Advanced methods for matching-adjusted indirect comparison

Antonio Remiro-Azócar, PhD
Methods & Outreach, Novo Nordisk
ISPOR 2024, Atlanta
8th May 2024

Agenda

1. Background
2. Matching-adjusted indirect comparison
3. Two-stage matching-adjusted indirect comparison
4. Weight truncation
5. Simulation study
6. Alternative variance reduction approaches
7. Concluding remarks

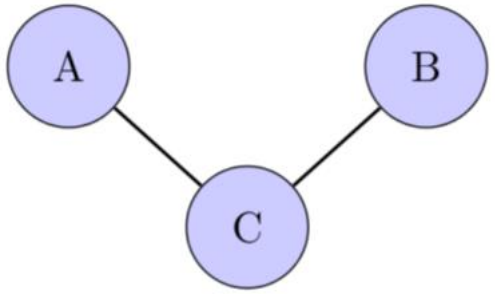
Background

Covariate-adjusted indirect treatment comparisons (ITCs)

The following setting is common in health technology assessment:

- An active treatment (treatment *A*) needs to be compared against a competitor (treatment *B*)
- No head-to-head trial between treatments *A* (*T*=1) and *B* (*T*=2)
- We have individual patient data (IPD) for study *A* ("index" study) but not for study *B* ("competitor" study)
- There are differences in baseline characteristics between study *A* (*S*=1) and study *B* (*S*=2)
- Transportability problem: transfer inferences from study *A* to study *B* for an unbiased ITC in study *B*

ANCHORED COMPARISON



$$\hat{\Delta}_{AB}^{(B)} = \hat{\Delta}_{AC}^{(B)} - \hat{\Delta}_{BC}^{(B)}$$

UNANCHORED COMPARISON

$$\hat{\Delta}_{AB}^{(B)} = g(\hat{\mu}_A^{(B)}) - g(\hat{\mu}_B^{(B)})$$

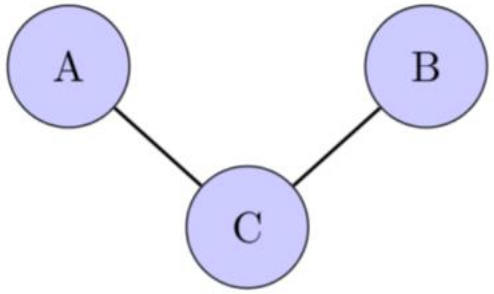
Background

Covariate-adjusted indirect treatment comparisons (ITCs)

Matching-adjusted indirect comparison (MAIC):

- The most widely used covariate-adjusted ITC method
- Weighting approach: the IPD is weighted so that there is cross-study balance in covariate moments
- Vulnerable to poor precision where covariate overlap is poor and the effective sample size after weighting is small, in which case a few extreme weights have an undue impact
- The above scenarios are pervasive in health technology appraisals

ANCHORED COMPARISON



$$\hat{\Delta}_{AB}^{(B)} = \hat{\Delta}_{AC}^{(B)} - \hat{\Delta}_{BC}^{(B)}$$

UNANCHORED COMPARISON

$$\hat{\Delta}_{AB}^{(B)} = g(\hat{\mu}_A^{(B)}) - g(\hat{\mu}_B^{(B)})$$

Matching-adjusted indirect comparison (MAIC)

- Logistic model for trial assignment

$$\ln(w_i) = \ln[w(\boldsymbol{z}_i)] = \ln\left[\frac{Pr(S=2 \mid \boldsymbol{z}_i)}{1-Pr(S=2 \mid \boldsymbol{z}_i)}\right] = \alpha_0 + \boldsymbol{z}_i\boldsymbol{\alpha_1}$$

- Entropy balancing approach; covariate balance is viewed as a convex optimization problem

$$Q(\boldsymbol{\alpha_1}) = \sum_{i=1}^{n} \exp\left(\boldsymbol{z}_i^*\boldsymbol{\alpha_1}\right) \qquad \boldsymbol{z}_i^* = \boldsymbol{z}_i - \boldsymbol{\theta}_z$$

- The estimated weights represent the conditional odds of assignment to study *B*

$$\hat{w}_i = \exp(\boldsymbol{z}_i^*\hat{\boldsymbol{\alpha}}_{\mathbf{1}})$$

- Marginal mean outcomes and/or relative effects for study *A* treatment(s) estimated in study *B*

$$\hat{\mu}_t^{(2)} = \frac{\sum_{i=1}^{n_t} y_{i,t}\hat{w}_{i,t}}{\sum_{i=1}^{n_t} \hat{w}_{i,t}} \qquad \hat{\Delta}_{10}^{(2)} = g\left(\hat{\mu}_1^{(2)}\right) - g\left(\hat{\mu}_0^{(2)}\right)$$

- The objective of the “odds weights” is to account for covariate differences and attain balance **across studies**
- Key assumptions: conditional transportability and overlap **across studies**

Two-stage matching-adjusted indirect comparison (2SMAIC)

Modular extension to MAIC

- Additional logistic model for treatment assignment in the index trial, fitted to the IPD

$$\text{logit}[e_i] = \text{logit}[e(\boldsymbol{x}_i)] = \text{logit}[Pr(T = 1 \mid \boldsymbol{x}_i)] = \beta_0 + \boldsymbol{x}_i\boldsymbol{\beta_1}$$

- Having fitted the model, e.g., using maximum-likelihood, predict propensity scores

$$\hat{e}_i = \text{expit}[\hat{\beta}_0 + \boldsymbol{x}_i\hat{\boldsymbol{\beta}}_{\mathbf{1}}]$$

- Estimate "inverse probability of treatment" weights (IPTWs) and combine them with the odds weights

$$\hat{\omega}_i = \frac{t_i\hat{w}_i}{\hat{e}_i} + \frac{(1-t_i)\hat{w}_i}{(1-\hat{e}_i)}$$

$$\hat{\mu}_t^{(2)} = \frac{\sum_{i=1}^{n_t} y_{i,t}\hat{\omega}_{i,t}}{\sum_{i=1}^{n_t} \hat{\omega}_{i,t}} \qquad \hat{\Delta}_{10}^{(2)} = g\left(\hat{\mu}_1^{(2)}\right) - g\left(\hat{\mu}_0^{(2)}\right)$$

- The IPTWs seek to balance covariates **between the index trial treatment groups**; the combined weights seek to attain balance **between the index trial treatment groups and across studies**
- Limitation: because it relies on a treatment assignment model for the index trial, 2SMAIC is not applicable in the unanchored case

Rationale for 2SMAIC with index RCT

- In an RCT, the true treatment assignment mechanism and propensity scores are fixed and known, due to randomization
- Randomization guarantees covariate balance on expectation, in large samples
- Senn (2004): “over all randomizations the groups are balanced; for a particular randomization they are unbalanced” – there may still be finite-sample imbalances due to chance
- Estimating the propensity scores is beneficial to correct for residual imbalances between treatment arms, particularly where the index trial sample size is small
- Motivation for covariate adjustment: to increase efficiency by gaining precision, not to reduce bias!

Rationale for 2SMAIC with observational index study

- One no longer relies on the internal validity of the index study; covariate adjustment between treatment arms is necessary for confounding control
- Strong assumption #1: conditional exchangeability over treatment assignment → compromised if the treatment assignment model excludes potential confounders
- Strong assumption #2: positivity of treatment assignment → compromised by deterministic positivity violations, such as different selection criteria into the treatment groups
- Randomization is no longer leveraged to meet the strong assumptions above
- Motivation for covariate adjustment: to reduce internal validity bias due to confounding

Truncation

A simple approach for variance reduction

- Restricts the influence of extreme weights by capping the highest estimated weights at a given percentile
- The ideal truncation level will vary on a case-by-case basis and can be set empirically, e.g. by progressively truncating the weights. Density plots are helpful to assess the dispersion of the weights and identify an optimal cutoff point.
- There is a clear trade-off from a bias-variance standpoint: precision improvements always come at the cost of sacrificing balance and accepting bias
- Prior transportability/generalizability literature uses a 95th percentile cutoff; lower thresholds further reduce variance at the cost of more bias and further shifting the target population or estimand
- Limitations:

1. Shifts the target estimand definition (population or analysis set attribute)
2. Requires arbitrary ad hoc decisions on cutoff thresholds

Simulation study

Setting

- Anchored indirect treatment comparison across two RCTs
- Small sample sizes for index trial ($N \in \{140, 200\}$)
- 3 strongly prognostic and effect-modifying covariates
- Varying levels of deterministic overlap between the target populations of the RCTs
- Continuous outcome, linear outcome generating model

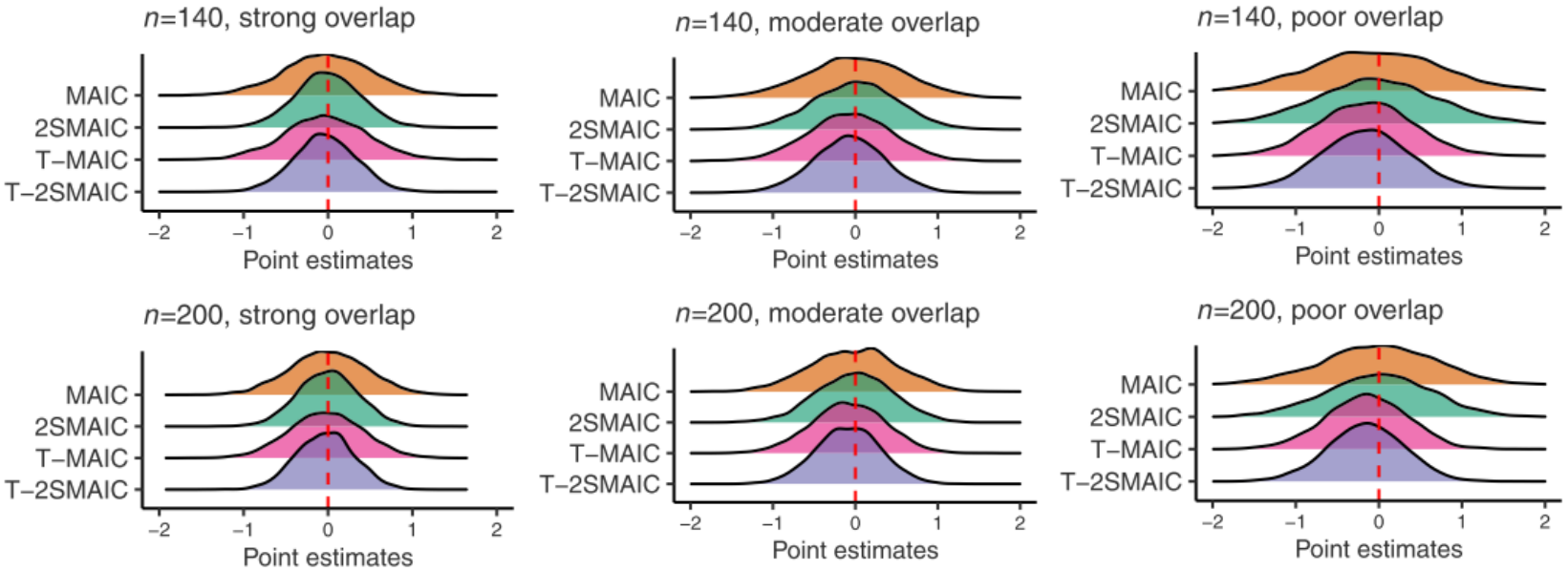
Methods

- Standard matching-adjusted indirect comparison (MAIC)
- Two-stage matching-adjusted indirect comparison (2SMAIC)
- MAIC combined with weight truncation (T-MAIC), capping the estimated weights at the 95th percentile
- 2SMAIC combined with weight truncation (T-2SMAIC), capping the estimated weights at the 95th percentile

No unmeasured covariates and cross-study balance attained for all effect-modifying moments (means)

Simulation study results

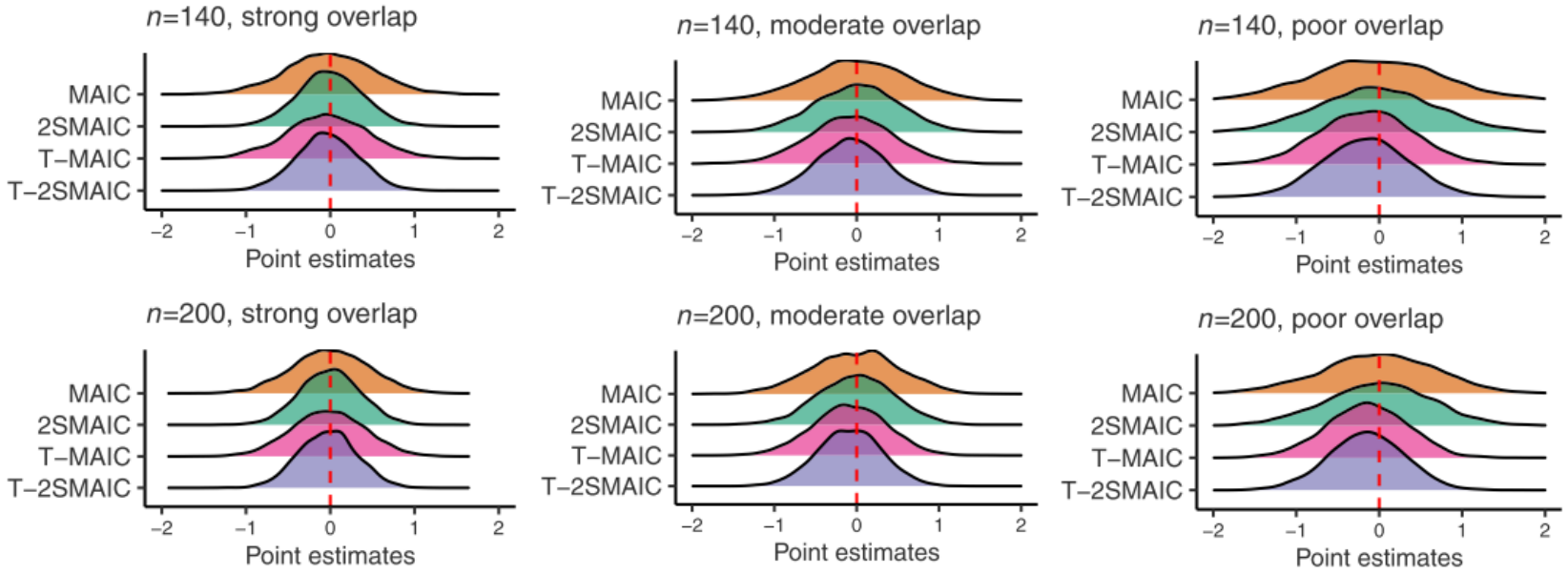
- The two-stage approaches yield improved precision and efficiency with respect to their one-stage counterparts, with similar bias
- The two-stage approaches are more effective with lower index trial sample sizes, due to greater empirical imbalances between treatment arms
- The enhanced performance of the two-stage methods is strongly linked to the prognostic strength of covariates
- Performance gains of the two-stage approaches are attenuated where overlap is poor, due to high extremity of the weights



Remiro-Azócar, A., 2022. Two-stage matching-adjusted indirect comparison. BMC medical research methodology, 22(1), p.217.

Simulation study results

- With strong covariate overlap, truncation offers modest improvements in precision and efficiency, while inducing bias
- As overlap weakens, truncation notably improves precision by reducing the influence of extreme weights, but induces considerable bias
- The combination of the two-stage method and weight truncation (T-2SMAIC) offers the best performance in terms of precision and efficiency, with the increase in precision offsetting the increase in bias
- Truncation is less necessary and bias-variance trade-offs less favorable to variance reduction where there is good overlap, the weights are well-behaved and the effective sample size after weighting is sizeable



Remiro-Azócar, A., 2022. Two-stage matching-adjusted indirect comparison. BMC medical research methodology, 22(1), p.217.

Alternative variance reduction approaches

Jackson et al. (2021) weighting scheme

- Weight estimation procedure that satisfies the conventional method of moments while explicitly maximizing the effective sample size
- Minimizes dispersion of the weights, with more stable weights improving precision at the expense of inducing bias

Reduce the number of moment-balancing conditions

- Exclude less influential covariates
- Exclude higher-order moments, e.g., only balance means and not variances
- There are bias-variance trade-offs
- These options lead to increased overlap, lower likelihood of extreme weights, less drastic reductions in effective sample size and precision
- These options also lead to residual bias (Vo 2023), particularly as marginal treatment effects generally depend on the full joint covariate distribution, including that of purely prognostic covariates (Remiro-Azócar 2024)

Alternative variance reduction approaches

Weight trimming

- Excludes subjects with outlying weights
- Shares many of the limitations of truncation: arbitrary cutoff points, change in the estimand (target population/analysis set)
- Less appealing than truncation: information directly discarded → precision loss

Weight stabilization

- Not applicable where the covariate-adjusted marginal effect is derived from the treatment coefficient of a weighted model of outcome on time-fixed binary treatment
- The fitted model is “saturated” (cannot be mis-specified)
- For saturated models, stabilized and unstabilized weights give identical results
- Potentially useful where the weighted outcome model is unsaturated, e.g., with dynamic or continuous-valued treatment regimens

Alternative variance reduction approaches

Overlap weighting

- Estimates treatment effects in a subsample with good overlap
- Challenging to implement where subject-level data are unavailable for the competitor study
- Also changes the population or analysis set attribute of the target estimand

Avoid weighting

- With weak overlap, methods based on modeling the outcome expectation, e.g., STC, model-based standardization (G-computation), ML-NMR, exhibit greater precision and efficiency than MAIC…
- …but are prone to extrapolation, which may lead to severe bias under model misspecification (Vo 2023)
- Outcome modeling is a good option where feasible numerical solutions to MAIC do not exist due lack of covariate overlap

Concluding remarks

- We have explored two strategies to improve the precision and efficiency of MAIC:
  1. Modeling the treatment assignment mechanism in the index study
  2. Truncating the weights that are above a certain level

- Weighting is inherently modular:
  1. Just like we have combined the two-stage approach with truncation, other variance reduction approaches can be incorporated, potentially in combination
  2. The estimation procedure for the trial assignment weights does not necessarily need to be entropy balancing or method of moments, alternative methods could be used
  3. Further weighting modules could be incorporated to account for missingness and non-compliance, e.g., dropout or treatment switching, in the index trial

References

- Jackson, D., Rhodes, K. and Ouwens, M., 2021. Alternative weighting schemes when performing matching-adjusted indirect comparisons. Research Synthesis Methods, 12(3), pp.333-346

- Remiro-Azócar, A., 2022. Two-stage matching-adjusted indirect comparison. BMC medical research methodology, 22(1), p.217

- Remiro-Azócar, A., 2024. Transportability of model-based estimands in evidence synthesis. In press, Statistics in Medicine

- Senn, S., 2004. Controversies concerning randomization and additivity in clinical trials. Statistics in medicine, 23(24), pp.3729-3753

- Vo, T.T., 2023. A cautionary note on the use of G-computation in population adjustment. Research Synthesis Methods, 14(3), pp.338-341