Comparing treatments evaluated in studies forming disconnected networks of evidence: A review of methods

John W Stevens

Reader in Decision Science

University of Sheffield

EFPSI

European Statistical Meeting on Evidence Synthesis

2016

CENTRE FOR BAYESIAN STATISTICS IN HEALTH ECONOMICS

Acknowledgements

- Amgen:
  - For initiating the methodological review
- Chrissie Fletcher and Gerry Downey:
  - For supporting publication of the methodological review and as co-authors
- Anthea Sutton:
  - For conducting the systematic literature review of methods and as a co-author of the publication

Outline

Centre for Bayesian Statistics
in Health Economics

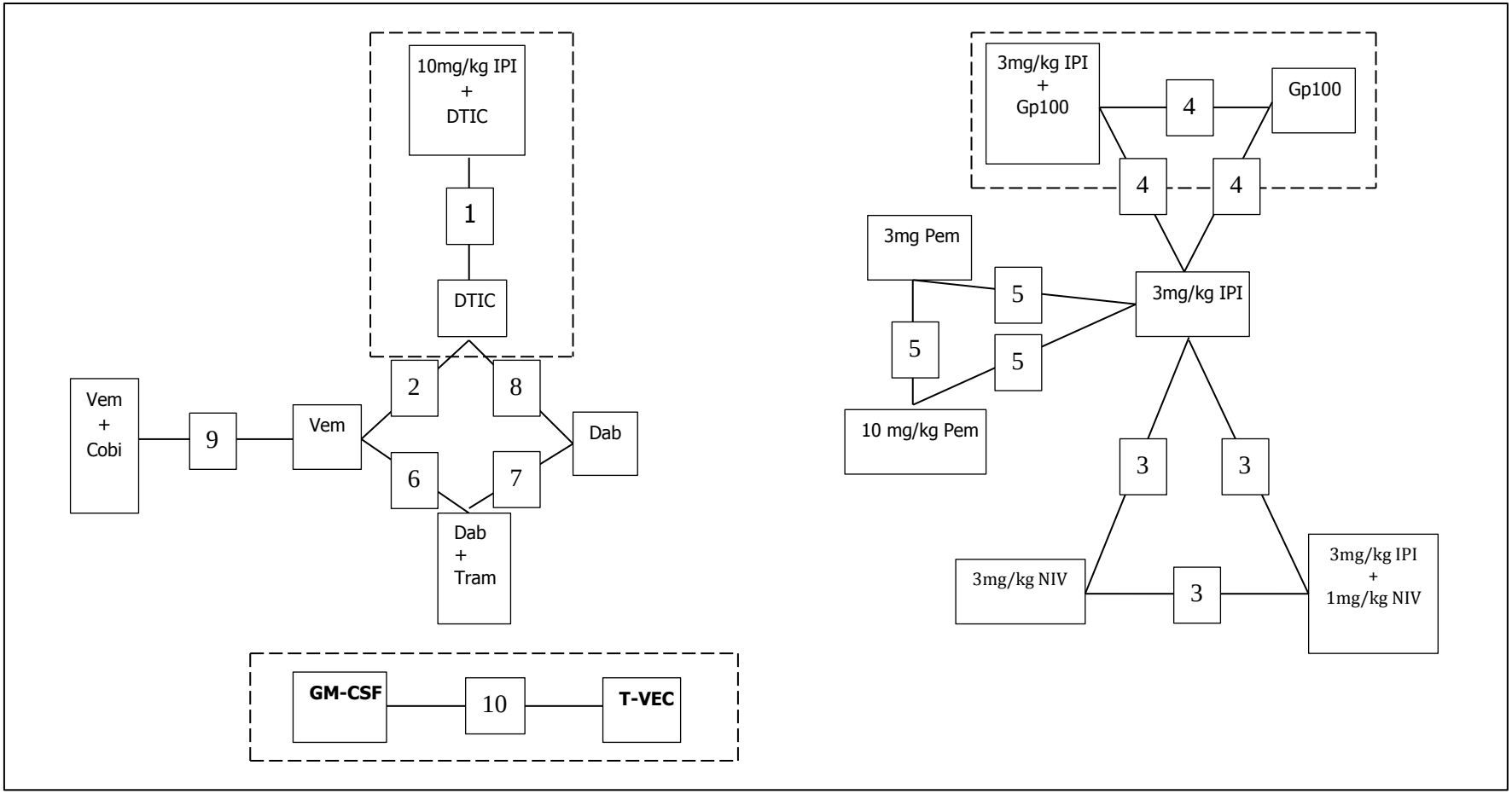
- Background
- The Problem
- Systematic Review of Methods
- Taxonomy of Methods
- Discussion and Recommendations

Background (1)

Centre for Bayesian Statistics in Health Economics

- NICE is responsible for making recommendations on the use of new treatments by the NHS in England
- Amgen was invited to submit evidence to support the use of T-VEC in metastatic melanoma
  - Comparators of interest were dacarbazine (DTIC), ipilimumab, vemurafenib and dabrafenib
- Amgen conducted a systematic literature review of published RCTs (and non-RCTs)

Background (2)



The Problem

- Perform a naïve or unadjusted indirect treatment comparison
  - Ignores differences in patient characteristics between studies and assumes that the data on each treatment arose from a single study
- Perform a conventional contrast-based network meta-analysis such that $d_{XY} = d_{ZY} - d_{ZX}$

  Not possible to compare treatments across networks without making additional assumptions

Systematic Review of Methods

Centre for Bayesian Statistics in Health Economics

- A two-stranded approach
  - Keyword searching
    - Including terms "no head-to-head" , "absence of head-to-head", "disconnected network", "meta-analysis"
    - Identified 23 articles
    - No new relevant articles were found
  - Pearl growing
    - Based on 11 published articles, including articles on model-based meta-analysis (which will not be discussed further)
    - Identified 343 articles; 258 relating to one article
    - 28 unique, relevant articles were found

Taxonomy of Methods

Centre for Bayesian Statistics in Health Economics

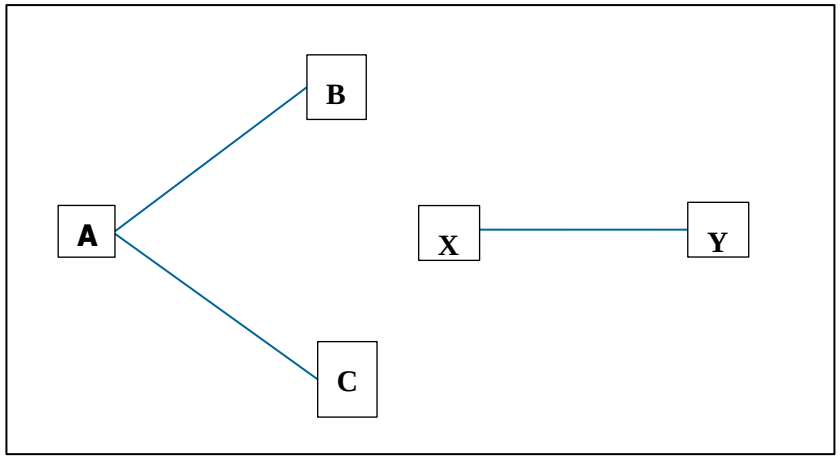
Simultaneous comparison between treatments in a heterogeneous population	**Use of external controls**
	Shared parameter model
	Random baseline model
Pair-wise comparisons in an homogeneous population	**Adjusted treatment response**
Add hoc methods	**Multivariate meta-analysis**
	Class effects

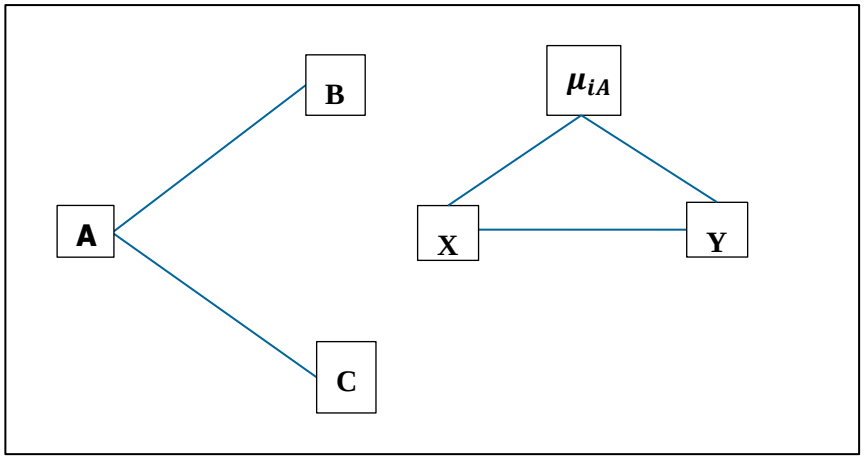
SIMULTANEOUS COMPARISON BETWEEN TREATMENTS IN A HETEROGENEOUS POPULATION

Use of External Controls (1)

Centre for Bayesian Statistics in Health Economics

- Formulate a prior distribution for a parameter (e.g. the log odds for a binary outcome) for the reference treatment in study i in at least one study in each group of disconnected studies





Use of External Controls (2)

Centre for Bayesian Statistics in Health Economics

- Korn et al (2008) proposed a method to create an external control as a benchmark in future single-arm studies in patients with metastatic Stage IV melanoma
  - Data from 2100 patients in 42 RCT and single-arm Phase 2 studies
  - External survivor function of an untreated group generated as:
    - $\bar{S}(t) = \frac{1}{n}\sum_{i=1}^{n} S_i(t)$ where $S_i(t) = [S_0(t)]^{HR}$

Use of External Controls (3)

Centre for Bayesian Statistics in Health Economics

- Limitations associated with the use of the Korn et al (2008) estimates:
  - Parameter estimates are sample statistics
  - Estimates of variances and covariances are not provided
  - It is unlikely that patient-level data will be available for comparator treatments
    - In non-linear models the expectation of a function is not the same as the function evaluated as its expectation i.e. $E_X[f(X)] \neq f(E[\bar{X}])$.
- More about the Korn et al (2008) model later

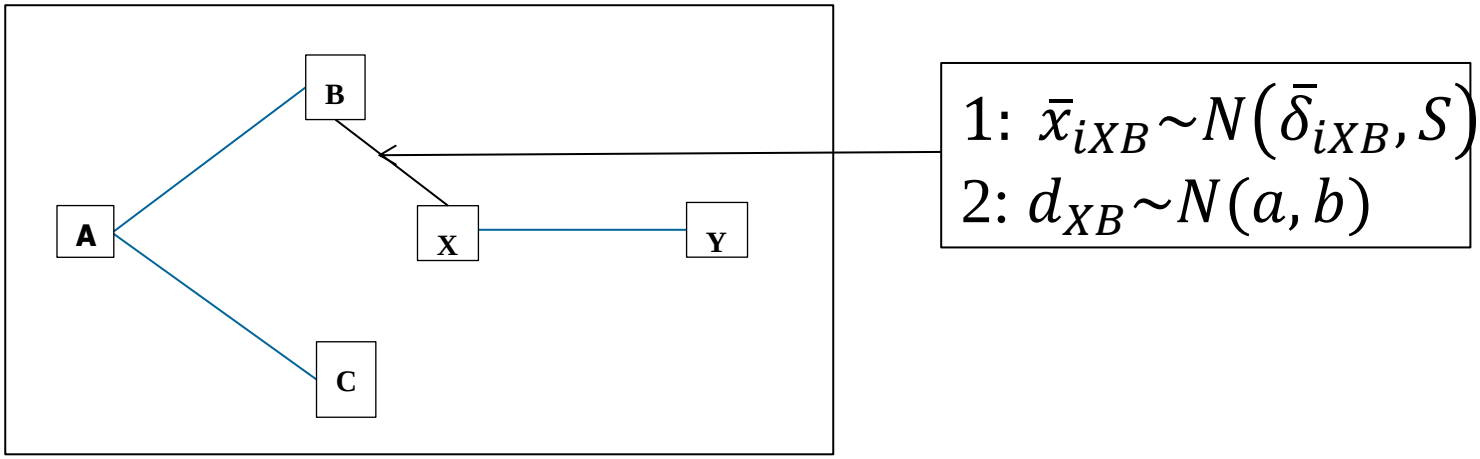
Use of External Controls (4)

Centre for Bayesian Statistics in Health Economics

- In the absence of any empirical evidence, use elicitation of experts' beliefs to formulate the required prior distributions.

Shared Parameter Model

- Abrams et al (2016) used observational data
- Alternatively, generate a prior distribution for the population effect of two treatments in different networks



Random Baseline Models

Centre for Bayesian Statistics in Health Economics

- Conventional meta-analyses combine relative treatment effects across studies
  - Baselines are treated as fixed within studies and unrelated across studies
- Random baseline models assume that the baseline are related across studies
  - A criticism of them is that they assume that patients are randomised across studies as well as within studies
- Thom et al (2015) used a random baseline model to connect disconnected networks

PAIRWISE COMPARISONS IN AN HOMOGENEOUS POPULATION

Adjusted Treatment Response

- Adjusted treatment response methods:
  - Generate adjusted responses for at least one treatment arm
  - Indirect estimates are derived as if the treatments had been included in the same study
- Inferences will generally differ from a random effects NMA depending on the patient population characterised by one of the studies
- We are aware of five methods that have been proposed

External Evidence-Based Adjustment

Centre for Bayesian Statistics in Health Economics

- Adjustments based on prediction models
  - Korn et al (2008) and modified Korn model
  - The adjustment factor, HR_{Adj}, for a comparator treatment is the hazard ratio for the new treatment, HR_N, divided by the hazard ratio for the comparator treatment, HR_C i.e. $HR_{Adj} = {HR_N}/{HR_C}$.
  - Adjusted survivor functions for the comparator treatment can then be generated as:

$$S_{Adj}(t) = S_C(t)^{HR_{Adj}}.$$

  - Assumes no unmeasured confounds and coefficients are independent and estimated without uncertainty

Propensity Score Matching Methods (1)

Centre for Bayesian Statistics in Health Economics

- Propensity score: the probability of treatment assignment conditional on observed covariates.
- Four ways in which a propensity score can be applied:
  - matching, with the most common approach being pair-matching
  - inverse probability of treatment weighting (IPTW)
  - Stratification
  - covariate adjustment

Propensity Score Matching Methods (2)

Centre for Bayesian Statistics in Health Economics

- Limitations
  - Estimates of treatment effect may be biased when there are unmeasured confounders
  - Model misspecification can also arise when ignoring interaction effects
  - Extreme weights can arise as the effect of covariates on treatment selection increases
  - Implementation requires access to patient-level data on the new and comparator treatments

Matching-Adjusted Indirect Comparisons (MAIC)

Centre for Bayesian Statistics in Health Economics

- MAIC
  - Uses IPD from a reference treatment in one study
  - Weights the data so that the average baseline characteristics matches those of a treatment in a different study
  - Approach similar to propensity score weighting
  - Limitations
    - Similar to propensity score matching
    - Inferences apply to the population defined by the comparator treatment
      - The target patient population can vary with each comparator

Simulated Treatment Comparisons (STCs)

Centre for Bayesian Statistics in Health Economics

- STCs are similar to MAICs
  - Use IPD from a reference treatment in one study
  - Uses a prediction model as a function of baseline characteristics
    - Adjusted responses based on the average baseline characteristics in the comparator study
  - Limitations
    - Ignores unobserved confounders
    - Introduces bias in non-linear models
    - Inferences apply to the population defined by the comparator treatment
      - The target patient population can vary with each comparator

Centre for Bayesian Statistics
in Health Economics

AD HOC METHODS

Multivariate Meta-Analysis

Centre for Bayesian Statistics in Health Economics

- Studies may form a connect network but individual outcomes may form disconnected networks
  - It might be possible to borrow strength across outcome measures using a multivariate NMA (MNMA)
  - A developing area of research that typically synthesises sample estimates of treatment effect using a multivariate normal likelihood function
  - We are not aware of any published work on MNMA of time-to-event outcomes in more flexible models that do not assume proportional hazards

Class Effects

Centre for Bayesian Statistics in Health Economics

- Treatments could be classified according to their drug class
  - Assumes there is no treatment effect within drug class variability
  - Might be useful when treatments are clinically equivalent
  - Pairwise studies comparing treatments in the same class are excluded
- This approach was used by Dequen et al., 2012

Discussion and Recommendations (1)

Centre for Bayesian Statistics in Health Economics

- Network meta-analysis (of RCTs)
  - Allows a synthesis of direct and indirect evidence
  - A simultaneous comparison of all treatments
- Disconnected networks
  - Indirect comparisons, even after adjustment, have been criticised as being a type of naïve indirect comparison
    - "its results are not worthy of consideration" Hoaglin, 2013
  - Statistical modelling is an important part of the armamentarium used to make inferences
  - Decision-makers must make a decision
  - Require alternative methods of analysis

Discussion and Recommendations (2)

Centre for Bayesian Statistics in Health Economics

- Methods can be classified according to whether:
  - they allow simultaneous comparisons between treatments in a heterogeneous population
  - pair-wise comparisons will be made between treatments in an homogeneous population
  - they are based ad hoc methods
- External controls and shared parameter models
  - Preserve the ability to make simultaneous comparisons between treatments
  - Prior distributions can be based on empirical evidence or expert opinion

Discussion and Recommendations (3)

Centre for Bayesian Statistics in Health Economics

- Adjusted treatment responses
  - MAIC and STCs may be useful in some contexts but may not be appropriate when the patient population in the comparator treatment's study is different to the target population
  - Proposals typically only account for sampling variation not parameter or structural uncertainty
  - Generating posterior distributions should be seen as an important aim in health technology assessment to represent uncertainty about inputs to decision analytic models

Discussion and Recommendations (4)

- All methods have limitations (some more than others) and there is a need for further research
  - to evaluate the robustness of results and assess the properties of frequentist methods
  - to generate examples using a Bayesian approach to reflect parameter uncertainty not just sampling variation
- Having made a decision, companies should be required to generate empirical evidence
  - Using value of information
  - To update evidence