University of York

Centre for Reviews and Dissemination

Borrowing strength across evidence sources in Health Technology Assessment: challenges and opportunities

Sofia Dias

Bayes 2026, Leiden 23 October 2026



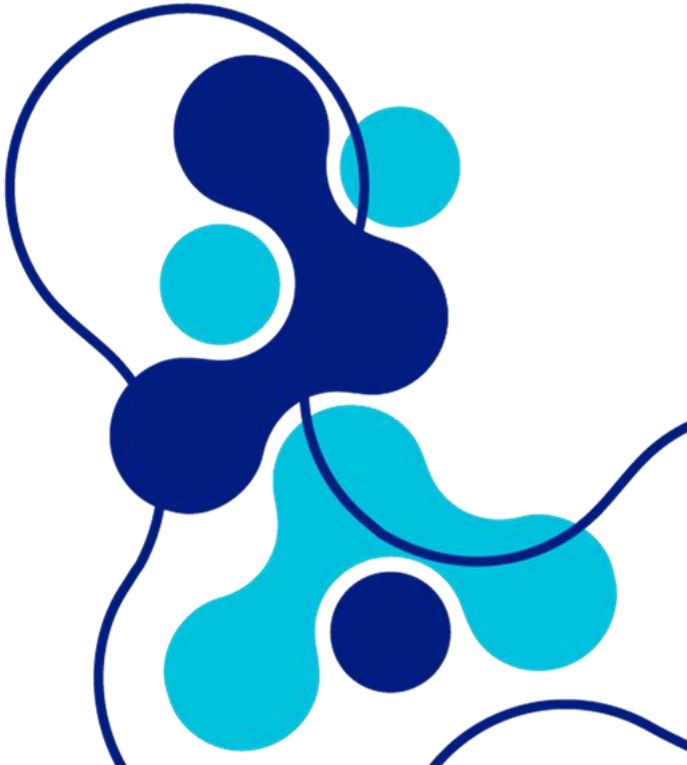
UNIVERSITY of York

Centre for Reviews and Dissemination

Disclaimer

I have no actual or potential conflicts of interest in relation to this presentation.

I acknowledge travel funding from the Conference organisers.



Health Technology Assessment (HTA)

University of York
Centre for Reviews and Dissemination

- After marketing authorisation granted, an intervention can be sold privately
  - It is considered to be safe and effective
- However, this is not sufficient to ensure funding through national health services or public insurance systems
- HTA evaluates whether the technology is value for money compared to existing technologies (i.e. standard of care)
  - Carried out by national or regional HTA bodies (such as NICE in the UK) or joint EU frameworks.
  - Analyse clinical data and real world performance
  - In the UK NICE considers economic models and evidence dossiers along with a critique by an external (independent) assessment group are presented to a committee who determine cost-effectiveness of new technologies and may recommend for funding
- A positive HTA recommendation leads to public reimbursement, inclusion in national formularies, and price negotiations
  - Patient access

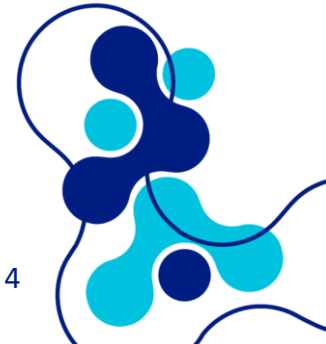
NICE: National Institute for Health and Care Excellence

Evidence sparseness

University of York
Centre for Reviews and Dissemination

- Evidence requirements for HTA are typically broader than for marketing authorisation
  - Need to **quantify** magnitude of the effect compared to the current standard of care, typically one or more active treatments, instead of placebo.
- Direct available evidence is often insufficient to inform the relative effects of the new technology versus all the existing comparators
  - RCTs typically against placebo or an older technology
  - Multiple comparators
  - Single-arm trials carried do not allow quantification of the relative effect
  - Available evidence may be biased due to internal validity or applicability/generalisability concerns
  - Even if direct, relevant evidence is available, estimates may be imprecise
- Can we use other relevant evidence to improve estimation of key parameters whilst limiting heterogeneity?

RCT: Randomised controlled trial



The Bayesian approach

University of York
Centre for Reviews and Dissemination

- Where HTA already considers a decision modelling framework, advantages of Bayesian analysis are easy to see
  - Correct propagation of uncertainty and parameter correlation
  - Direct simulation from posterior distributions (MCMC)
- Acknowledgement that all **relevant** evidence should be used
  - Exchangeability/hierarchical models, informative prior distributions etc allow adequate propagation of heterogeneity and uncertainty
- Are they accepted by decision-makers?
  - It depends…



Opportunities

University of York
Centre for Reviews and Dissemination

- Decision-makers recognise the need to borrow information to strengthen inferences
  - In cases where direct evidence is sparse discarding other relevant evidence is not appropriate
- Methods developments allow different assumptions to be considered
  - Model flexibility allows multiple explorations and sensitivity analyses
  - More acceptable assumptions
- Bayesian framework allows natural checking for conflict between different data sources
- Assumptions can be transparently communicated
  - Can validate and check acceptability with all stakeholders (e.g. clinical experts)



Challenges

University of York
Centre for Reviews and Dissemination

- Which evidence to borrow from?
  - Comprehensive literature review, existing external evidence, others?
  - Borrow evidence across categories (internally) e.g. basket trials, multi-cancer tests
- Models need to account for heterogeneity
  - but presence of heterogeneity can limit borrowing
- How to justify borrowing in the presence of heterogeneity?
  - Explaining heterogeneity can be difficult
- Using informative prior distributions
  - How are they derived?
  - Will they be accepted by decision-makers?



In practice…

University of York
Centre for Reviews and Dissemination

- **Aim:** to get best possible estimates of relative treatment effects to inform decisions
  - Unbiased, precise, robust
- Use “all relevant evidence” which may include
  - Different doses (but only licensed dose can be recommended)
  - Other (related) outcomes
  - Other (related) populations (e.g. children vs adults)
  - Different study types (high and low risk of bias, observational and RCTs etc)



Examples



Additional treatments

UNIVERSITY of York
Centre for Reviews and Dissemination

- Inclusion of **additional treatments** may provide relevant indirect evidence on the comparisons of interest
- Relative treatment effect quantification for all comparisons is essential for decision-making
  - Key input into decision model

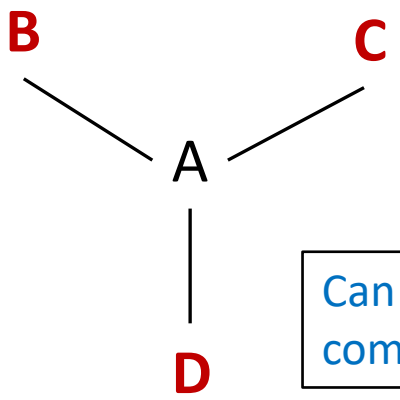


Additional treatments: network meta-analysis

Centre for Reviews and Dissemination

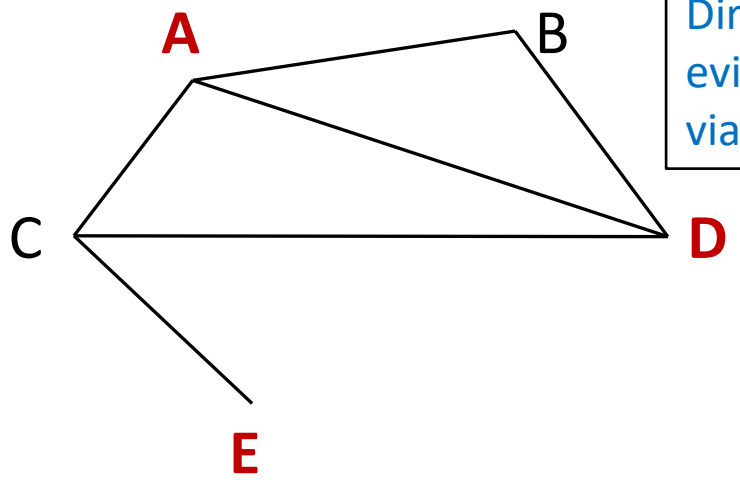
- Lack of RCTs directly comparing a new technology to the current standard(s) of care

- Example 1. interested in comparing treatments B, C and D, but not directly compared: Placebo (A) also included



Can form indirect comparisons via A

- Example 2. interested in comparing treatments A, D and E but include additional treatments B and C to increase precision

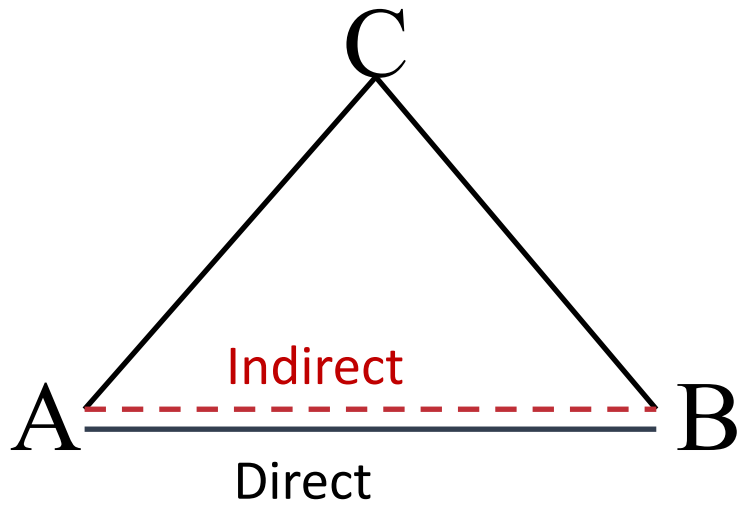


Direct and indirect evidence available via multiple paths

Network meta-analysis and indirect treatment comparisons

UNIVERSITY of York
Centre for Reviews and Dissemination

- Combines all evidence from RCTs that form a connected network
  - May include loops (both direct and indirect evidence)
  - Pairwise meta-analysis is a special case
- Indirect comparisons and NMA have gained popularity in HTA over the last ~20 years with lots of methodological developments
- Now commonly used in HTA
  - Reimbursement, clinical guidelines
  - Clinical research publications



Ades et al (2024) **Twenty years of network meta-analysis: Continuing controversies and recent developments** https://doi.org/10.1002/jrsm.1700

Network meta-analysis and indirect treatment comparisons

UNIVERSITY of York
Centre for Reviews and Dissemination

- Lack of RCTs directly comparing a new technology to the current standard(s) of care
  - Relative treatment effect quantification for all comparisons is essential for decision-making
- “Borrow” information from indirect evidence, under assumption of exchangeability of all relative treatment effects across all studies (consistency)
- First implementations of the models were within a Bayesian framework, directly motivated by the need to inform economic models
  - Methods to check for **conflict** between direct and indirect sources of evidence
- Many frequentist approaches now also exist but in HTA Bayesian models remain popular.



Bayesian information sharing models

UNIVERSITY of York
Centre for Reviews and Dissemination

- Can be used to synthesise evidence across different study designs, populations tumour types and/or stages of disease.
- Do not need to assume that evidence from different sources is estimating the exact same parameter
- Can impose different levels of sharing information based on different assumptions about the relationships across the evidence sources.
- Structural and exchangeability-based relationships can be used and combined with prior information
- However, unexplained heterogeneity can limit precision gains

Nikolaidis *et al.* (2021) Classifying information-sharing methods. BMC Med. Res. Methodol.;21:107.
Hemming *et al.* (2012) Pooling systematic reviews of systematic reviews: a Bayesian panoramic meta-analysis. Stat. Med. 2012;31:201-16.

Additional studies: risk of bias

UNIVERSITY of York
Centre for Reviews and Dissemination

- RCTs are considered “high quality” evidence
- Studies can be considered at risk of bias
  - Trial conduct issues, small study bias, sponsorship bias, novel agent (optimism) bias, missing data
  - Including studies at high risk of bias will increase heterogeneity and bias the estimated pooled treatment effect
- Should studies at risk of bias be excluded? Are they not relevant?
  - Can use a synthesis model that adjusts and downweighs the studies at risk of bias appropriately
  - More reliable and more precise estimates



Bias adjustment in (N)MA

UNIVERSITY of York
Centre for Reviews and Dissemination

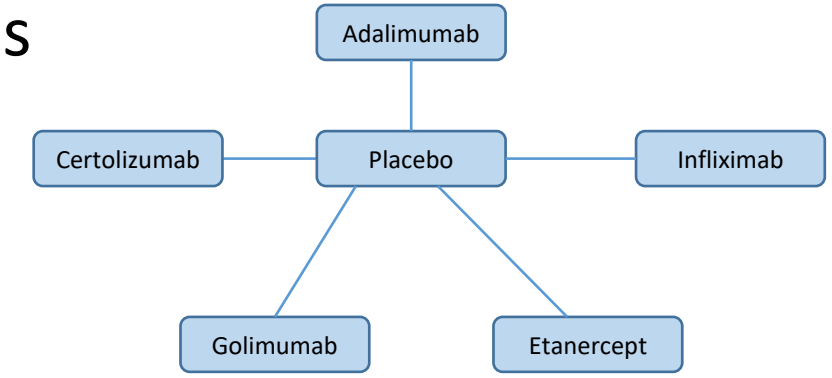
- Estimate bias associated with particular trial characteristics and adjust for it
- True treatment effect
  - Unbiased studies → δ
  - Assume $\beta \sim \text{Normal}(b,\kappa)$

  Biased studies → $\delta + \beta$
- Where does information on β come from?
  - External sources provide prior distributions for the bias parameters
  - formal elicitation from experts (Turner et al JRSS A, 2009)
  - meta-epidemiological studies (Savovic et al AJE, 2018; Reis-Pardal et al, 2026)
- Relative effect estimates can be adjusted for bias

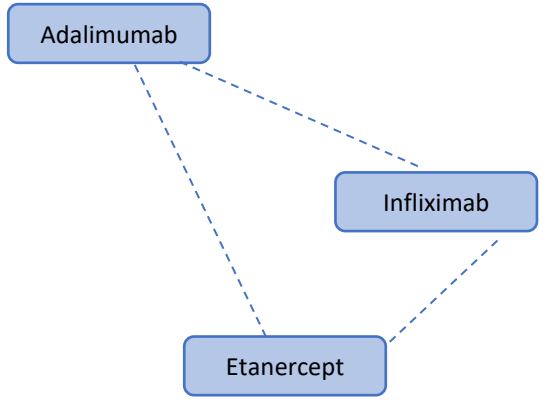
Additional studies: non-RCTs

- NMAs informed mainly by RCTs but data on some comparisons may be sparse
  - observational studies can be used to strengthen evidence in such cases
- Need to balance the desired increase in precision with the potential for introducing bias

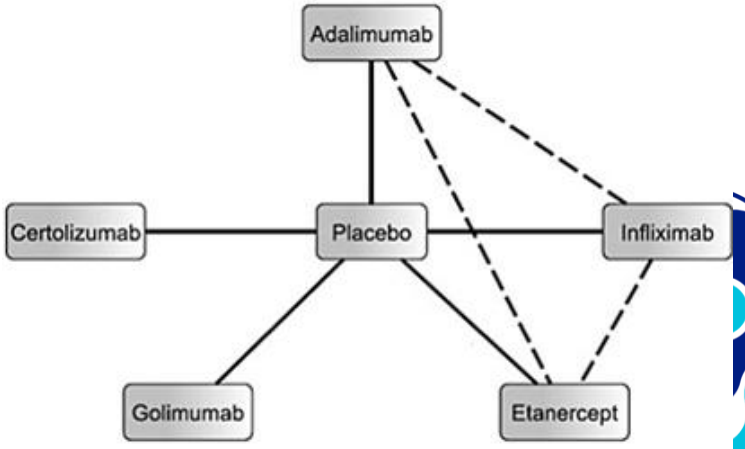
RCT



Observational



RCT + Observational



Network adapted from: Schmitz et al. (2013) https://doi.org/10.1002/sim.5764

17

Borrowing information: Models

University of York
Centre for Reviews and Dissemination

- Separate synthesis model
  - RCT and observational evidence are synthesised separately (no information sharing).
- Shared Parameter Models
  - Can share heterogeneity or treatment effect parameters or both (the lumped model)
- Multi-level/ Hierarchical Models
  - assume exchangeability between parameters informed by RCT and observational data
- Prior-based Models
  - Various types of model to be explored
- Bias Adjustement Models
  - Assume observational data are subject to an additional bias not in the RCTs
  - Can **adjust and downweigh** evidence

Downweigh evidence but do not adjust for bias

Hussein et al (2023) **Hierarchical NMA models for synthesis of evidence from randomised and non-randomised studies** https://doi.org/10.1186/s12874-023-01925-5)

18

Additional studies: populations

University of York
Centre for Reviews and Dissemination

- Directly relevant evidence from RCTs carried out in the target patient population (narrowly defined)
- But if evidence is sparse may include external evidence on
  - Related populations, e.g. use evidence in adults to inform MA in children (Walker et al 2023)
  - Indications, e.g. borrow information from relative effects of the same treatments in a different indication (Singh et al 2024)
  - If studies in breast cancer present more mature and precise evidence, can we borrow information for lung cancer decisions?
  - “Panoramic meta-analysis” (Hemming et al 2011)



Borrowing information across doses

UNIVERSITY of York
Centre for Reviews and Dissemination

- 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

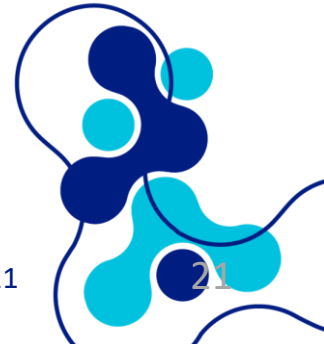


20

Dose-response NMA

UNIVERSITY of York
Centre for Reviews and Dissemination

- 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
  - Assess consistency assumption
- Uses additional evidence from studies (or arms) of doses not of primary interest
  - borrow strength through consistency/exchangeability relationship ***and*** dose-response relationship
- Allows estimation and prediction of treatment effects at multiple doses



Mawdsley et al (2016) **Model-Based Network Meta-Analysis: A Framework for Evidence Synthesis of Clinical Trial Data** https://doi.org/10.1002/psp4.12091

21

Multi-cancer tests

- Multi-cancer detection tests can detect circulating tumour DNA and identify the presence of up to 50 different cancers from a single blood draw
- The Galleri (GRAiL) test currently has the most developed evidence base, including a recently published RCT (NHS-Galleri)
- Proposed value of the test is to diagnose cancers early, before symptoms develop
- An economic model is being developed to provide an initial estimate of the cost-effectiveness of a screening programme using the Galleri test in the UK NHS
  - Evidence needs to be robust
- Accuracy of the Galleri test to detect each different cancer type at different stages is a key parameter for economic evaluation
  - Different treatment options with different effectiveness and costs

Galleri test: Data features

University of York
Centre for Reviews and Dissemination

- Variation in test sensitivity across cancer types and stages due to known differences in ctDNA expression
  - Different cancer types will express different levels of ctDNA which is what the Galleri test detects
- For a large proportion of the cancer types, the numbers of observed patients are small
  - significant uncertainty in test sensitivity at each stage (and overall)
- Where the test is expected to have similar sensitivity between two or more tumour types, or across different stages, data on one cancer type at a particular stage, can be used to **share information and strengthen inferences on other cancer types/stages**.



Sharing models

- Models allow different levels of sharing of information between cancer types
  - Exchangeability models
  - Mixture models
  - Class/Cancer group models
- Found substantial heterogeneity of test sensitivity across cancer types and stages, particularly for early-stage cancers
  - Suggests there are some cancer types that the Galleri test is more likely to find early than others.
- Models share information only weakly but can provide some increases in precision
- Additional sharing assumptions implemented using flexible sharing models were supported by the data
- Further understanding of the heterogeneity would allow models with stronger sharing assumptions to be implemented, and evidence to be used more efficiently.

- Note parallel with Bayesian models for borrowing information in basket trials

Preprint arXiv:2504.21517 – models currently being updated

Are these methods accepted by decision makers?

UNIVERSITY of York
Centre for Reviews and Dissemination

- Bayesian NMA and indirect comparisons regularly accepted at NICE technology appraisals and guidelines
- Bayesian information sharing models being considered for
  - Evaluation of histology-independent technologies at NICE (Larotrectinib)
  - Estimation of sensitivity of the Galleri test, to be considered by NHS-England and the National Screening Committee in their evaluation of the test
- Combining evidence from different populations (adults/children) when no other alternative has been used at NICE
- Combining observational and RCT evidence and bias-adjustment methods – recognition of need for methods to combine evidence
  - some use in practice, not very common.



Challenges and opportunities

University of York
Centre for Reviews and Dissemination

Opportunities

- Decision-makers recognise the need to borrow information to strengthen inferences
  - In cases where direct evidence is sparse discarding other relevant evidence is not appropriate
- Methods developments allow different assumptions to be considered
  - Model flexibility allows multiple explorations and sensitivity analyses
  - More acceptable assumptions
- Bayesian framework allows natural checking for conflict between different data sources
- Assumptions can be transparently communicated
  - Can validate and check acceptability with all stakeholders (e.g. clinical experts)

Challenges

- Which evidence to borrow from?
  - Comprehensive literature review, existing external evidence, others?
  - Borrow evidence across categories (internally) e.g. basket trials, multi-cancer tests
- Models need to account for heterogeneity
  - but presence of heterogeneity can limit borrowing
- How to justify borrowing in the presence of heterogeneity?
  - Explaining heterogeneity can be difficult
- Using informative prior distributions
  - How are they derived?
  - Will they be accepted by decision-makers?



UNIVERSITY of York

Centre for Reviews and Dissemination

THANK YOU!

sofia.dias@york.ac.uk

References

1. Ades et al. (2024) Twenty years of network meta-analysis: Continuing controversies and recent developments. Research Synthesis Methods. https://doi.org/10.1002/jrsm.1700

2. Hemming et al. (2012) Pooling systematic reviews of systematic reviews: a Bayesian panoramic meta-analysis. Statistics in Medicine, 31: 201–216. https://doi.org/10.1002/sim.4363

3. Hussein et al. (2023) Hierarchical NMA models for synthesis of evidence from randomised and non-randomised studies. BMC Medical Research Methodology, 23: 107. https://doi.org/10.1186/s12874-023-01925-5

4. Mawdsley et al. (2016) Model-Based Network Meta-Analysis: A Framework for Evidence Synthesis of Clinical Trial Data. CPT: Pharmacometrics & Systems Pharmacology, 5: 393–401. https://doi.org/10.1002/psp4.12091

5. Nikolaidis et al. (2021) Classifying information-sharing methods. BMC Medical Research Methodology, 21: 107. https://doi.org/10.1186/s12874-021-01297-y

6. Reis-Pardal et al. (2026) Meta-epidemiological studies informing bias priors in evidence synthesis.

7. Savović et al. (2018) Association of various potential sources of bias with treatment effect estimates in randomized clinical trials. American Journal of Epidemiology.

8. Schmitz et al. (2013) Network meta-analysis combining randomised controlled trials and observational studies. Statistics in Medicine, 32(23): 3943-3960. https://doi.org/10.1002/sim.5764

9. Singh et al. (2024) Borrowing information across indications in meta-analysis.

10. Turner et al. (2009) Bias modelling in evidence synthesis. Journal of the Royal Statistical Society: Series A (Statistics in Society).

11. Walker et al. (2023) Extrapolating evidence from adult to pediatric populations in evidence synthesis.

12. arXiv:2504.21517 – Models for multi-cancer detection test sensitivity evaluation.