POPULATION-ADJUSTED TREATMENT COMPARISONS IN HEALTH TECHNOLOGY ASSESSMENT: AN OVERVIEW OF APPROACHES AND PERSPECTIVES

Providing an industry perspective

Presented by

Mark Belger, Principal Research Scientist, Eli Lilly

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.

Acknowledgements

♦ Alan Brnabic
♦ Zbigniew Kadziola
♦ Helmut Petto
♦ Daniel Saure

3

Outline

♦ 1. *Background*
♦ 2. *Challenges*
♦ 3. *Matching Adjusted Indirect Comparisons*
♦ *4. Limitations*
♦ *Next Steps*

Background

♦ The majority of HTA company submissions will include a section on treatments that are currently approved for the indication that the new intervention is being considered for. Comparisons will be made on efficacy, safety and cost effectiveness

♦ Where possible Head to Head (H2H) studies are used. In the absence of H2H studies then Indirect Comparisons and Network Meta-analysis are performed.

♦ These analysis require a number of assumptions to be met However there will always be some differences between studies in the Network, and methods for adjusting for these differences (random effects models, meta-regression) may not be sufficient to negate the concerns

♦ Methods such as Matching Adjusted Indirect Comparisons have been suggested to alleviate some of these concerns

The 3 assumptions

♦ **Heterogeneity**

Variation in the same relative treatment effect between studies

= Trials are homogenous

♦ **Similarity**

patients included should be sufficiently similar in the two sets of common comparator controlled trials

E.g. Placebo creep

♦ **Consistency**

No difference between direct and Indirect comparison

Ref: http://www.whatisseries.co.uk/what-is-indirect-comparison/
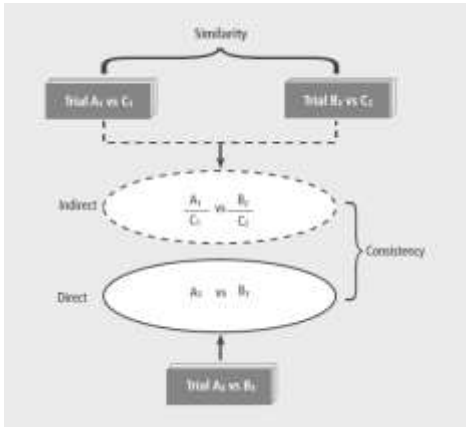
Transitivity assumption

♦ *"Transitivity implies that the treatment comparisons within the indirect comparison do not differ with respect to the distribution of <u>known treatment effect modifiers (EMs)</u>."*

♦ EM occurs when the magnitude of the effect of the primary exposure on an outcome (i.e., the association) differs depending on the level of a third variable.

♦ Effect modifiers are variables which influence the relative treatment effects

♦ EMs are the main reason for using matching-adjusted indirect comparisons (MAIC)

Matched Adjusted Indirect Comparison (MAIC)

♦ Basic idea: If you have individual patient data (IPD) then re-weight patients outcomes of IPD trial before conducting the adjusted indirect comparison (AIC). Weights are generated by matching to baseline characteristics (and effect modifiers) from published trials

How is it conducted? 3 Steps:
1. Clinical trial selection (SLR)
2. Identification of outcome measures (Should be aligned)
3. Matching trial populations
   Inclusion/exclusion, matching baseline/ EMs

Step 3 MAIC: Matching IPD to AGR
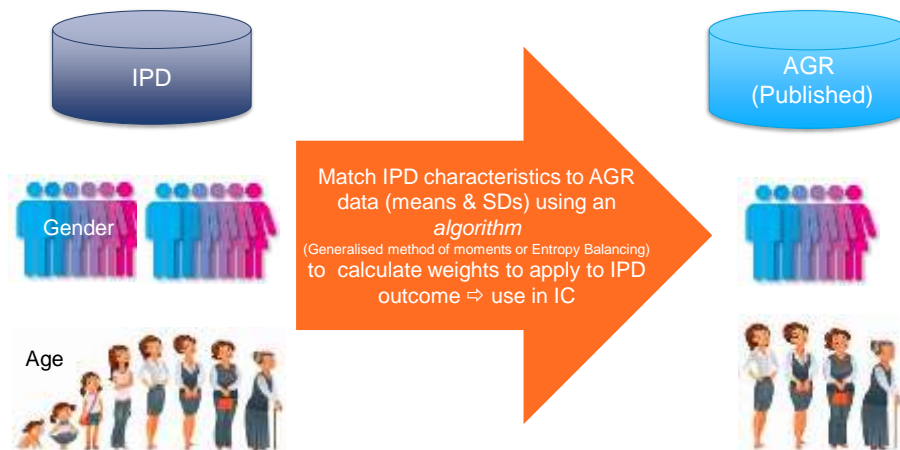
AGR = aggregate; IPD = individual patient data; MAIC = matching-adjusted indirect comparison

Ref: Signorovitch et al. *Pharmacoeconomics* 2010; 28:935–945

Limitations in performing MAIC

♦ Often IPD data is only available for the treatment being assessed, and the comparator data is aggregated, and so reliant on the how much information is published

♦ The current MAIC methods force you to perform the MAIC in the population of the published comparator study

♦ If comparing against older studies then study designs may have changed a lot over the years, and the common comparators may not be so "Common"

♦ Treatment effect modifiers may not be known and information on them is not available in the published comparator study

♦ Published methods currently only look at one common comparator, but you may have more than one common comparator in the studies available to you.

Methods in simulation project

Weighting methods

Different approaches to weights calculation: Weighting as originally proposed by Signorovitch (WS) vs. EB. WS is calculated by method of moments.[4] EB relies on a maximum entropy reweighting scheme[5]

Simulations

Simulated data sets of treatment A vs treatment C (AC) IPD trials and treatment B vs treatment C (BC) AGR trials are used to assess the performance of different weighting methods.

Each trial consisted of a sample size of 300; with M=500 replications were simulated with 1000 bootstraps within each replication for variance estimation. The simulations presented in this poster look at a continuous outcome variable. Two different types of baseline covariates were created:

- – Xs which were available for analysis in both the IPD and AGR studies
- – Zs which were only available for analysis in the IPD studies

Assessing Performance of Weighting methods

The performance of each weighting method was assessed through:

- – Estimated treatment (IC) effect and bootstrap 95% Confidence Interval
- – Bias, difference from "True treatment effect" $BIAS(\hat{\Delta}) = M^{-1} \sum_{m=1}^{M}(\widehat{\Delta_m} - \Delta_m)$
- – Mean Standard Error (MSE) $MSE(\hat{\Delta}) = M^{-1} \sum_{m=1}^{M}(\widehat{\Delta_m} - \Delta_m)^2$

Increasing the Number of Continuous Treatment Effect Modifiers

Introduce into the model further effect modifiers (do not include additional covariates which are not treatment modifiers)
Two treatment modifiers X1 and X2

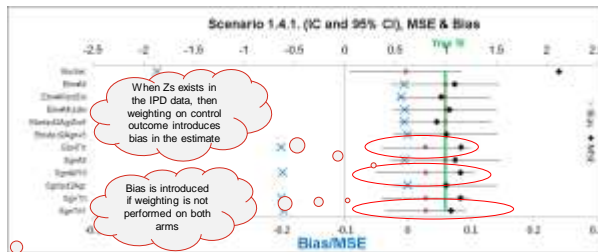Include three treatment modifiers X1, X2, and X3

Increasing the number of treatment effect modifiers does not appear to have an impact on the estimate or the variance

Does the size of the treatment modifiers effect the performance of the weighting method used? X3 is a larger treatment modifier than X1 and X2
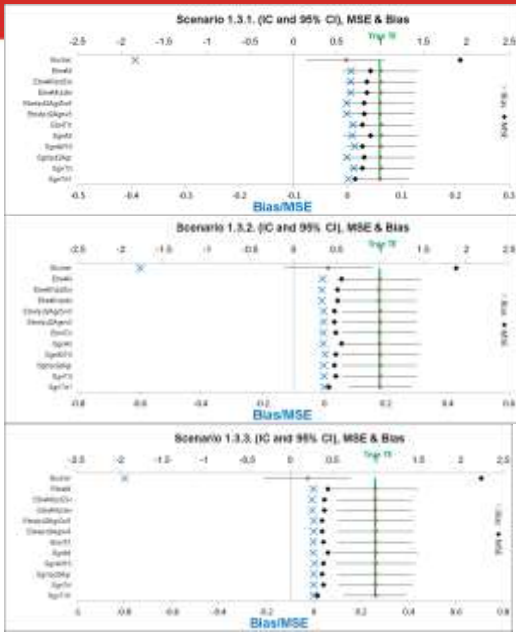
Influence of Unmeasured Continuous Effect Modifiers

What is the influence of unmeasured confounders, so factors known to influence treatment effect, but are not included in the matching? Just include one of these variables Z_1, but consider different size of effect modifier. (Keep model simple, just two Xs included) – compare against S.1.3.1
In this first simulation, Z_1 is not an effect modifier

When Zs exists in the IPD data, then weighting on control outcome introduces bias in the estimate

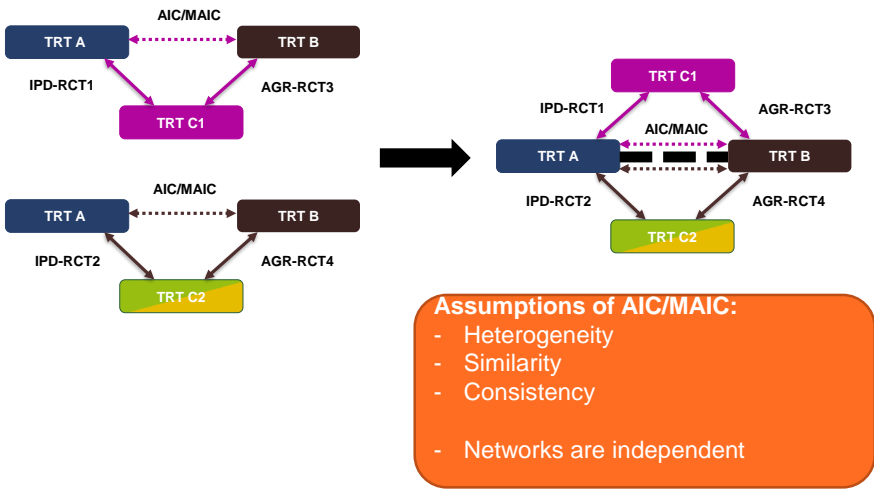Bias is introduced if weighting is not performed on both arms

In Scenario 1.4.1, the introduction of unmeasured confounder Z_1 introduces bias into the reweighted estimates and increases the width of the CI. The exception is when weighting is performed on both arms separately
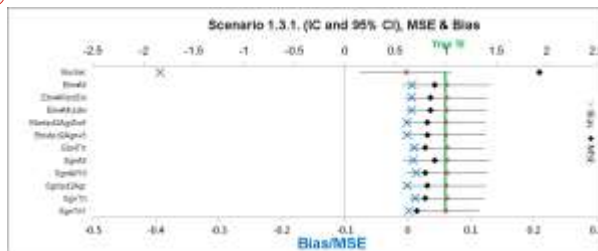
Recommendations from simulation project

- If information is available on all treatment effect modifiers in both the IPD and AGR data sets, then all weighting methods perform well

- The number of treatment effect modifiers included does not appear to effect any of the weighting methods to obtain a true estimate of the IC.

- As you introduce more variables that are not effect modifiers, it is important that weights are applied to individual arms (weighting on both arms separately, or weighting on just the treatment arm)

- Bias is introduced when weighting only on treatment arm and prognostic variables are present in the IPD and not in the AGR

- Bias is introduced when weighting on the outcome in the control arm of the AGR, and prognostic variables are present in the IPD and not in the AGR

- All weighting methods fail to estimate the true treatment effect when treatment effect modifiers are present in the IPD and are not reported in the AGR

Suggestion: Merging *via* meta-analysis

Assumptions of AIC/MAIC:
- Heterogeneity
- Similarity
- Consistency

- Networks are independent

Next Steps

- ♦ Look at ways of adjusting to a common population rather than to the population of the published comparator study

- ♦ Introduce methods to assess the impact of unmeasured confounders

- ♦ Extend beyond the Indirect Comparison to Network Meta-Analysis models

- ♦ If combining different models, develop test for consistency

- ♦ Consider ways of applying these to sub-groups of the IPD