Evidence synthesis for decision making: making best use of relevant evidence

Sofia Dias, PhD

Professor of Health Technology Assessment
Centre for Reviews and Dissemination
University of York, York, UK

RSS Sheffield Local Group Webinar
28 November 2023

With thanks to: Nicky J Welton & Hugo Pedder, University of Bristol

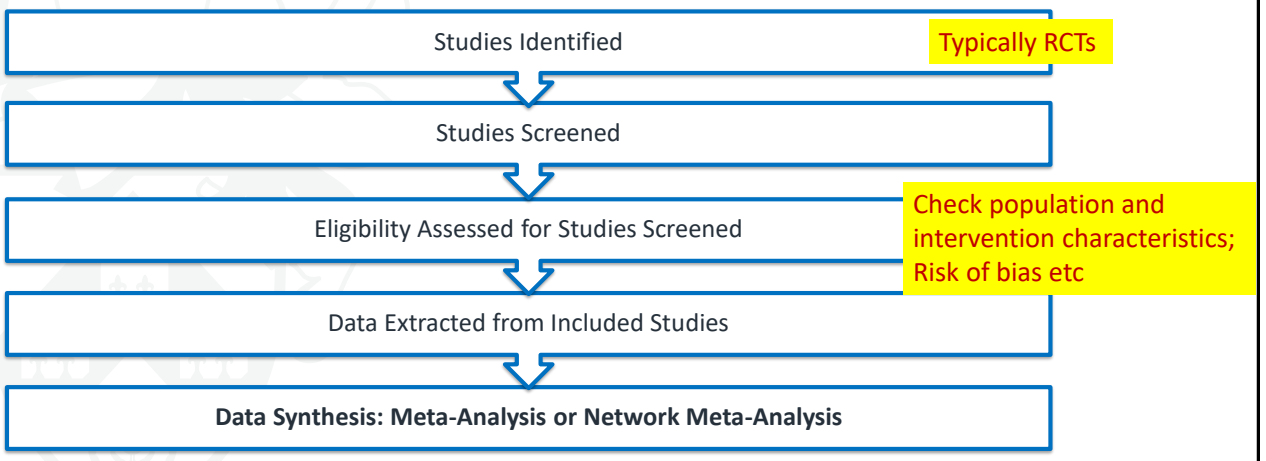
@sdias_stats
sofia.dias@york.ac.uk

Context

- Health technology assessment (HTA)
  "…a multidisciplinary process that uses explicit methods to determine the value of a health technology at different points in its lifecycle. The purpose is to inform decision-making in order to promote an equitable, efficient, and high-quality health system."
  Health Technology Assessment International http://www.htai.org/index.php?id=428
- In the UK
  - Manufacturer submissions to NICE for reimbursement decisions on new technologies
  - NICE writes guidelines for managing different conditions
  - Academic or other agency led projects…
- Usually involve a systematic review to collect and synthesise the relevant evidence
  - Typically a meta-analysis will be done
- Results of the evidence synthesis used in a decision model
  - Consider costs and benefits of alternative interventions
- New methods or extensions to current methods appear all the time
  - NICE Decision support Unit: critical review of existing and emerging methods for evidence synthesis on clinical effectiveness for decision-making in HTA, April 2020
    https://www.sheffield.ac.uk/nice-dsu/methods-development/chte2020-sources-and-synthesis-evidence

NICE: National Institute for Health and Care Excellence

Evidence selection



RCT: randomised controlled trial

3

Evidence synthesis in HTA

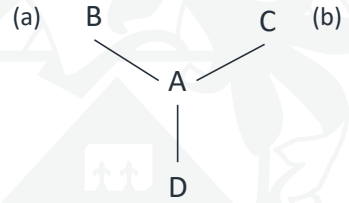
- Most HTAs will involve a meta-analysis of studies comparing the interventions of interest for the decision problem
  - Typically randomised controlled trials considered
  - Relative effects pooled across studies
- Often **multiple treatments** are of interest
- Network meta-analysis (NMA) is an extension of standard meta-analysis to incorporate direct and indirect evidence on multiple treatment comparisons in a coherent way
  - Increases precision and robustness of results as multiple sources of evidence are used to estimate the same relative treatment effect.
- Often most directly relevant evidence is sparse, leading to imprecise estimation of key parameters

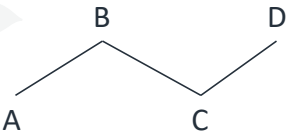
4

What is Network Meta-analysis?

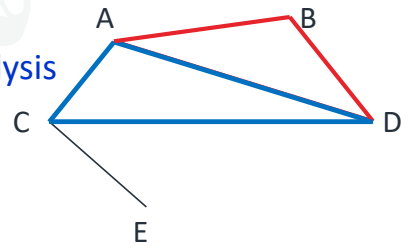
Pair-wise Meta-Analysis

Indirect Comparisons





Network meta-analysis



- The existence of "evidence loops" means that there is both DIRECT evidence and INDIRECT evidence on the same contrast
- More data → estimates more precise, more robust (less sensitive to any one source of data)
- Possible to estimate additional parameters.

5

Basic & Functional parameters

1. Four treatments 1, 2, 3, and 4
2. Take treatment 1 as the reference treatment
3. Then the treatment effects (log odds ratios) of 2, 3, and 4 relative to 1 are the basic parameters
4. Give them priors: $d_{12}, d_{13}, d_{14} \sim N(0,100^2)$

Remaining contrasts are functional parameters

$$d_{23} = d_{13} - d_{12}; \quad d_{24} = d_{14} - d_{12}; \quad d_{34} = d_{14} - d_{13}$$

- Consistency equations define these relationships: $d_{ck} = d_{1k} - d_{1c}$
- Consistency follows from the decision to synthesise the data and confidence in study selection criteria
  - Can/should be statistically checked when possible

Lu & Ades (2004) https://doi.org/10.1002/sim.1875

6

Inconsistencies

In a pre-specified population…

- The **true** treatment effects **must be** consistent
- But there may be inconsistencies in the **EVIDENCE**

- Check for conflict across different evidence sources
- Different methods proposed for
  - Global checks
  - Local checks
  - Active area of research

Dias et al. https://doi.org/10.1177/0272989X12455847

Notation for the Model: binary data

- Define

$r_{i,k}$ – the number of events in arm k of trial i
$n_{i,k}$ – the number of patients in arm k of trial i
$p_{i,k}$ – the probability of an event in arm k of trial i
$t_{i,k}$ – the treatment given in arm k of trial i

$k = 1,2,3, \ldots, na_i$
$i = 1, 2, \ldots, ns$

na_i=number of arms in study i
ns=number of studies
nt=number of treatments in the network

Random Effects NMA Model

For treatment in arm k in study i

Likelihood: $r_{i,k} \sim Binomial(p_{i,k}, n_{i,k})$

Probability of an event in arm *k* of study *i*

Model:

$$\text{logit}(p_{i,k}) = \begin{cases} \mu_i & k = 1 \\ \mu_i + \delta_{i,k} & k = 2,\ldots,na_i \end{cases}$$

log-odds of event in arm 1 of study *i* (note treatment in that arm)

log-odds ratio for treatment in arm *k* of study *i* compared to treatment in arm 1 of study *i*

$$\delta_{i,k} \sim N\left(d_{t_{i,k}} - d_{t_{i,1}}, \sigma^2\right)$$

Consistency equation, ensures correct comparison made

Between study variance (heterogeneity): Assumed **common** across treatment comparisons

μ_i is considered a nuisance parameter, not of interest
d_k is the log-odds ratio for treatment *k* compared with treatment 1
Set d_1 =0 (log-odds ratio for treatment 1 compared with itself)

9

Bayesian framework

PRIORS DISTRIBUTIONS

- Typically vague for relative treatment effect parameters
  - Informative prior distributions for these parameters are rarely used and would require strong justification.
  - Log-odds ratios can take any value

$$d_k \sim N(0, 100^2) \quad k > 1 \qquad (d_1 = 0)$$

$$\mu_i \sim N(0, 100^2) \quad i = 1,\ldots,ns$$

- Between studies heterogeneity often given a wide (scale-dependent) Uniform prior
  - eg Uniform(0,5) or Uniform(0,2)
- Specification of vague priors on variance components is complex area
- Bayesian approach allows incorporation of external evidence on heterogeneity, based on empirical evidence
  - Higgins and Whitehead (1996) https://pubmed.ncbi.nlm.nih.gov/8981683/
  - Turner et al. (2012, 2015) https://doi.org/10.1093/ije/dys041 , http://dx.doi.org/10.1002/sim.6381
  - Rhodes et al. (2014) https://doi.org/10.1016/j.jclinepi.2014.08.012
- Frequentist estimation is also possible

10

Generic NMA RE model

THE UNDERLYING MODEL IS

$$\theta_{ik} = \mu_i + \delta_{ik}$$

- θ_{ik}: continuous measure of the treatment effect
- μ_i: effect of baseline treatment in trial i
- δ_{ik}: trial-specific treatment effect of treatment in arm k relative to the treatment in arm 1

Priors

$$\sigma \sim U(0,5)$$

$$d_{1k} \sim N(0,100^2) \quad k > 1$$

$$\mu_i \sim N(0,100^2) \quad i = 1,\ldots,N$$

Set $\delta_{i1} = 0$

$$\delta_{ik} \sim N\left(d_{t_{i1},t_{ik}}, \sigma^2\right), k > 1$$

$$d_{t_{i1},t_{ik}} = d_{1,t_{ik}} - d_{1,t_{i1}}$$

RE: random effects

Making best use of evidence

RELEVANT EVIDENCE

- Often restrict dataset to only include treatments of interest
- Leads to sparse networks
  - Few comparisons (no indirect evidence)
  - Few studies per comparison
- Wastes evidence that has been generated and could provide more meaningful inferences
  - Particularly important when quantification of relative treatment effects and associated uncertainty is used for decision-making

Making best use of evidence

BASIC PRINCIPLES TO INCREASE PRECISION

- Evidence synthesis must use all relevant evidence
- Avoid arbitrary selection of studies or outcomes
- Assumptions must be transparent and acceptable to all stakeholders
  - Often requires clinical opinion to validate
- Use "all relevant evidence" which may include
  - Different doses (but only licensed dose can be recommended)
  - Other (related) outcomes
  - Other (related) populations (children vs adults)
  - Different study types (observational as well as RCTs??)

13

Making best use of evidence

BORROWING INFORMATION ACROSS DOSES

- Intervention (treatment) is defined as a drug/agent **given at a particular dose**
- Whilst interventions at unlicensed doses may not strictly be of interest for decision-making, studies that compare different doses can provide additional, **relevant**, evidence for synthesis.
- Use a synthesis model that "borrows" information across doses
  - requires data on different doses of a drug of interest to be available
  - Combine with NMA to compare different agents, given at different doses

14

Model-based NMA (MBNMA)

- Model-based meta-analysis (MBMA) used in drug development to inform decision-making and future trial designs
  - uses plausible physiological time-course or dose-response models
  - Tends to be arm-based and **not** respect randomisation
- Model-based NMA combines MBMA with NMA
  - works at the level of the relative effects so respects randomization
  - allows estimation and prediction of treatment effects at multiple time points or doses
  - Allows assessment of evidence consistency across comparisons
- R packages available on CRAN:
  - MBNMAtime https://cran.r-project.org/package=MBNMAtime
  - MBNMAdose https://cran.r-project.org/package=MBNMAdose

Pedder et al. (2019) https://doi.org/10.1002/jrsm.1351
Mawdsley et al (2016) https://doi.org/10.1002/psp4.12091

15

Dose-response NMA

- Information sharing via "model-based" approach that functionally incorporates a dose-response relationship
  - Bayesian framework
- Dose-response function fitted to study-specific **relative effects**
  - Preserves within-study randomisation
  - Model fit compared to "split" NMA (where possible)
  - Assess consistency assumption
- Uses additional evidence from studies (or arms) of doses not of primary interest
  - borrow strength through consistency relationship ***and*** dose-response relationship

Mawdsley et al (2016) https://doi.org/10.1002/psp4.12091

16

Generic NMA RE model

Recap

THE UNDERLYING MODEL IS

$$\theta_{ik} = \mu_i + \delta_{ik}$$

- θ_{ik}: continuous measure of the treatment effect
- μ_i: effect of baseline treatment in trial i
- δ_{ik}: trial-specific treatment effect of treatment in arm k relative to the treatment in arm 1

Priors

$$\sigma \sim U(0,5)$$

$$d_{1k} \sim N(0,100^2) \quad k>1$$

$$\mu_i \sim N(0,100^2) \quad i=1,\ldots,N$$

Set $\delta_{i1} = 0$

$$\delta_{ik} \sim N\left(d_{t_{i1},t_{ik}}, \sigma^2\right), k>1$$

$$d_{t_{i1},t_{ik}} = d_{1,t_{ik}} - d_{1,t_{i1}}$$

RE: random effects

MBNMA dose-response model

$$\theta_{ik} = \mu_i + \delta_{ik}$$

- θ_{ik}: continuous measure of the treatment effect
- μ_i: effect of baseline treatment in trial i
- δ_{ik}: trial-specific treatment effect of treatment in arm k relative to the treatment in arm 1

treatement $t_{i,k} = \left(x_{i,k}, a_{i,k}\right)$

defined by dose $x_{i,k}$ of agent $a_{i,k}$

$$\delta_{i,k} \sim N\left(d_{(x_{i,1},a_{i,1}),(x_{i,k},a_{i,k})}, \sigma^2\right)$$

We define the relative effect of treatment (x_k, a_k) relative to reference (x_1, a_1) via a dose-response curve:

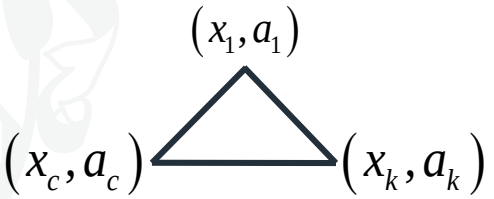
$$d_{(x_1,a_1),(x_k,a_k)} = f\left(x_k, \beta_{a_k}\right)$$

Treatment Level Consistency

FOR ALL TREATMENTS

- Apply consistency equation at the level of the dose response curve

$$d_{(x_c,a_c),(x_k,a_k)} = d_{(x_1,a_1),(x_k,a_k)} - d_{(x_1,a_1),(x_c,a_c)}$$



- Methods to check consistency available
  Pedder et al. (2022) https://doi.org/10.1002/sim.9270

19

Dose-response: Emax model

Dose-response function:

$$E_0 + f(x, \beta_a)$$

= 0 for placebo

treatement defined by dose x of agent a

$$f(x, \beta_a) = \frac{Emax_a \, x}{ED50_a + x}$$

20

Dose-response MBNMA

Dose-response function:

$$E_0 + f(x, \beta_a)$$

treatement defined by dose *x* of agent *a*

$$\theta_{i,k} = \begin{cases} \mu_i & k = 1 \\ \mu_i + \delta_{i,k} & k \geq 2 \end{cases}$$

i = study
k = arm

Dose-response function on relative effects

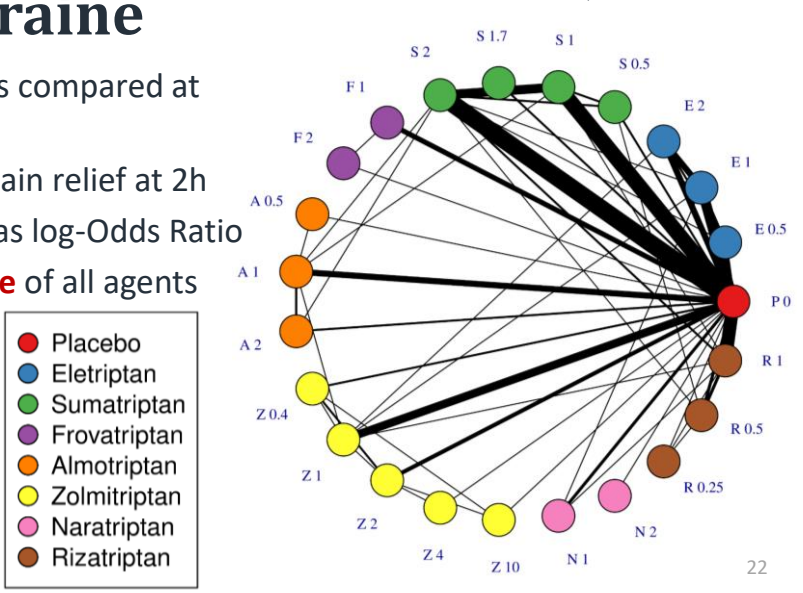
$$\delta_{i,k} = (E_{0,i} + f(x_{i,k}, a_{i,k})) - (E_{0,i} + f(x_{i,1}, a_{i,1}))$$

$$\delta_{i,k} = f(x_{i,k}, a_{i,k}) - f(x_{i,1}, a_{i,1})$$

21

Illustrative dataset: Triptans in migraine

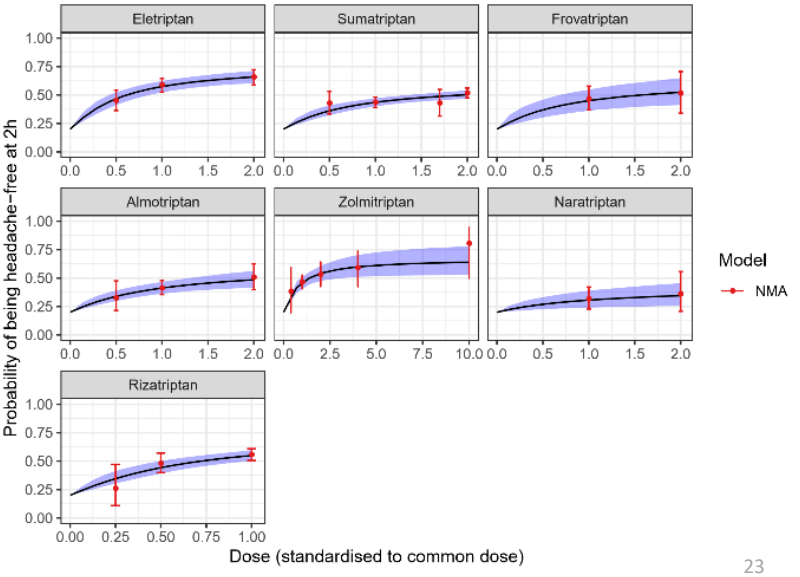
- 70 studies of 8 interventions compared at multiple doses
- Outcome: % patients with pain relief at 2h
- Treatment effect modelled as log-Odds Ratio
- Placebo treated as **zero dose** of all agents



22

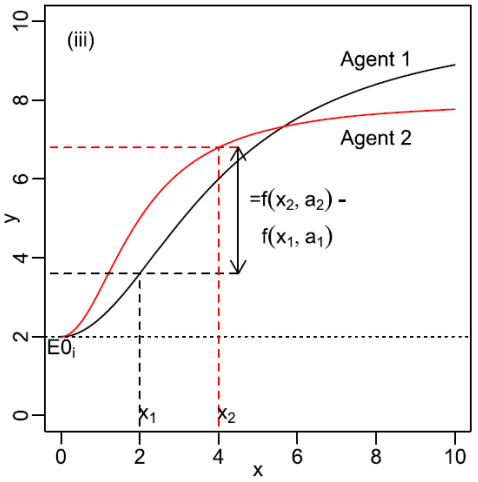
Triptans: connected network

- Fitting structural dose-response function increases precision versus standard "split" NMA…
  - assess fit of dose-response model by comparing to split NMA results
- Assumes that dose-response relationship is correctly specified
  - Here used **Emax function** but could use others



23

Relative effect at any dose



Dose x_2 of Agent a_2 compared to Dose x_1 of Agent a_1

$$\delta = f(x_2, a_2) - f(x_1, a_1)$$

24

Implementation

- Compare:
  - Lumped NMA: all doses assumed to have same effect (ignore dose)
  - Split NMA: no relationship assumed across doses of the same agent
  - Different model-based NMA
    - E.g. linear, Emax, others
- Assess goodness of fit (e.g. using DIC or similar), residual deviance and heterogeneity
- Check consistency

25

MBNMAdose

https://cran.r-project.org/package=MBNMAdose

- Fits Bayesian dose-response MBNMA
- Models different dose-response functions
- Can check consistency in data
- Allows for
  - class effects (sharing of parameters within a class)
  - Including study-level covariates (meta-regression)
- Produces
  - summary tables, treatment ranks and plots of key parameters
  - Outputs for comparing results from two models e.g. MBNMA vs NMA
  - Outcome predictions at different doses
- Includes several example networks, including the triptans network
- **Models also relevant for non-drug interventions**
  - E.g. physical activity: dose is exercise intensity

26

Optimal dose and type of physical activity to improve functional capacity and minimise adverse events in acutely hospitalised older adults (Gallardo-Gómez et al, 2023)

- Dose response not monotonic
- User-specified function

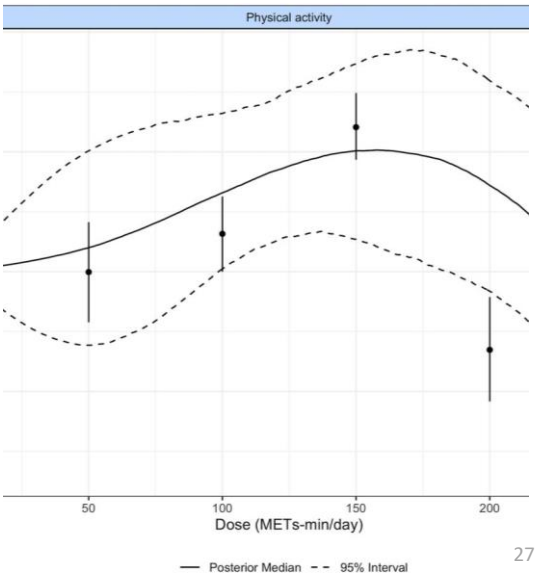


Figure 3 Dose-response relationship between physical activity dosage and functional capacity.

- Point estimates and credible intervals from a 'split' network meta-analysis in which each dose of physical activity is treated as an independent intervention.

British Journal of Sports Medicine

http://dx.doi.org/10.1136/bjsports-2022-106409

27

Comments

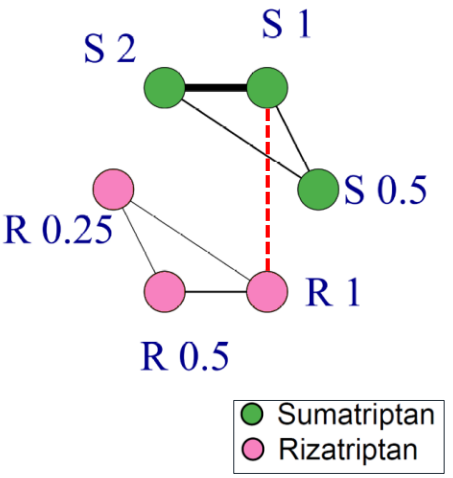
- Sharing of information via dose-response relationship can:
  - improve precision
- Uses relevant evidence on the interventions of interest
  - If there is a dose-response relationship, then evidence on an agent at one dose provides evidence that is relevant for other doses
- Availability of evidence at different doses is key
  - Phase II and non-licensed dose studies should be included in systematic review
  - Will increase burden of data extraction, but can strengthen inferences
- It may be possible to share dose-response parameters from different populations based on understanding of pharmacometrics
  - E.g. adults to children
- Can be useful to link **disconnected** networks of evidence

28

Joining the dots

CONNECTING NETWORKS

- Comparisons between disconnected treatments not possible without making strong assumptions.
- Dose-response MBNMA can be used to connect networks using the assumed dose-response relationship
  - Estimates functional relationships for dose-response models (eg Emax model)
- Allows interpolation to predict outcomes for doses not in the original trials

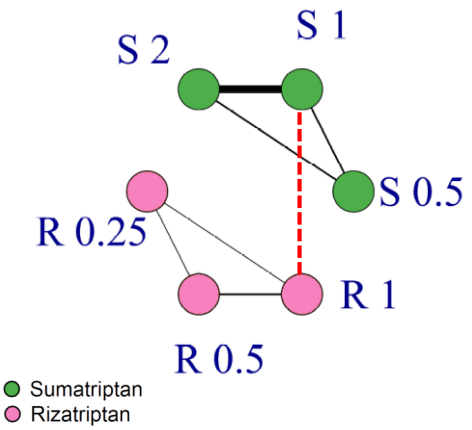


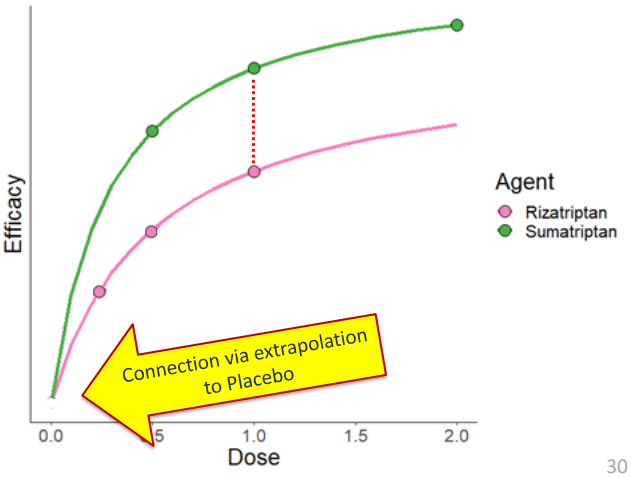
Pedder et al. (2021) https://doi.org/10.1177/0272989X20983315

29

Joining the dots via Placebo

- Interest in comparing licensed doses of Sumatriptan and Rizatriptan: S1 vs R1
- Disconnected network





30

Does it work?

- It can work well when the dose-response function is well estimated in the disconnected network components
  - Needs doses close enough to Placebo, along the curvature and towards the asymptote of the Emax function
- Typically unable to check assumptions since no data
  - See Pedder et al. (2021) for an example where a large network was artificially disconnected https://doi.org/10.1177/0272989X20983315
- Ongoing work to explore different scenarios.

31

Class models

- Sometimes there are many treatment options, but treatments fall into classes
- Treatments in the same class are assumed to have similar (but not identical!) effects
  - Eg if one SSRI works for depression, the others are likely to work too, to a similar extent
- A class model borrows strength across treatments in the same class
  - Effects of treatments in a class are distributed around a common class mean with a within-class variance
  - Treatment effects are shrunk to class mean (like a random effects meta-analysis)
- Estimates are more precise
- The original treatment definitions are retained
- Can help **connect networks** when disconnected treatment(s) are within a class

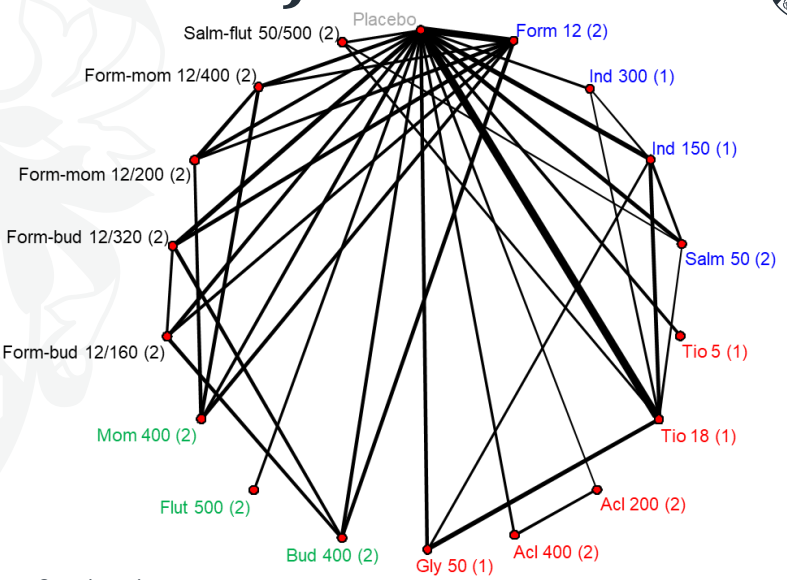
32

Example: COPD

- Treatments for chronic obstructive pulmonary disease (COPD)
  - Cochrane Review: Kew at al 2014
    http://dx.doi.org/10.1002/14651858.CD010844.pub2
- Outcome: mean difference on St George's Respiratory Questionnaire (SGRQ) total score at 6 months
- 25 trials, 18 treatments compared
  - 39 pairwise comparisons made
- Treatments belong to 5 classes: Placebo, Long-active β_2-agonists (LABA), Long-acting muscarinic antagonists (LAMA), Inhaled corticosteroids (ICS) and LABA+ICS.

33

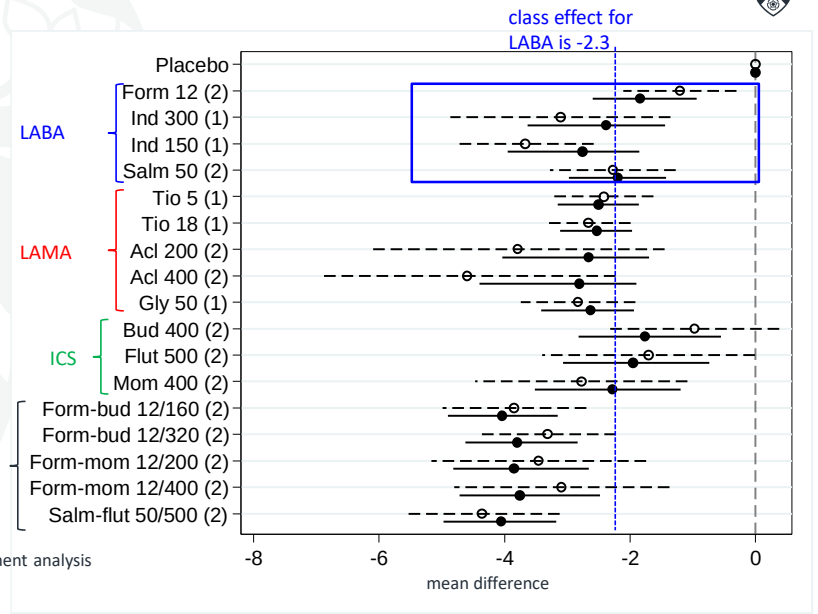
SGRQ (6 months)



SGRQ: St George's Respiratory Questionnaire

34

NMA with and without class effects



Dashed lines: individual treatment analysis
Solid lines: class analysis

35

Discussion

ADVANTAGES

- Models with functional relationships or borrowing of information across studies (e.g. by class) rely on correct specification of the dose-response function or borrowing conditions
  - Expert knowledge is required to assess suitability
- Estimates more precise and allow for better decisions
  - Subject to model assumptions
- Need to have *a priori* clinical plausibility as usually very few data/class elements to check assumptions
  - Especially useful when data sparse or when certain combinations are missing
- But needs to be convincing
  - Expert opinion and biological plausibility will be crucial
- Can be used to connect networks
  - Assumptions may be more plausible than other methods to connect networks

Thom et al. (2022) https://doi.org/10.1177/0272989X221097081

36

Discussion

DISADVANTAGES

- Requires sufficient data
  - Doses per treatment
  - Treatments per class
- May not work when evidence is too sparse
  - When few doses available, dose-response function parameters estimates will be too uncertain
  - Requires sufficient "spread" of doses across dose-response function
- Can use additional information on dose-response function from early phase dose-finding studies
  - To specify functional form?
  - To inform prior distributions on some parameters?
- Class assumptions can be combined with dose-response modelling
  - Class assumptions on treatment effects or dose-response parameters
- Additional data searching and extraction burden for all doses of relevant drugs:
  When is it worth it?

37

Thank you

@sdias_stats
sofia.dias@york.ac.uk