ‘B’ arm

Transitivity is…

- A valid IC requires that the sets of AC and BC studies are similar in their distributions of effect modifiers – in which case the intervention effects can be assumed as transitive
- can be viewed as the extension of clinical and methodological homogeneity to comparisons across groups of studies that compare treatments.
- Note that similarity within each comparison in the network is not sufficient to justify the transitivity assumption.

Source: Cipriani, *Ann Int Med* 2013

Consistency is…

- ‘Direct’ and ‘indirect’ evidence are assumed to provide estimates of the same parameter (transitivity)
- What does this imply?
- The treatment effect, d_{AB}, estimated by the AB trials, is expected to be similar than the treatment effect estimated by the AC and BC trials if they had included B and A arms, respectively

= direct (AB) and indirect evidence (AC and BC) is not expected to differ beyond what can be explained by heterogeneity

- Consistency can only be tested when a loop in the evidence network exists

Summary points

- The key assumptions of NMAs are: consistency, exchangeability/ homogeneity, similarity and transitivity

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Population decision models: effectiveness evidence
3.8: Network meta-analysis: its role and examples

Pedro Saramago, PhD
Research Fellow, CHE

Objectives

- Detailed example:
  - The difference between direct and indirect treatment evidence
  - Combining direct and indirect evidence in a network
  - Benefits of the NMA approach for decision making

The role of network meta-analysis

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.

Angioplasty balloon device used to unblock and widen arteries

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

The need to combine direct and indirect evidence

Example: Thrombolysis for MI – problems with pairwise

- 7 (k) different treatments, 21 [k.(k-1)/2] possible pairwise comparisons (aka contrasts)
- Evidence on 10 direct pairwise comparisons
- Outcome: 35 day mortality

No of trials	Streptokinase	Alteplase-	Acclerated alteplase	Streptokinase +alteplase	Reteplase	Tenecteplase	PCTA
Boland et al[15]:							
8	P	P					
1	P		P	P			
1	P			P			
1	P				P		
2			P		P		
1			P			P	
Keeley et al[16]:							
8	P						P
3		P					P
11			P				P

PCTA = primary percutaneous transluminal coronary angioplasty.

Example: Thrombolysis for MI – problems with pairwise

35 day mortality, OR and 95% CI

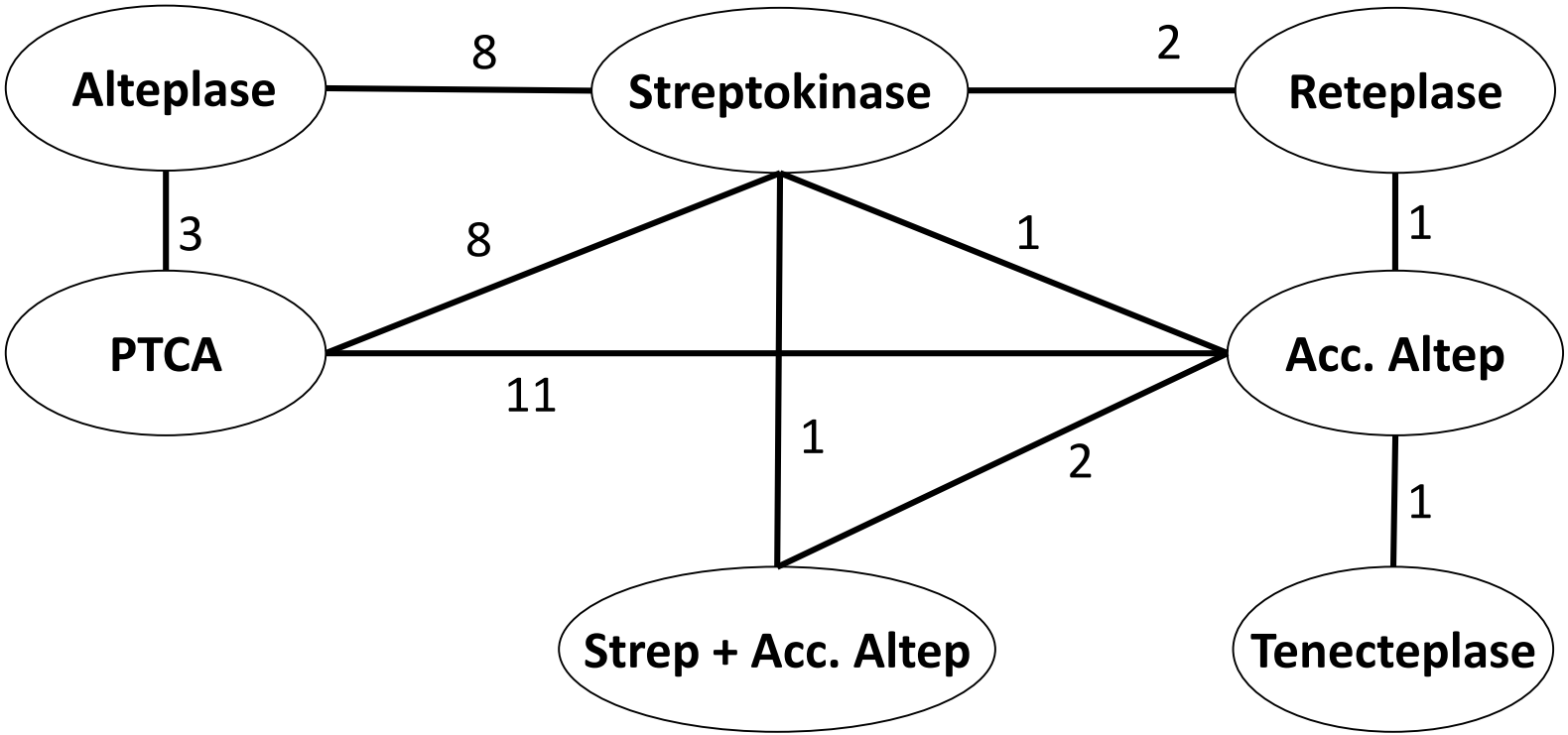
Treatment comparison	Direct comparisons
Streptokinase *v*:	
Alteplase	1.00 (0.94 to 1.06)
Accelerated alteplase	0.86 (0.78 to 0.94)
Streptokinase+alteplase	0.96 (0.87 to 1.05)
Reteplase	0.95 (0.79 to 1.12)
Tenecteplase	
PCTA	0.52 (0.36 to 0.73)
Alteplase *v*:	
Accelerated alteplase	
Streptokinase+alteplase	
Reteplase	
Tenecteplase	
PCTA	0.63 (0.25 to 1.29)

Treatment comparison	Direct comparisons
Accelerated alteplase *v*:	
Streptokinase+alteplase	1.12 (1.00 to 1.25)
Reteplase	1.02 (0.90 to 1.16)
Tenecteplase	1.01 (0.88 to 1.14)
PCTA	0.81 (0.64 to 1.02)
Streptokinase+alteplase *v*:	
Reteplase	
Tenecteplase	
PCTA	
Reteplase *v*:	
Tenecteplase	
PCTA	
Tenecteplase *v* PCTA	

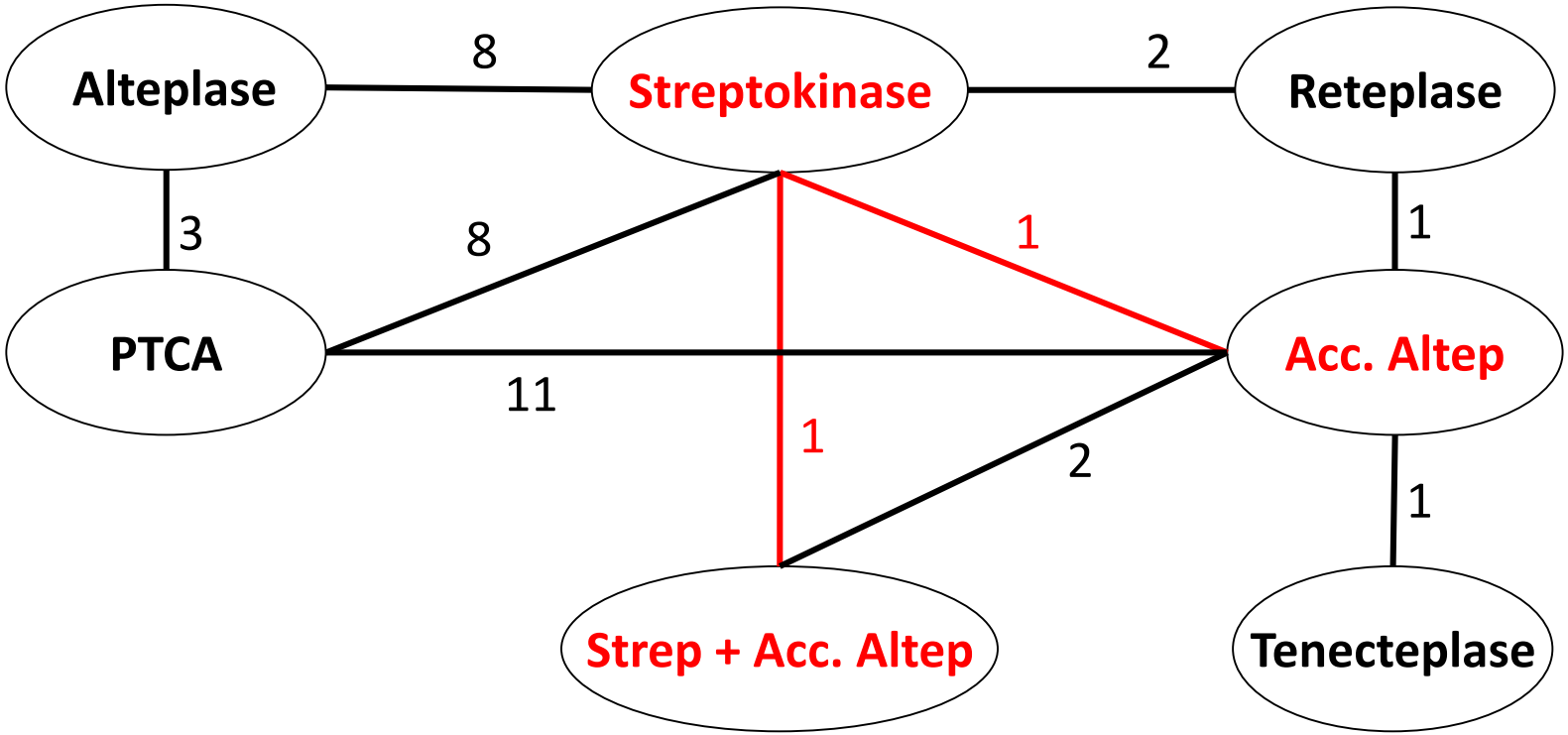
Thrombolysis for MI – interpretation example results?

- “Definitive conclusions on efficacy are that streptokinase is as effective as accelerated alteplase, that tenecteplase is as effective as accelerated alteplase, and that reteplase is at least as effective as streptokinase.”
  - Difficult to draw a conclusion about which treatment is ‘best’
  - Only represents subset of relevant alternatives
- Require simultaneous comparison of all relevant options to establish the most effective (and cost-effective)

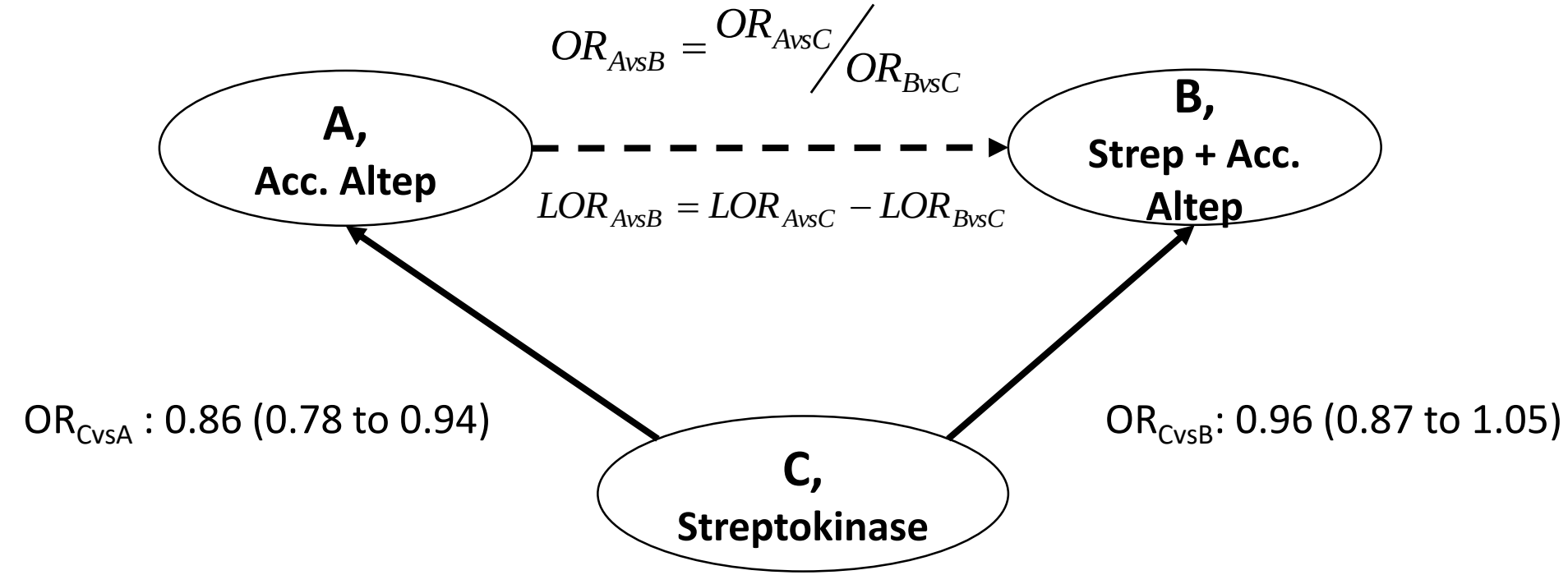
Example: Thrombolysis for MI – Network of evidence



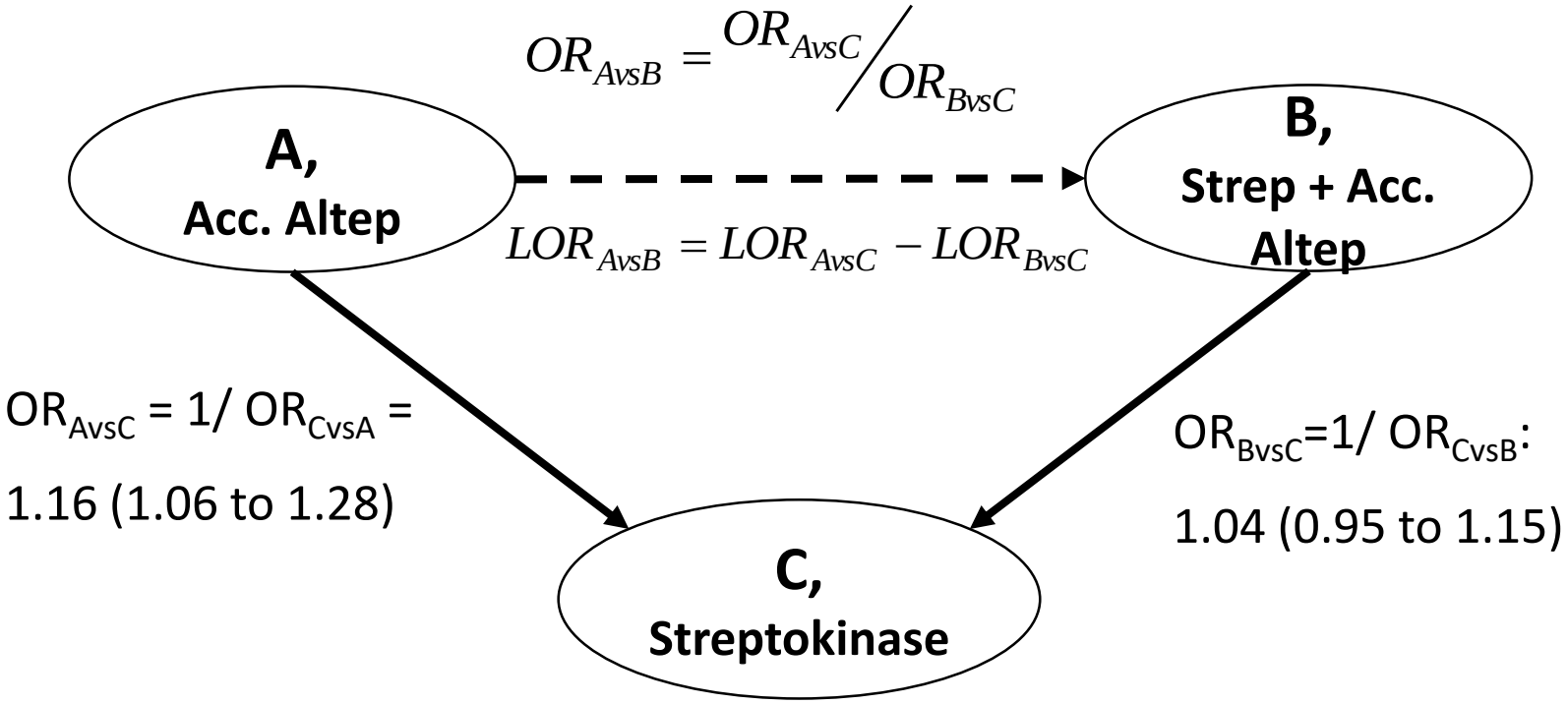
Example: Indirect comparisons (IC)



Example: Indirect Comparison (IC)



Example: Indirect Comparison (IC)



Uncertainty in Indirect Comparisons - *Bucher et al.*

- Variance of the indirect comparison:

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

- Standard Error (SE) on the Log scale:

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

- Mean and 95% Confidence interval for the OR:

$$mean = \exp\left(\ln OR_{AvsB}\right)$$

$$95\%CI = \exp(\ln OR_{AvsB} \pm 1.96 \times SE(\ln OR_{AvsB}))$$

Source: Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997; 50(6): 683-691

Example: calculations for an Indirect Comparison

1) Calculate mean difference in log OR

$$\ln\left(OR_{AB}\right) = \ln\left(OR_{AC}\right) - \ln\left(OR_{BC}\right)$$

$$0.11 = 0.15 - 0.04$$

2) Calculate standard errors for mean difference in log OR

$$SE\left(\ln OR_{AB}\right) = \sqrt{SE\left(\ln OR_{AC}\right)^2 + SE\left(\ln OR_{BC}\right)^2}$$

$$0.065 = \sqrt{0.05^2 + 0.05^2}$$

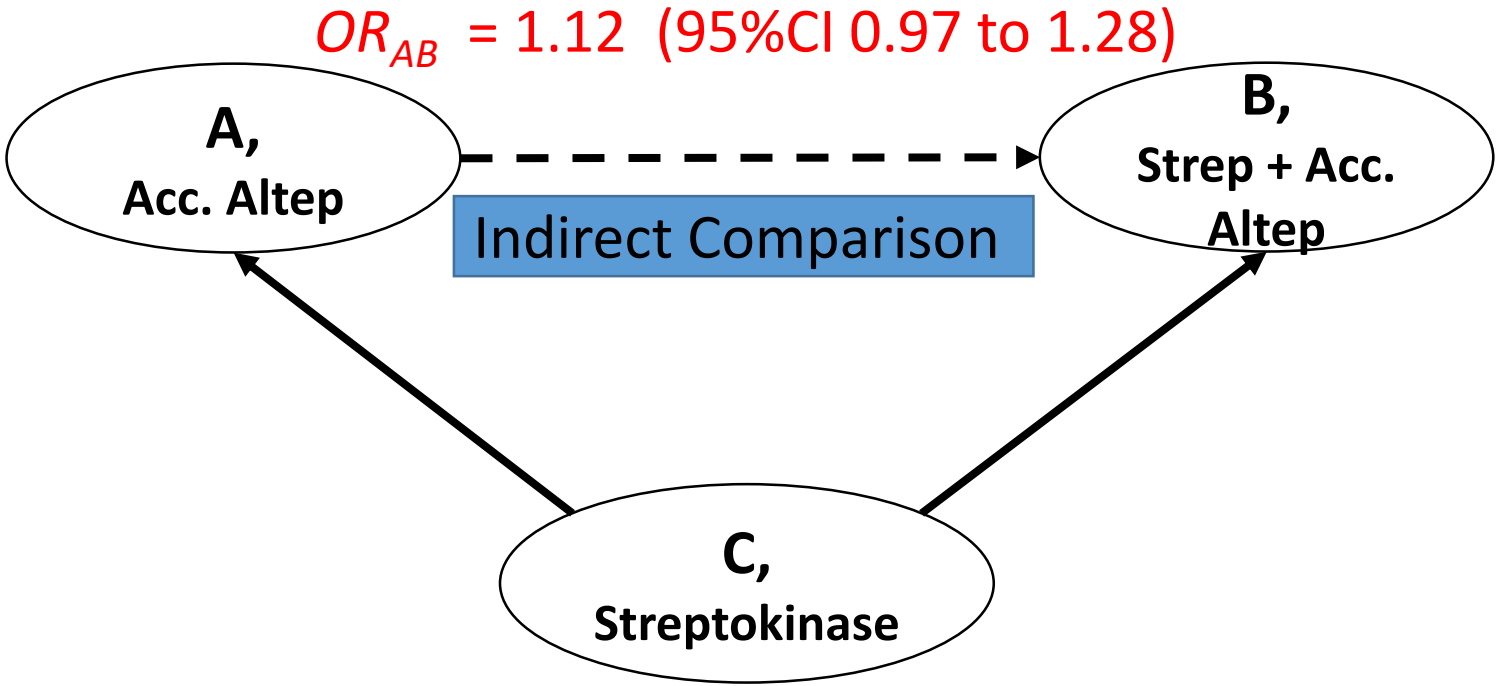
Note: for AvsC
ln(1.06)=0.06
ln(1.28)=0.25
CI width on ln scale=
0.25-0.06=0.19
Ln(SE(OR_{AC}))=
0.19/(2*1.96)=0.05

3) Exponentiate to get OR

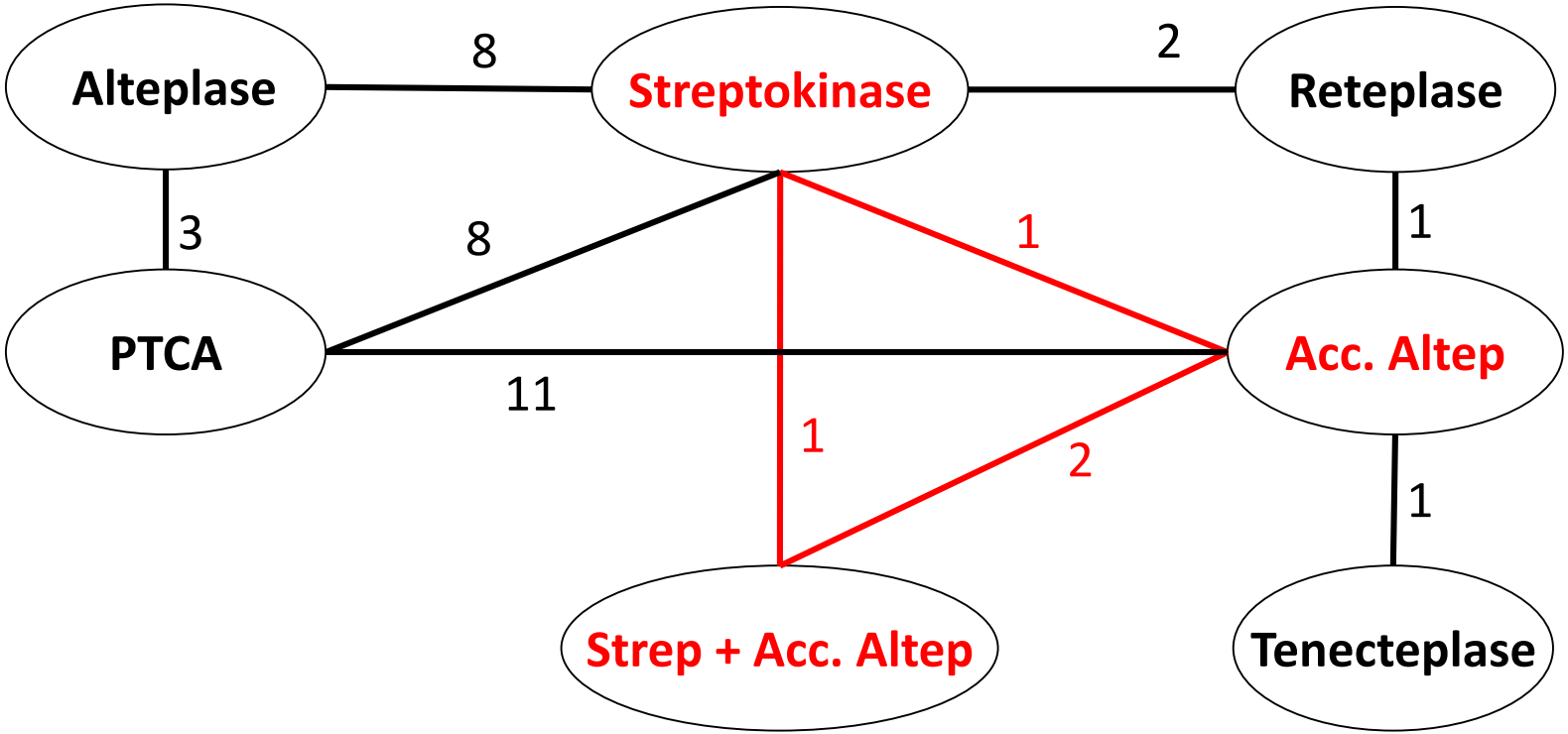
ln(OR_{AB}) = 0.11 (SE=0.065)

OR_{AB} = 1.12 (95%CI 0.97 to 1.28)

Example: Indirect Comparison



Example: Network meta-analysis



Example: Direct *vs* Indirect evidence and mixed estimates

Analysis	**OR** **Mean (95% CI)**	**LOR** **Mean (SE)**
Indirect *via* Streptokinase	1.11(0.97 to 1.28)	0.110 (0.064)
Direct (2 Trials)	1.12 (1.00 to 1.25)	0.113 (0.058)
Combined estimate	**1.12 (1.01 to 1.24)**	**0.112 (0.043)**

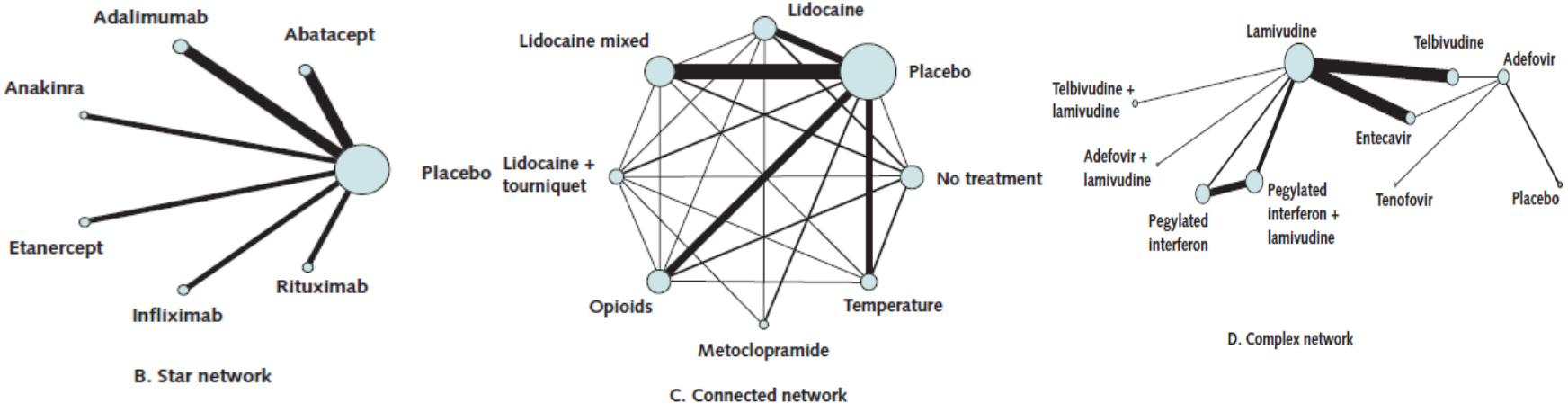
- Combined estimate should reflects both direct and indirect evidence, with relative weight dependent on the variance components

Example: Thrombolysis for MI – NMA results

- NMA 'fills in the blanks', i.e. estimates a relative treatment effect between all treatments of interest, simultaneously using all available evidence

	Fixed effect	
Treatment comparison	Direct comparisons	Multiple comparison
Streptokinase *v*:		
Alteplase	1.00 (0.94 to 1.06)	0.99 (0.94 to 1.06)
Accelerated alteplase	0.86 (0.78 to 0.94)	0.86 (0.78 to 0.93)
Streptokinase+alteplase	0.96 (0.87 to 1.05)	0.96 (0.87 to1.05)
Reteplase	0.95 (0.79 to 1.12)	0.90 (0.80 to 1.01)
Tenecteplase		0.86 (0.74 to 1.00)
PCTA	0.52 (0.36 to 0.73)	0.63 (0.52 to 0.77)
Alteplase *v*:		
Accelerated alteplase		0.86 (0.77 to 0.95)
Streptokinase+alteplase		0.96 (0.86 to 1.07)
Reteplase		0.90 (0.79 to 1.02)
Tenecteplase		0.86 (0.73 to 1.01)
PCTA	0.63 (0.25 to 1.29)	0.64 (0.51 to 0.77)
Accelerated alteplase *v*:		
Streptokinase+alteplase	1.12 (1.00 to 1.25)	1.12 (1.01 to 1.24)
Reteplase	1.02 (0.90 to 1.16)	1.05 (0.94 to 1.17)
Tenecteplase	1.01 (0.88 to 1.14)	1.01 (0.89 to 1.14)
PCTA	0.81 (0.64 to 1.02)	0.74 (0.61 to 0.89)
Streptokinase+alteplase *v*:		
Reteplase		0.94 (0.82 to 1.07)
Tenecteplase		0.90 (0.76 to 1.05)
PCTA		0.66 (0.53 to 0.81)
Reteplase *v*:		
Tenecteplase		0.96 (0.82 to 1.13)
PCTA		0.71 (0.57 to 0.87)
Tenecteplase *v* PCTA		0.74 (0.58 to 0.92)

Evidence networks can become very complicated



Source: Cipriani, Higgins, Geddes and Salanti. Conceptual and technical challenges in network meta-analysis. *Annals Int Med* 2013 Jul 16; 159(2): 130-7

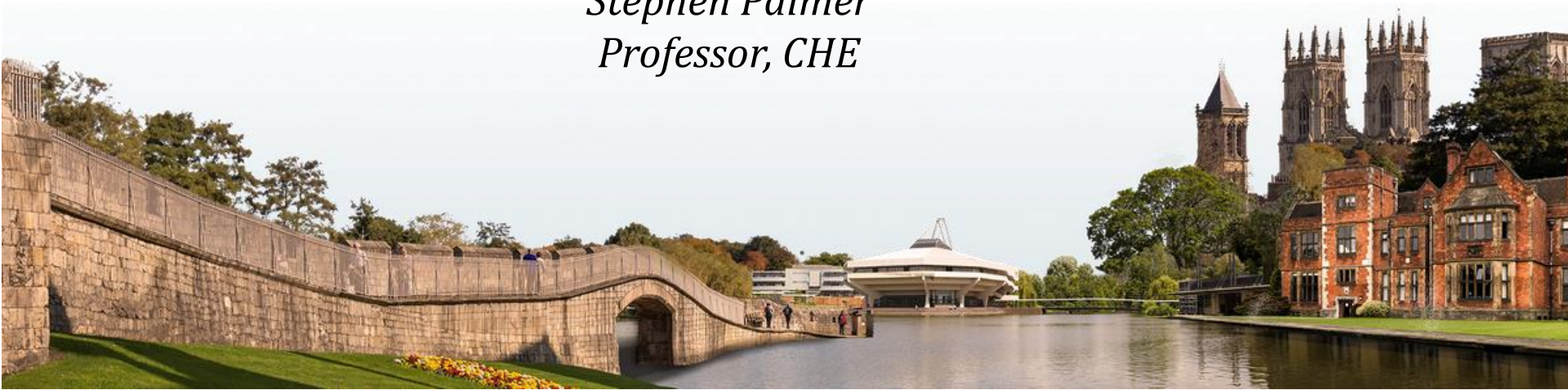
Summary points

- NMA uses all relevant evidence simultaneously in a single model
- NMA provides effect estimates for all comparisons of interest
- NMA provides all relevant information for decision making

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating models - effectiveness
3.9: Incorporating effectiveness evidence in decision models

Stephen Palmer
Professor, CHE



Objectives

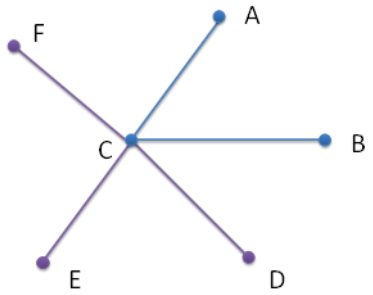
- Understand the general principles of applying (network) meta-analysis in decision models
- Identify appropriate methods to translate relative effects to absolute estimates of outcomes
- Understand how (network) meta-analysis results can be incorporated into different model structures
- Explore the role of synthesis for other types of input parameters

General principles of using (network) meta-analysis within models

- (Network) meta-analysis provides estimates of relative treatment effects compared to a reference treatment
  - e.g. standard of care, placebo, an active treatment
- Relative effect may be on a range of scales
  - e.g. risk difference (RD), relative risk (RR), odds ratio (OR), hazard ratio (HR)
- Models require absolute estimate of outcome
  - e.g. probability of death, probability of response

General principles of using (network) meta-analysis within models

- To translate NMA results from relative to absolute effects require absolute measure of outcome for one comparator
- Consider the following network



- A binary endpoint NMA on relative risk (RR) scale produces:

  RR_{AC}, RR_{BC}, RR_{DC}, RR_{EC}, RR_{FC}

  If p_C is the probability of the event for intervention C

  $p_A = RR_{AC} * p_C$

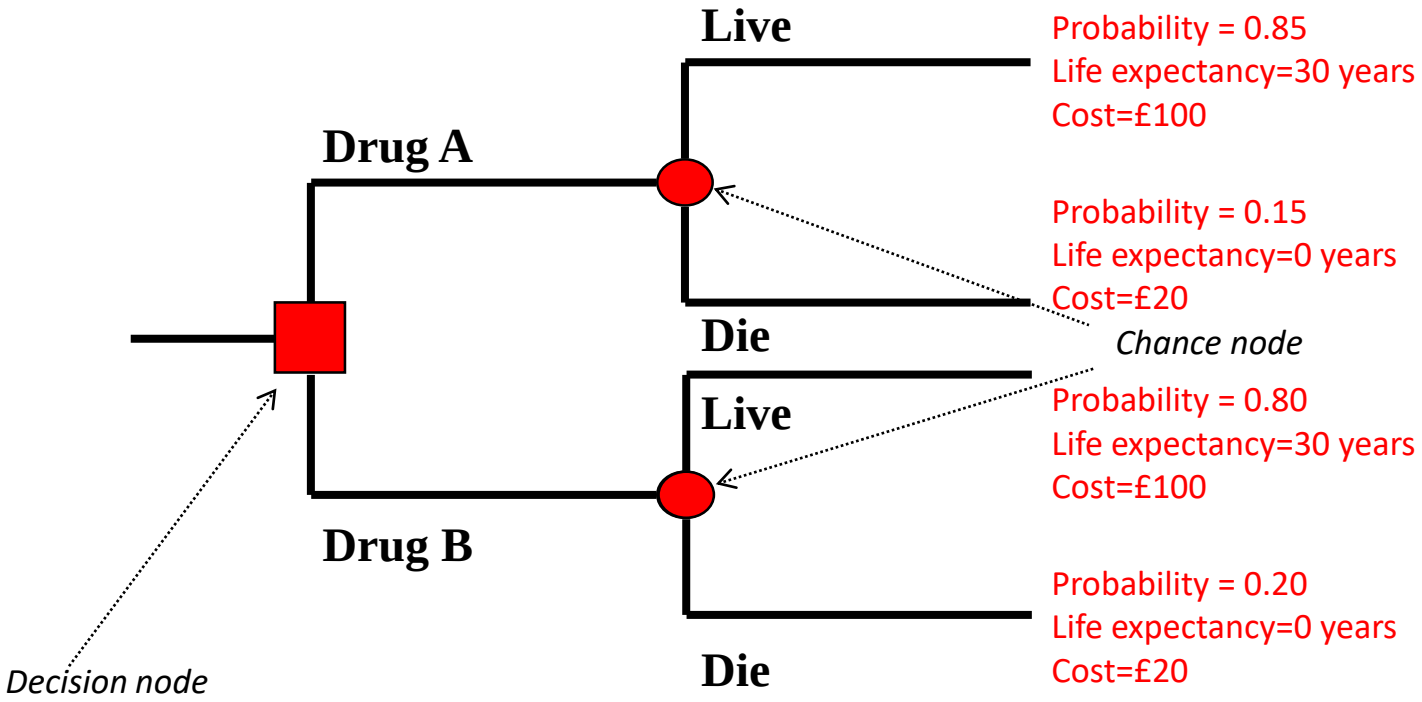
  $p_B = RR_{BC} * p_C$ etc.

Methods for translating from relative effects to absolute scale

Binary endpoint	
Risk difference	$p_A = p_C + d_{AC}$
Relative risk	$p_A = p_C \cdot d_{AC}$
Odds ratio	$odds_C = p_C/(1-p_C)$ $odds_A = odds_C \cdot d_{AC}$ $p_A = odds_A / (odds_A + 1)$
Continuous data	$\mu(t)_A = \mu(t)_C + d_{AC}$
Time to event data	$h(t)_A = h(t)_C \cdot d_{AC}$ A set of formulae exist to convert from hazards to survival curves or transition probabilities

Where p represents a probability, $\mu(t)$ the mean outcome at time t, and $h(t)$ the hazard at time t

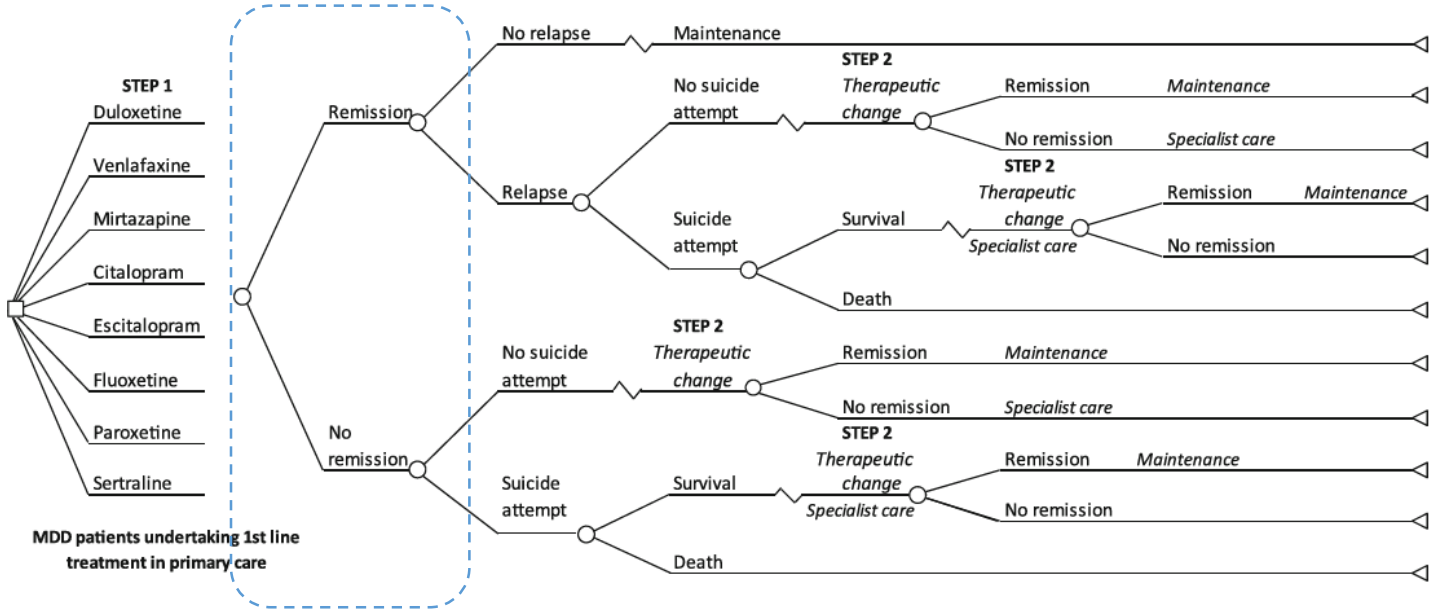
Incorporating NMA results in a decision tree



Life years A =
0.85*30+0.15*0=25.5
Cost A = 0.85*100+0.15*20=£88

Life years B =
0.80*30+0.20*0=24
Cost B = 0.80*100+0.20*20=£84

Example: pharmacological treatment for major depressive disorder



NMA informs probability of remission for each alternative

Annemans et al. Pharmacoeconomics 2014, DOI 10.1007/s40273-014-0138-x

Example: treatment for major depressive disorder

- Network meta-analysis of remission endpoint on odds ratio (OR) scale
  - Produced OR for each treatment vs fluoxetine
- Fluoxetine estimated to have remission probability of 40.21%
- Decision model requires probability of remission for each treatment

Example: treatment for major depressive disorder

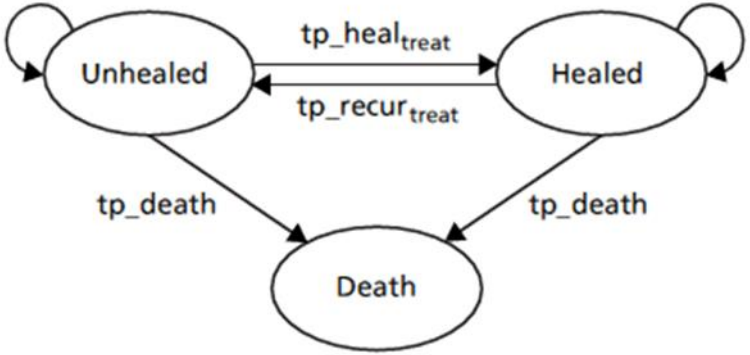
- OR for venlafaxine vs. fluoxetine from NMA: 1.25 (95% CI 1.12, 1.39)
- What is the probability of remission for patients receiving venlafaxine?
  1. Odds of remission for fluoxetine = 0.4021/(1-0.4021) = 0.6725
  2. Odds of remission for pts. receiving venlafaxine = 0.6725*1.25 = 0.8407
  3. Probability of remission for pts. receiving venlafaxine = 0.8407/(1+0.8407)=0.4568

Clinical data	Data (%)	References
Remission rate in first therapeutic step		
SNRIs[a]		
Duloxetine	44.99	Wessling and Ramsberg (TLV) [1]
Venlafaxine	45.68	Wessling and Ramsberg (TLV) [1]
Mirtazapine	45.08	Wessling and Ramsberg (TLV) [1]
SSRIs		
Citalopram	40.50	Wessling and Ramsberg (TLV) [1]
Escitalopram	47.56	Wessling and Ramsberg (TLV) [1]
Fluoxetine	40.21	Wessling and Ramsberg (TLV) [1]
Paroxetine	42.70	Wessling and Ramsberg (TLV) [1]
Sertraline	43.02	Wessling and Ramsberg (TLV) [1]

Annemans et al. Pharmacoeconomics 2014, DOI 10.1007/s40273-014-0138-x

Incorporating NMA results in a Markov model

- Markov structure used to calculate time in each states
- Costs and health outcomes assigned to each state
- Differences in time in each state drive differences across interventions

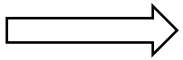


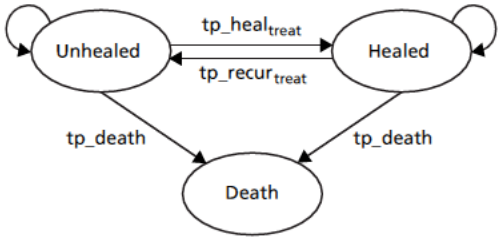
Example: Venus leg ulcers

- VenUS IV clinical trial compared four-layer bandage (4LB) with two-layer hosiery (HH) for management of venous leg ulcers.
- Decision problem included wider set of comparators
    1. Conducted NMA of alternative high-compression treatments
    2. Constructed decision model to assess cost-effectiveness
- Effects of treatment
    - Time to healing – all treatments have different rate (NMA)

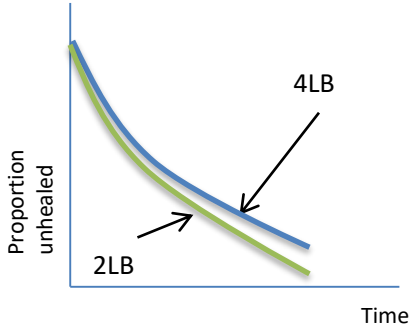
Example: Venous leg ulcers

		Base-case analysis Including VenUS IV (1)	
HRs		HR	95% CrI
Treatment effects (vs. 4LB)	SSB	0.88	0.76 to 1.03
	HH	1.05	0.85 to 1.29
	Paste	0.77	0.41 to 1.42
	2LB	1.40	0.65 to 3.05
	Ba	1.19	0.43 to 3.47
	BHeH	0.93	0.34 to 2.62
	BzeaH	1.33	0.42 to 4.51
	HV	1.00	0.23 to 4.22





- NMA Hazard ratios modify hazard "rate" of healing
  Example: $h_{2LB} = h_{4LB} \cdot HR_{2LBvs.4LB}$
- Essentially "shifts" the survival curve for time to healing
- Computationally achieved by modifying transition probabilities



Synthesis of other parameters

- Focus so far on relative effectiveness
  - Drive incremental differences between interventions
- Multiple source of evidence may also be available for other parameters
  - Natural history or 'baseline' event rate (e.g. remission rate with fluoxetine)
  - Costs/resource use
  - Utilities (health related-quality of life weights)

Reference: Saramago P, Manca A, Sutton AJ. Deriving input parameters for cost-effectiveness modeling: taxonomy of data types and approaches to their statistical synthesis. Value in Health. 2012 Aug 31;15(5):639-49.

Evidence synthesis for the baseline (natural history)

- Synthesis methods - effectively weighted pooling
- Recommend conducted separately from NMA
  - So that NMA unaffected by assumptions in baseline model
- Exploration of heterogeneity may be important
  - May be possible if subgroup data or individual patient data (IPD) available

Reference: Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Medical Decision Making. 2013 Jul;33(5):657-70.

Evidence synthesis of cost and resource use data

- Rarely done
- Require data specific to a jurisdiction
- Potential limitations to synthesis:
  - Methodological heterogeneity e.g. data collection strategies, methods of measurement
  - Temporal changes e.g. prices, technology
  - Differences between health systems

Evidence synthesis of utility data

- Methodological heterogeneity is a major issue
  - Different instruments (descriptive systems, valuation methods/populations)
- Guidance from NICE recommends synthesis only if studies share similar populations and use the same instrument and valuation
- Synthesis methods not well developed, generally weighted pooling

Reference: Papaioannou D, Brazier J, Paisley S. NICE DSU Technical Support Document 9: the identification, review and synthesis of health state utility values from the literature. University of Sheffield. 2011 Oct. Available from: http://www.nicedsu.org.uk/TSD9%20HSUV%20values_FINAL.pdf

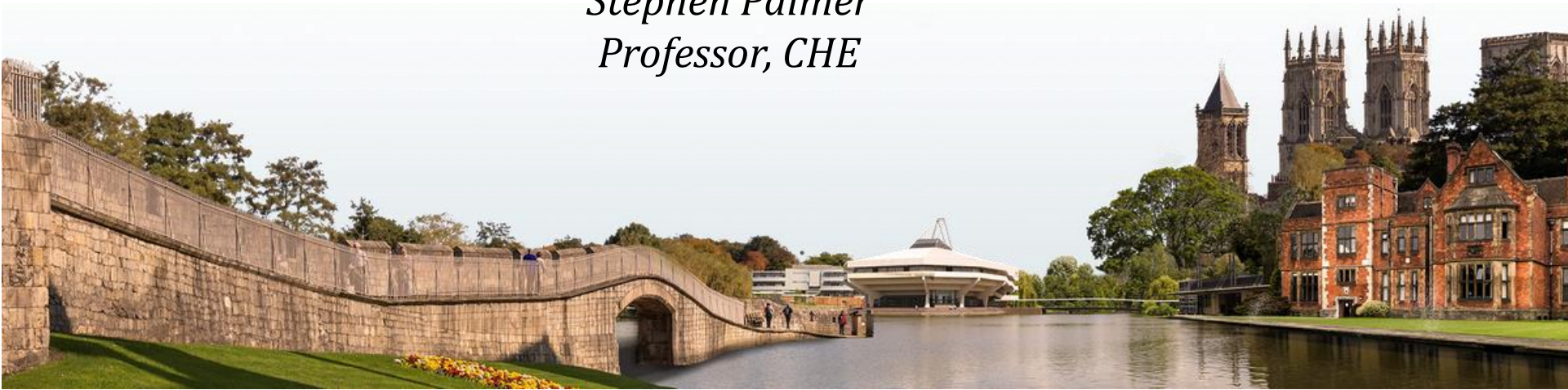
Summary

- Relative effects from (network) meta-analysis need to be translated into absolute estimates for decision making
- The specific methods depend on the outcome scale and the model structure
- Decision model provides framework to incorporate results from (network) meta-analysis and translate these into policy relevant outcomes (costs, QALYs)
- Need to consider role and appropriateness of synthesis for other input parameters

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating models - effectiveness
3.10: Reflecting heterogeneity in decision models

Stephen Palmer
Professor, CHE



Objectives

- Understand the importance of heterogeneity for policy decisions
- Identify different sources of heterogeneity for cost-effectiveness analysis
- Understand the importance of subgroup analysis and principles for cost-effectiveness analysis

Importance of heterogeneity: Policy context

- Growing demands to maximise value from limited health care budgets
- Political pressure to find a middle ground between refusal to reimburse and reimbursement as per license
- Restricted use: give to the sub-groups in which therapy cost-effective
- Traditional caution of trialists and EBM towards sub-group analysis

Subgroups and relative effectiveness

- Commonly used to explore difference in treatment effects between subgroups
  - “subgroup effect”, “effect modification”, or “interaction between a subgroup variable and treatment”
- Range of principles proposed for analysing and interpreting subgroup effects in clinical literature
  - Emphasis on relative effects
  - Use of baseline rather than post-randomisation characteristics
  - A priori specification (including direction)
  - Relatively few
  - Statistical significance
  - Biological rationale

Heterogeneity in cost-effectiveness analysis

- Individual patients vary in terms of different disease-related parameters
  - Underling risk of events (“baseline risk”)
  - Cost of an event
  - Utility of an event
- Individual patients vary in terms of treatment-related parameters
  - Relative risk for response to treatment (“treatment effects”)
  - Experience of adverse events
- Some of this variability can be explained by patient characteristics known at treatment initiation.
  - Age
  - Gender
  - Clinical characteristics

Importance of subgroups in cost-effectiveness

Heterogeneity in baseline risks and treatment effects

Example from RITA-3 economic analysis

	First quartile*	Second quartile*	Third quartile*	Fourth lower quartile*	Fourth upper quartile*
Age	45	52	52	61	66
Diabetes	0	0	0	0	1
Previous myocardial infarction	0	0	1	1	1
Smoker	0	1	0	1	0
Pulse	8	10	10	11	13
ST depression	0	0	1	1	1
Angina	1	0	1	0	0
Male	0	1	1	1	1
Left bundle branch block	0	0	0	0	0
ICER (no interaction)	49,754	22,145	20,765	11,682	12,490
ICER (interaction)	783,283	42,877	27,626	11,702	10,190

Henriksson *et al (2008)*. The cost-effectiveness of an early interventional strategy in non-ST-elevation acute coronary syndrome based on the RITA 3 trial. *Heart*; 94:717-23.

Ignoring heterogeneity - implications for population health

Subgroup (% patients)	QALYs	Costs	QALYs displaced (λ=£20,000/ QALY)	NHB in QALYs (λ=£20,000/ QALY)
Subgroup A (50%)	0.60	£10,000	⇨ 0.5	0.10
Subgroup B (50%)	0.30	£10,000	⇨ 0.5	-0.20
Average patient	0.45	£10,000	⇨ 0.5	-0.05

- Reimbursement as per license results in a <u>negative</u> population health gain per treated patient
- Restricting reimbursement to Subgroup A results in 0.10 QALY population health gain per treated patient
- Restricted decisions can improve population health

Principles of subgroup analyses for cost-effectivenes

- Relevant subgroups identified in terms of contribution to absolute treatment effect (baseline and relative effect)
- Pre-specification prior to data analysis based on clinical and economic plausibility
- Appropriate quantification of uncertainty rather than statistical significance
- Magnitude of population health gains (losses)
- Implementation issues in routine practice
  - Feasibility, costs, constraints

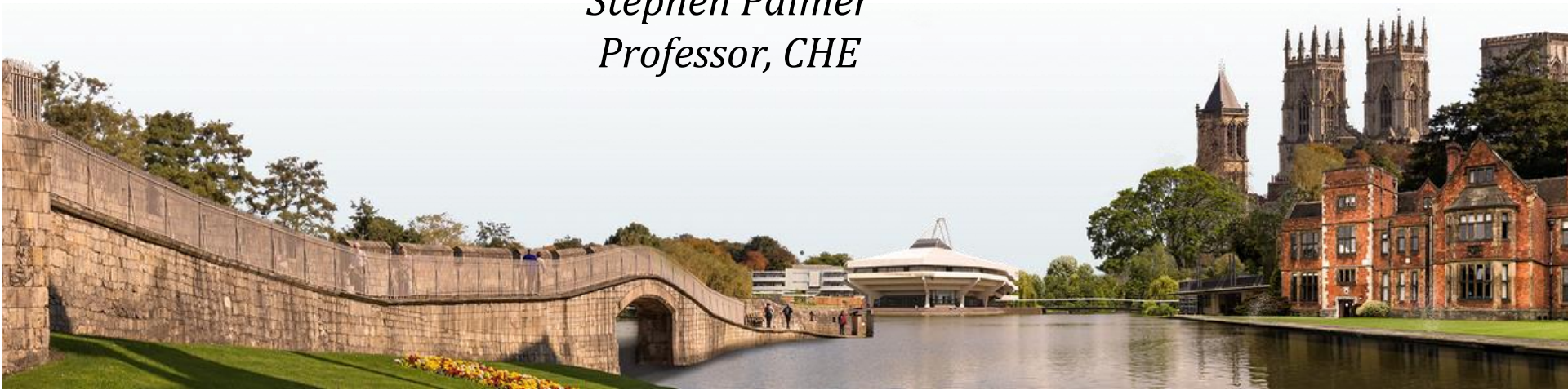
Summary

- Heterogeneity is an important issue for reimbursement and policy
- Heterogeneity manifests in different disease and treatment related parameters
- Important to establish principles of subgroups for cost-effectiveness
- Ignoring heterogeneity has implications for population health

Online Advanced Methods for Cost-Effectiveness Analysis

Presentation 3: Populating models - effectiveness
3.11: Summary and conclusions

Stephen Palmer
Professor, CHE



Summary

- Evidence synthesis is an important input to cost-effectiveness modelling
- Statistical tools, such as meta-analysis, enable the combination of quantitative results from multiple studies
- Subgroup analysis and meta-regression may be used to investigate heterogeneity
- Network meta-analysis can be used when there are multiple treatments providing estimates for all comparisons of interest
- Estimation of absolute outcomes required for evidence synthesis to be used within models
- Range of possible sources of heterogeneity should be explored

Further reading

Methods and application of network meta-analysis

- Bucher H et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997; 50(6): 683-691
- Caldwell D et al. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*. 2005 Oct 15;331(7521):897-900
- NICE DSU TSDs on Evidence Synthesis: TSD 1 to TSD 7 (available from: http://nicedsu.org.uk/technical-support-documents/evidence-synthesis-tsd-series/)

Application of evidence synthesis approaches for HTA

- Sutton A, Ades AE, Cooper N, Abrams K. Use of Indirect and Mixed Treatment Comparisons for Technology Assessment. *PharmacoEconomics.* 2008; 26: 753-767

Subgroups and heterogeneity

- Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. *Pharmacoeconomics*. 2008;26(9):799-806