PSI workshop conference

Cross-design approaches combining observational and clinical trial data for HTA

Mark Belger
Principal Research Scientist
Eli Lilly

Disclaimer

The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Eli Lilly.

Abstract

The Innovative Medicines Initiative (IMI) “GetReal” project explored methods for combing Randomised Clinical Trials (RCT) data with non-RCT data within the same Network Meta-Analysis (NMA). Methods such as, the design-adjusted analysis, using informative priors and three-level hierarchical models have been summarised in the manuscript. “Combining randomized and nonrandomized evidence in network meta-analysis “[Orestis Efthimiou et al.]. We will discuss how to incorporate these methods within an HTA setting. Outlining the limitations in combining this type of evidence, and exploring how these methods are used to improve our understanding of how a new intervention will perform outside of the clinical trial environment.

Efthimiou O, Mavridis D1, Debray TP, Samara M, Belger M, Siontis GC, Leucht S, Salanti G; GetReal Work Package 4. **Combining randomized and non-randomized evidence in network meta-analysis**. Stat Med. 2017 Apr 15;36(8):1210-1226. doi: 10.1002/sim.7223. Epub 2017 Jan 12

Presentation contents

- IMI “GetReal” and context
- Summary of methods available
- Recommendations from IMI
- Further research/ challenges

IMI “Get Real”

Statistics in Medicine

Research Article

Received: 13 June 2016, Accepted: 16 December 2016 Published online 12 January 2017 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.7223

Combining randomized and non-randomized evidence in network meta-analysis

Orestis Efthimiou,[a*†] Dimitris Mavridis,[a,b] Thomas P. A. Debray,[c,d] Myrto Samara,[e] Mark Belger,[f] George C. M. Siontis,[g] Stefan Leucht,[e] Georgia Salanti[a,h,i] and on behalf of GetReal Work Package 4

Non-randomized studies aim to reveal whether or not interventions are effective in real-life clinical practice, and there is a growing interest in including such evidence in the decision-making process. We evaluate existing methodologies and present new approaches to using non-randomized evidence in a network meta-analysis of randomized controlled trials (RCTs) when the aim is to assess relative treatment effects. We first discuss how to assess compatibility between the two types of evidence. We then present and compare an array of alternative methods that allow the inclusion of non-randomized studies in a network meta-analysis of RCTs: the naïve data synthesis, the design-adjusted synthesis, the use of non-randomized evidence as prior information and the use of three-level hierarchical models. We apply some of the methods in two previously published clinical examples comparing percutaneous interventions for the treatment of coronary in-stent restenosis and antipsychotics in patients with schizophrenia. We discuss in depth the advantages and limitations of each method, and we conclude that the inclusion of real-world evidence from non-randomized studies has the potential to corroborate findings from RCTs, increase precision and enhance the decision-making process.

Including non-randomized studies (NRSs) in a NMA

Background

- Evidence from NRS may **complement** evidence provided by RCTs, and potentially address some of their limitations
- NRS can be used to **improve connectivity** in the network of competing interventions by providing missing links between treatments, and make estimates more precise
- Interest in including NRS in the NMA synthesis and decision-making process is growing.

Step 1

Use the NRSs to obtain estimates on relative treatment effects

Several different approaches have been proposed in the literature. See Faria et. al for a recent review. These methods include regression adjustments, the use of propensity score function, matching techniques etc.

Step 2

Assessing the compatibility of the evidence

- ✓ Before pooling observational and randomized evidence in a NMA, one must first assess the extend of **compatibility** between the types of evidence.

- ✓ For each treatment comparison there may be up to **four different types of evidence**
  - ❑ Direct randomized
  - ❑ Indirect randomized
  - ❑ Direct observational
  - ❑ Indirect observational

The four sources of evidence are independent and they can be formally compared with statistical tests.

- Differences between direct randomized and indirect randomized: **inconsistency in the network of RCTs**
- Differences between direct non-randomized and indirect non-randomized: **inconsistency in the network of NRSs**
- Differences between randomized and non-randomized sources: might be due to residual confounding in the observational evidence, or important differences in characteristics of patients between RCTs-NRSs.

Important differences need to be explored. If a source of disagreement is identified researchers can perform analyses that account for it and improve comparability across the different sources of evidence

Step 3
Data synthesis

Four alternative generic approaches were explored

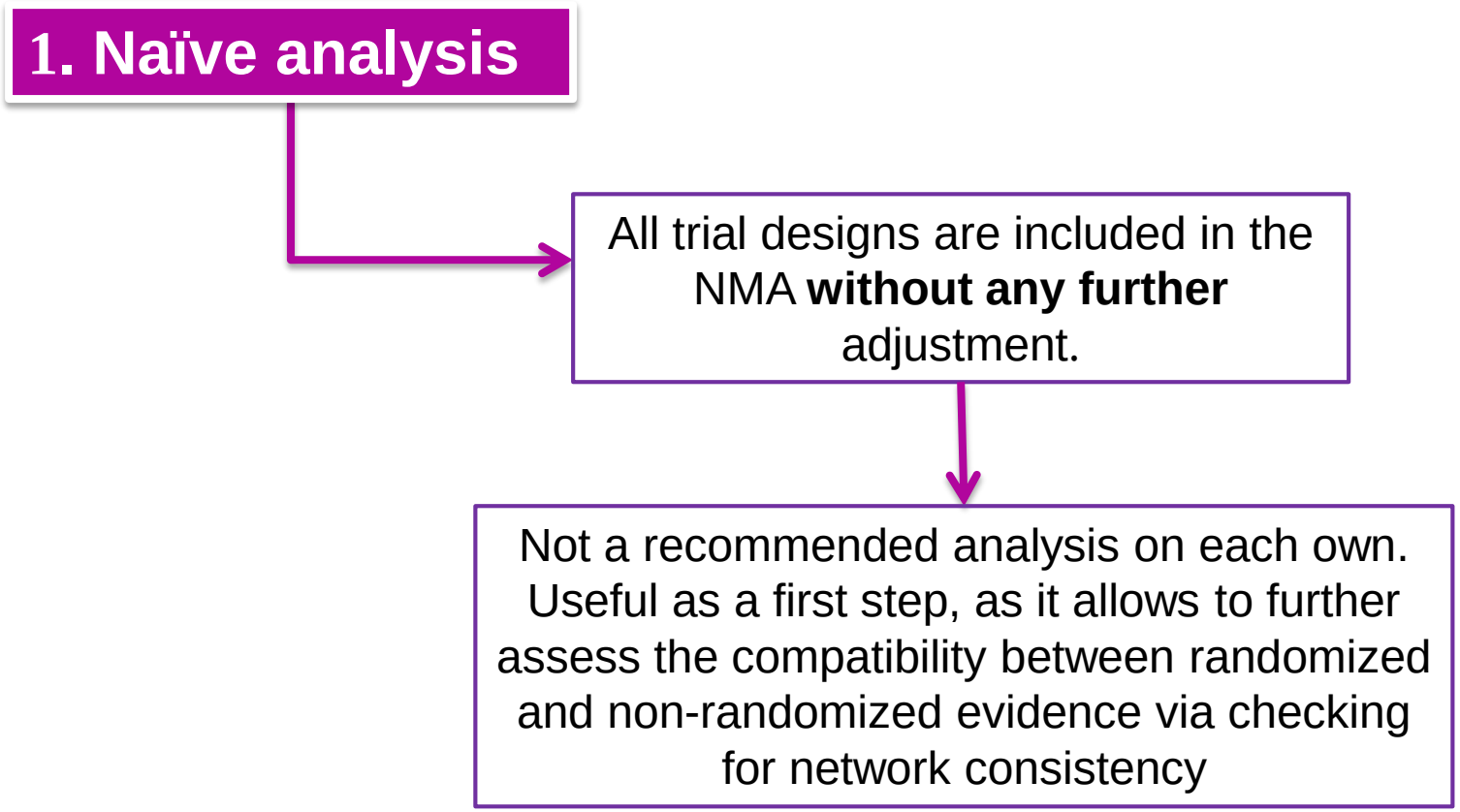
i. The naïve approach

ii. The design-adjusted analysis

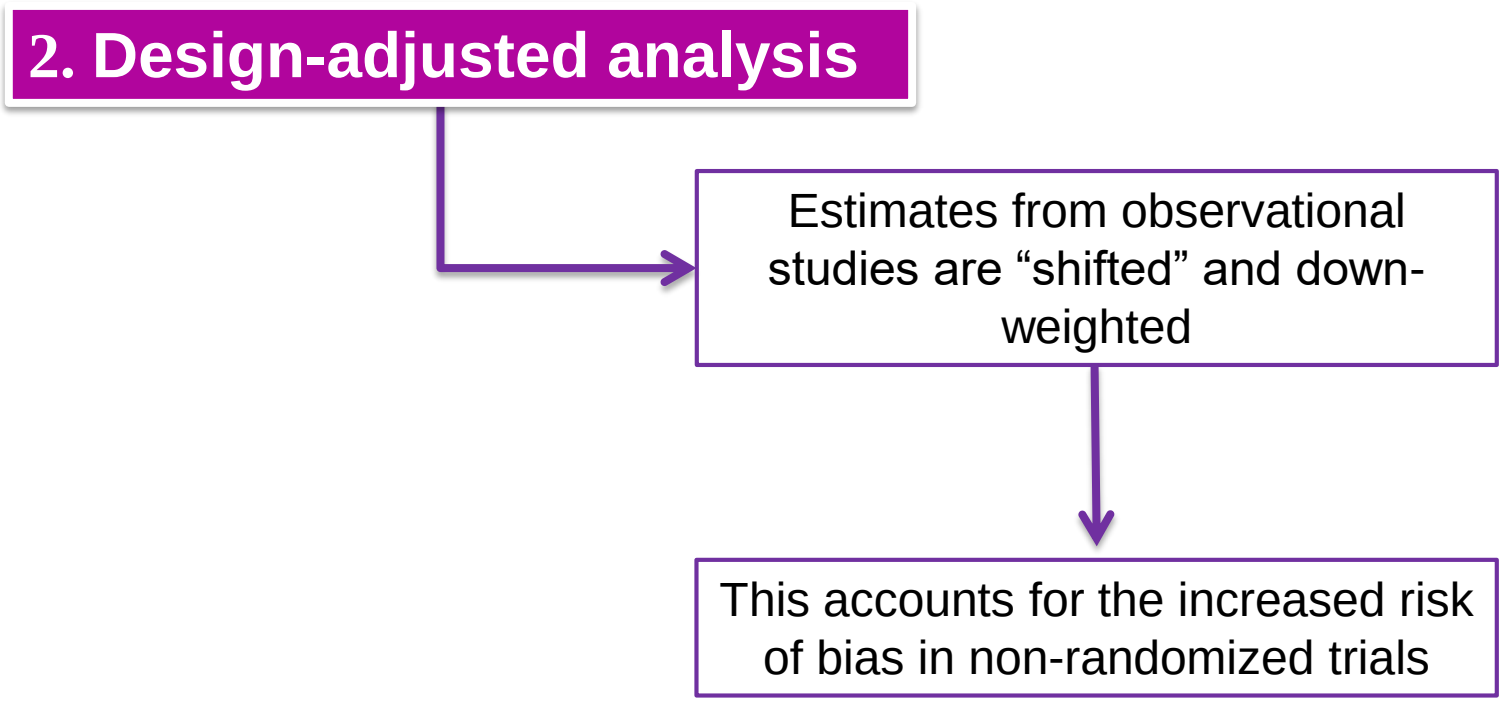
iii. Using informative priors

iv. Three-level hierarchical models

Statistical methods for combining RCTs and non-randomized studies



Statistical methods for combining RCTs and non-randomized studies



The usual random effects model

Consider a non-randomized study j comparing treatments X vs. Y. The usual random effects model assumes that:

$$d_{jXY} \sim N(\theta_{jXY}, s^2_{jXY})$$
$$\theta_{jXY} \sim N(\mu_{XY}, \tau^2_{XY})$$

The design-adjusted analysis

Consider a non-randomized study j comparing treatments X vs. Y. We assume that:

$$d_{jXY} \sim N(\theta_{jXY} + \beta_j, \frac{s^2_{jXY}}{w_j})$$
$$\theta_{jXY} \sim N(\mu_{XY}, \tau^2_{XY})$$

The point estimate is shifted by β_j

The variance is inflated by w_j $(0<w_j<1)$

These 'bias factors' need to be elicited using expert opinion, taking into account risk of bias and the credibility of each study's results.

Pinpointing exact values for β and w may be a difficult task → sensitivity analyses are necessary

By changing the value of w_j researchers can control the **amount of confidence** they want to place to the ^{jth} NRS and can easily perform sensitivity analysis.

- Setting $\boldsymbol{w_j = 1}$ corresponds to accepting the ^{jth} NRS at face value
- Setting $\boldsymbol{w_j = 0}$ corresponds to excluding the ^{jth} NRS from the analysis

Statistical methods for combining RCTs and non-randomized studies

3. Using non-randomized evidence as prior information

Set in a Bayesian framework: the non-randomized studies are analyzed separately, estimates used to formulate prior distributions for relative effects

Using non-randomized evidence as prior information

The analysis is performed in two stages.

- First, perform a (network) meta-analysis of the observational evidence
- The posterior distributions of the first step of the analysis are used as **prior distributions** for (some, or all of) the **basic parameters** of the NMA model, which includes only RCTs.
- Observational evidence can be ‘shifted’ and/or down-weighted

Most important differences between the **‘design-adjusted’** approach and using NRS as **prior information** :

- design-adjusted: adjust each study for bias **separately**, estimate τ using **all studies**
- As prior information approach: down-weights all NRS **collectively**; there is a **different** τ for RCTs and NRSs.

Statistical methods for combining RCTs and non-randomized studies

4. Three-level hierarchical models

First level: the study level.
Second level: the study-design level.
Third level: The overall level.

- The first level each study is analysed individually
- The second level synthesises studies of the same design
- The third level allows for design level heterogeneity

The three-level hierarchical models

More than two types of studies can be included in the analysis (e.g., different RCT designs, cohort studies, case-control etc.)

The basic assumption behind this model is that the underlying treatment effect is not fixed across designs

A design-level heterogeneity τ_{des} parameter enters the model, rendering estimates more conservative

When there are only few different designs in the data the estimation of τ_{des} will be difficult

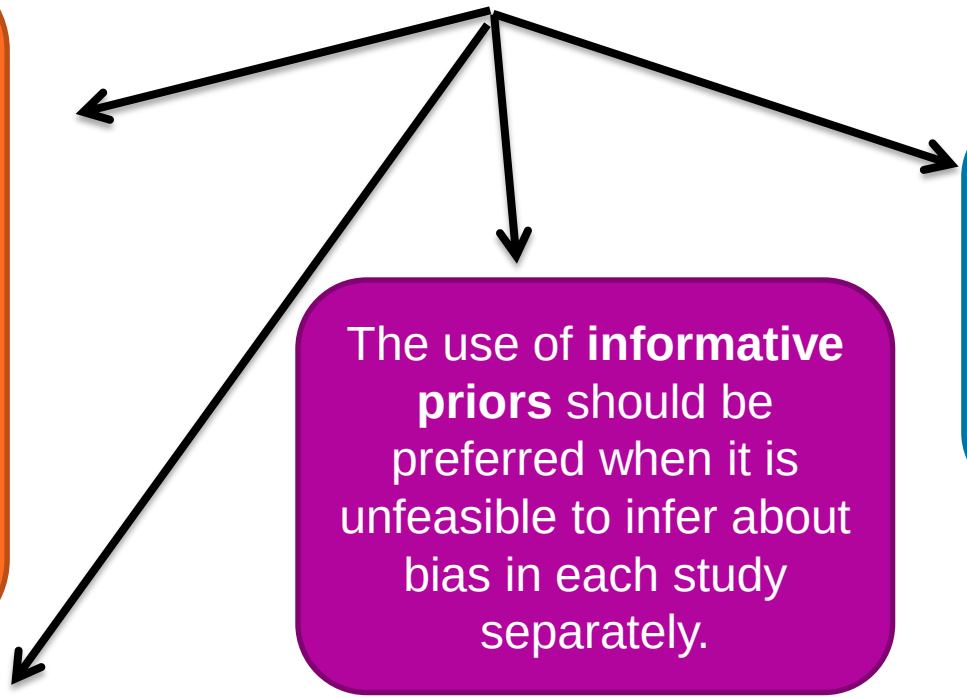
Step 4
Estimating the influence of the observational evidence in the NMA results

After the analysis has taken place, when assessing the **quality of evidence** provided by the NMA, researchers need to infer about the relative contribution of the various sources in the estimation of the pooled results, after all adjustments have taken place

- In a frequentist setting the relative contribution of each study to NMA estimates can be assessed by calculating the **hat matrix** (see Krahn et al.)
- In a Bayesian setting a measure similar to the multivariate I^2 statistic can be used (Jackson et al.)
- These approaches can quantify the **percentage contribution** of each study to the pooled results

Which method to use?

Naïve analysis is a good starting point and allows the use of considerations regarding network consistency to further assess the compatibility of the data



The use of **informative priors** should be preferred when it is unfeasible to infer about bias in each study separately.

Three-level hierarchical models should be used when there are studies pertaining to multiple designs

In the **design-adjusted** approach the estimates from each NRS are adjusted separately. Should be preferred when resources allow inference about bias in each study.

Why is it useful?

- **More direct and relevant results:** RCTs often have strict inclusion criteria, which may lead to the study populations differing from real-world populations. NRSs or other RWD sources may provide a more direct answer to the research question, and including RWE in the network may allow researchers to obtain more relevant answers.
- **Increased precision and power:** the inclusion of NRSs or other RWD sources can increase precision and power as compared with NMAs of RCTs.
- **Potentially corroborative to RCT evidence alone:** an NMA that includes both randomised and non-randomised evidence may corroborate conclusions drawn from an NMA of RCTs alone and reassure decision makers that study findings are transferable to real-world populations.

What are its limitations?

- **Risk of bias:** estimates of relative treatment effects obtained from NRSs or other RWD sources are considered to be at a higher risk of bias, due to the lack of randomisation and increased risk of biases in the data sources.
- **Difficulties in obtaining data:** obtaining individual participant data (IPD) from NRSs or other RWD sources might be difficult. Reported aggregated estimates on relative effects may be biased if non-optimal analysis methods have been used. Use of IPD from NRSs was examined in the GetReal case study – see here.
- **Reliability of results:** the inclusion of NRSs or other RWD sources in the network may make the underlying assumptions of the NMA model less plausible and the NMA results less reliable.
- **Increased effort to carry out:** including NRSs or other RWD sources in a NMA may greatly increase the workload of the review team.
- **Complex to carry out:** methods for including NRSs or other RWD sources in the NMA are complex and may be difficult to implement because they require additional software expertise.
- **Based on expert opinion:** most approaches for including NRSs or other RWD sources require expert input, which can be time-consuming. Possible biases in the estimates from these studies may be hard to predict, either in magnitude or in direction.

What do stakeholders say?

- There may be concerns about the plausibility of the underlying assumptions of an NMA or other RWD sources when NRSs are included.
- The inclusion of NRSs or other RWD sources is seen as a threat to the validity of NMA estimates, because of the increased risk of bias in the observational estimates.
- For some stakeholders the statistical methods might be difficult to understand or carry out.

Any Questions?

References

- Leucht S, Cipriani A, Spineli L, et al. (2013) Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 382:951–962. doi: 10.1016/S0140-6736(13)60733-3

- Efthimiou O, Debray TP et al. (2o16) GetReal in network meta-analysis: a review of the methodology. Research synthesis methods, doi: 10.1002/jrsm.1195.

- Faria R, Hernadez Alava M, Manca A, Wailoo AJ The use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. NICE DSU Tech. Support Doc. 17

- Krahn, U., Binder, H. & Konig, J. A graphical tool for locating inconsistency in network meta-analyses. BMC.Med.Res.Methodol. 13, 35 (2013).

- Jackson, D., White, I. R. & Riley, R. D. Quantifying the impact of between-study heterogeneity in multivariate meta-analyses. Stat. Med. 31, 3805–3820 (2012).