Predictive-adjusted indirect comparison (PAIC)

Antonio Remiro Azócar, Gianluca Baio, Anna Heath
Department of Statistical Science, University College London

12th June 2019

Outline

1. Background
2. Indirect treatment comparison methods
3. Predictive-adjusted indirect comparison (PAIC)
4. Simulation Study
5. Conclusion

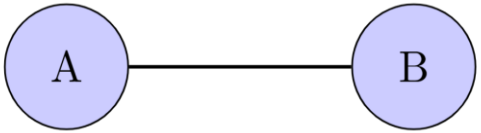
Background

- Population-adjusted indirect comparison methods are increasingly used to compare treatment outcomes across separate clinical trials and to inform health technology appraisals

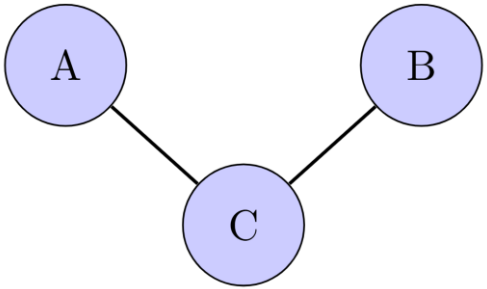
- Such methods are utilised when the following situation is encountered:
  1. No head-to-head trials have been conducted
  2. Individual-level data are available for an intervention, e.g. through the company's own trial, but are unavailable for "competing" comparator(s)
  3. There are cross-trial differences in baseline characteristics

Indirect treatment comparison

Suppose that we want to compare treatment A with treatment B. Ideally, a 2-arm randomised controlled trial would be conducted, and both treatments would be compared **directly**.



In the absence of a head-to-head trial, an **indirect** comparison is required. We shall assume that this is **anchored**: a connected treatment network is available through a common comparator C, e.g. placebo.



Indirect treatment comparison

- We have individual-level data (ILD) for trial AC (the index trial) and aggregate-level data (ALD) for trial BC, available from publications
- Standard methods for network meta-analysis and indirect comparisons, e.g. Bucher et al. (1997), allow for the use of aggregate-level data and maintain randomisation
- However, such methods are biased when the distribution of effect modifiers differs between trials
- **Population-adjusted indirect comparisons** adjust for cross-trial differences in observed covariates

Population-adjusted indirect comparison

Recently proposed approaches for population adjustment are:

1. **Matching-adjusted indirect comparison (MAIC)**

- MAIC (Signorovitch et al., 2010) reweights the patients in the AC trial so that the weighted means of their covariates match those in the published BC summary data
- Individual weights represent the odds of being enrolled in BC vs. AC
- Weights are derived from the ILD via a form of propensity score-like analysis and applied to the outcome model to generate adjusted outcomes

2. **Simulated treatment comparison (STC)**

- STC (Caro and Ishak, 2010) involves creating an outcome regression model
- A regression model is fitted in the AC population, including all effect modifiers in imbalance, and then used to predict outcomes in the BC trial population

Population-adjusted indirect comparison

Prognostic variables vs. effect modifiers:

- It is important to know the distinction between both
- Prognostic variables (PVs) are covariates that affect outcome
- Effect modifiers (EMs) are scale-specific and “interact” with treatment to affect outcome
- In the anchored setting, only differences in the distribution of effect modifiers affect inference
- National Institute of Health and Care Excellence (NICE) recommendation:
  - MAIC: only adjust for effect modifiers
  - STC: needs only adjusting for EMs but may include other covariates to increase precision

$$g(p(Y_i = 1|X_i, T_i)) = \beta_0 + \underbrace{\beta_1 X_{i1}}_{\text{PV}} + \underbrace{\beta_2 X_{i2} T_i}_{\text{EM}}$$

Predictive-adjusted indirect comparison (PAIC)

Rationale:

- We want to predict what outcomes we *might* see if we run a hypothetical AC trial on a pseudo-population exchangeable with that sampled in BC
- We can then estimate a treatment effect for A vs. C in the BC population, δ^*_{CA} and directly compare it with the B vs. C treatment effect, δ_{CB}, derived from the published data, to estimate an unbiased treatment effect for A vs. B, δ^*_{BA}

Some assumptions:

- Appropriate randomisation within trials, i.e. treatment assignment is independent of the covariates
- Complete information on relevant effect modifiers from both trials
- Prognostic variables influence outcome in the same way in both trials
- Effect modifiers influence outcome and interact with treatment in the same way in both trials

Predictive-adjusted indirect comparison (PAIC)

Available data:

- Available in ILD: $\{(X_i^{AC}, Y_i^{AC}, T_i^{AC})\}$
- Available in ALD: $\{(\theta_{BC}, \bar{Y}_T^{BC})\}$

where:

- Y_i^{AC} is the observed outcome for the *i*-th individual in AC, e.g. a continuous measurement (survival time) or a binary outcome (an adverse event)
- $X_i^{AC} = (X_{i1}^{AC}, \ldots, X_{iK}^{AC})$ is a vector of K baseline covariates observed for patient i in AC. These may be characteristics such as age, sex, comorbidities, etc.
- T_i^{AC} is the treatment indicator; 1 if patient i is in active treatment; 0 if under the common comparator
- $\{(\theta_{BC}, \bar{Y}_T^{BC})\}$ are the published BC summary data (proportions/averages) on covariates, e.g. average age, proportion of males, etc., and outcome (for each treatment), respectively
- The treatment effect for B vs. C, δ_{CB}, has either been published or is computed from the outcomes

PAIC algorithm

1. Use the AC individual-level data $\{(X_i^{AC}, Y_i^{AC}, T_i^{AC})\}$ to model the link between the outcome and the relevant prognostic variables and effect modifiers, via a GLM:

$$Y_i^{AC} \sim P(Y_i^{AC}|\mu_i, \alpha)$$
$$g(\mu_i) = X_i^{AC}\beta_X + X_{i,(EM)}^{AC}\beta_{(EM)}T_i^{AC}$$

For each draw of a large number of draws, e.g. $j = 1, \ldots, J$:

2. Simulate the process of treatment allocation for each patient:

$$T_{ij}^* \sim \text{Bernoulli}(\pi^{AC})$$

where π^{AC} is the proportion of patients under active treatment in AC.

3. "Forward sample" the aggregate θ^{BC} to predict the covariates X_{ij}^* that may be observed if an AC trial were to be run in the BC population:

$$X_{ij}^* \sim P(X_{ij}^*|\theta^{BC})$$

PAIC algorithm

4. Predict the outcome Y_{ij}^* for an AC trial conducted in a population exchangeable with that of BC, by simulating from the predictive distribution:

$$Y_{ij}^* \sim P(Y_{ij}^* | \mu_{ij}^*, \alpha)$$
$$g(\mu_{ij}^*) = X_{ij}^* \beta_X + X_{ij,(EM)}^* \beta_{(EM)} T_i^*$$

5. At this stage, the procedure has generated *J* simulated datasets $\{(X_{ij}^*, Y_{ij}^*, T_{ij}^*)\}$ These are used to estimate the A vs. C treatment effect δ_j^* for each draw:

$$Y_{ij}^* \sim P(Y_{ij}^* | \eta_{ij}^*, \alpha)$$
$$g(\eta_{ij}^*) = \delta_j^* T_i^*$$

6. Finally, the average A vs. C treatment effect is contrasted against that for B vs. C to estimate δ_{BA}^*:

$$\delta_{CA}^* = \frac{1}{J} \sum_{j=1}^{J} \delta_j^* \qquad \delta_{BA}^* = \delta_{CA}^* - \delta_{CB}$$

Simulation study (base case)

- 1,000 Monte Carlo replicates
- N=600 per trial (300 patients per arm)
- ILD generated for AC and BC; then aggregated for BC
- Binary outcome (the occurrence of an adverse event) simulated from Bernoulli distributions with probabilities generated from logistic regression
- 4 dichotomous covariates. All are prognostic variables; 2 also are effect modifiers
- Covariate marginal probabilities i.i.d. and drawn from Bernoulli distributions
- 1,000 pseudo-populations simulated for each replicate
- Strong treatment effect for active treatment (A or B) vs. placebo (C): log OR=-0.92
- Moderate prognostic variable and interaction effects (log OR=0.41)
- Strong cross-trial covariate imbalance (40% difference)

$$P(Y_i = 1|X_i = 0, T_i = 0) = 0.5 \qquad P(Y_i = 1|X_i = 1, T_i = 0) \approx 0.84$$

$$P(Y_i = 1|X_i = 0, T_i = 1) \approx 0.29 \qquad P(Y_i = 1|X_i = 1, T_i = 1) \approx 0.82$$

Simulation study (scenarios)

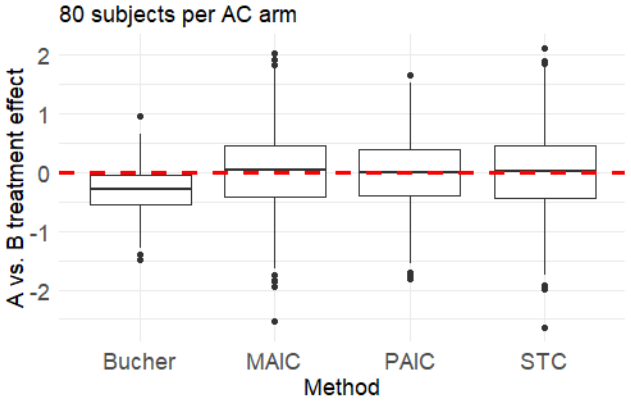
To benchmark against MAIC and STC, we vary several settings:

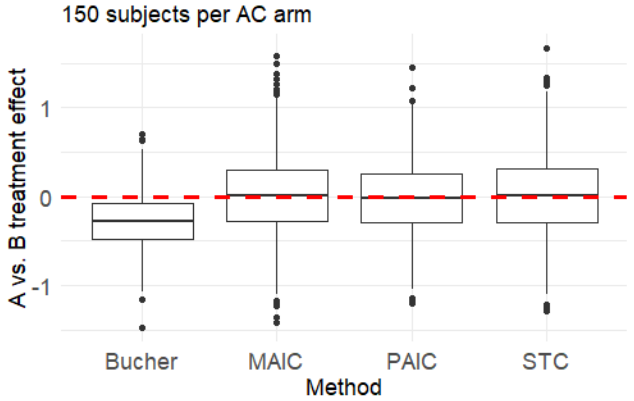
- AC/BC sample size
- Strength of A vs. C and B vs. C treatment effect (moderate/strong/none)
- Explanatory power of prognostic variables
- Strength of interaction of effect modifiers
- Degree of covariate imbalance

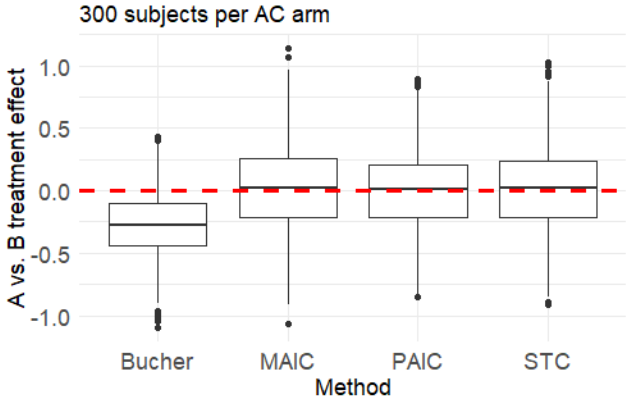
Potential model misspecification scenarios are explored:

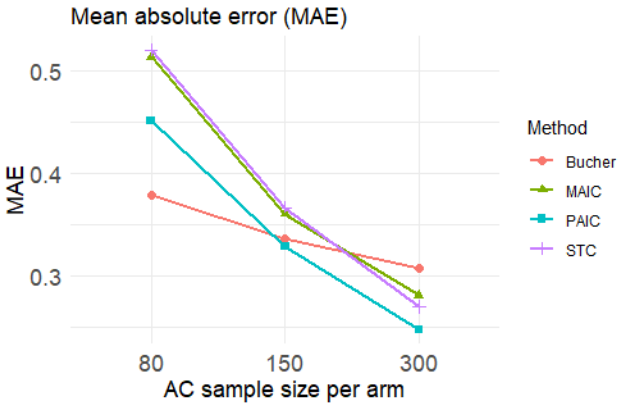
- Misspecification of EMs, e.g. due to poor selection
- Misspecification of PVs (and EMs) due to unavailable covariate information
- Effect modifiers interact differently with treatment A than with treatment B

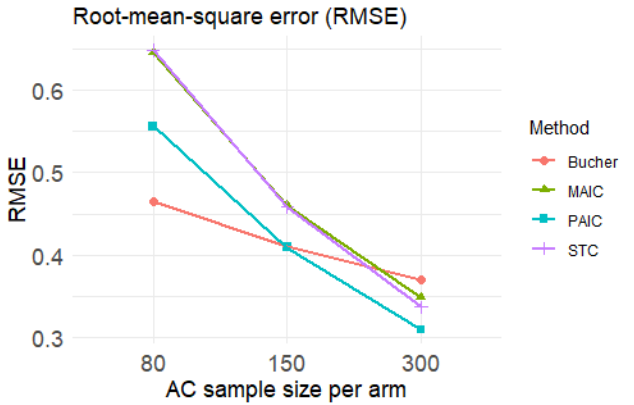
Effect of AC sample size

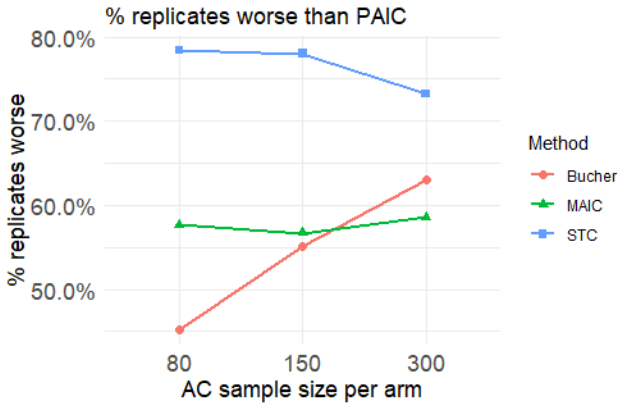




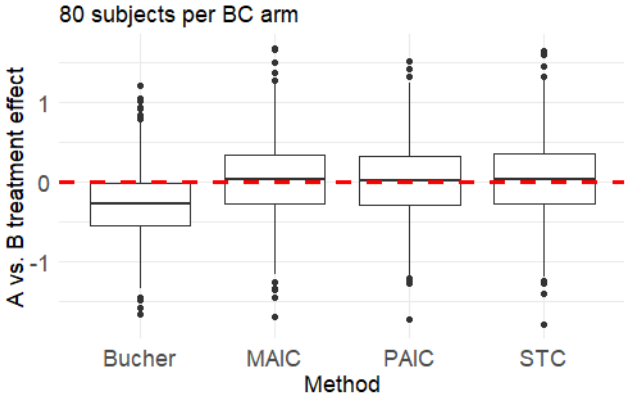


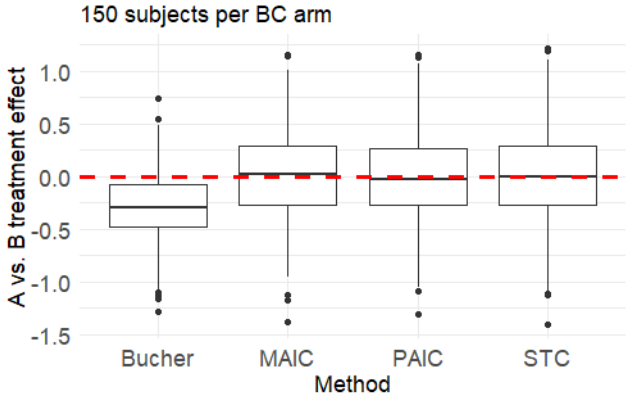


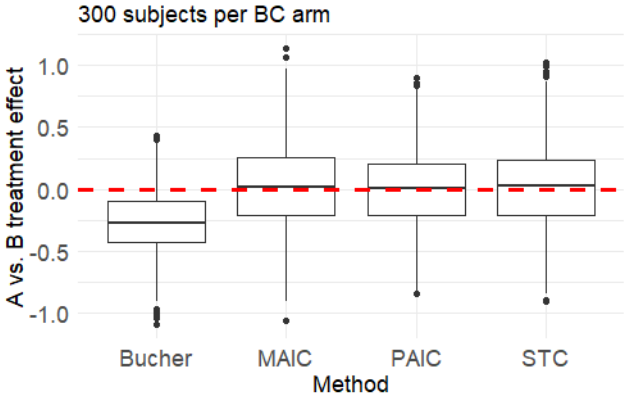


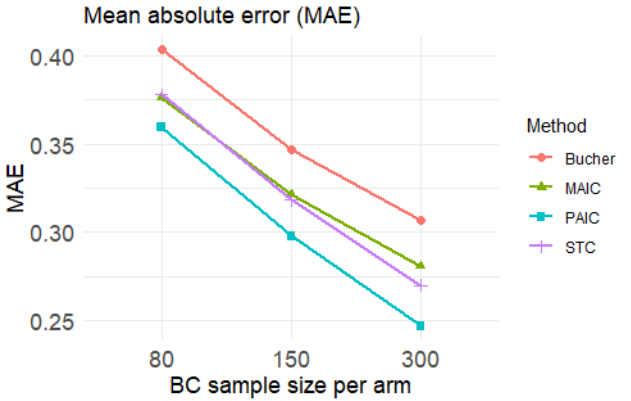


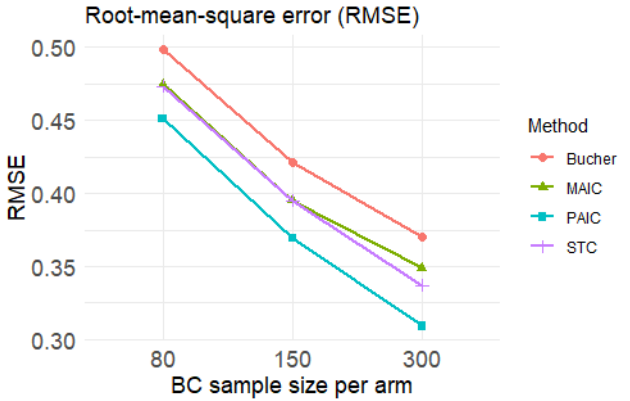
Effect of BC sample size

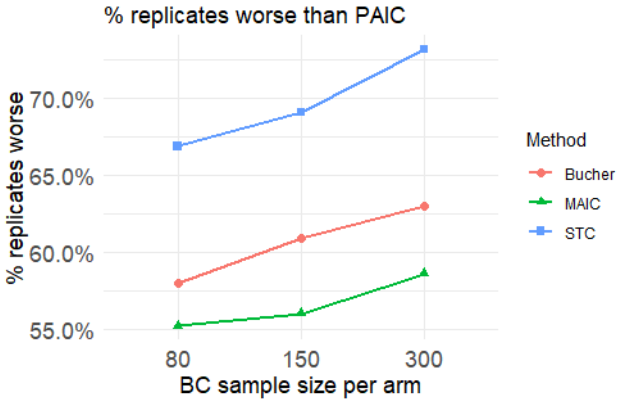




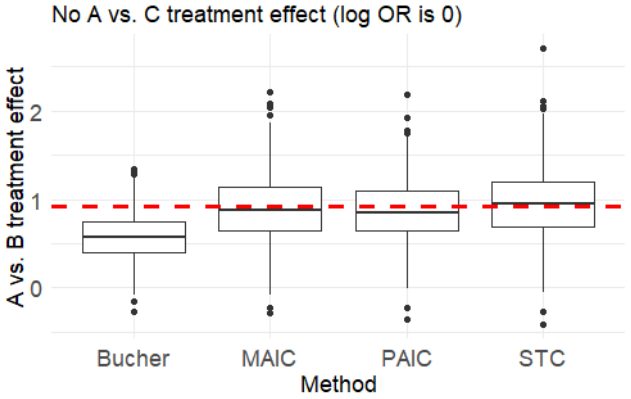


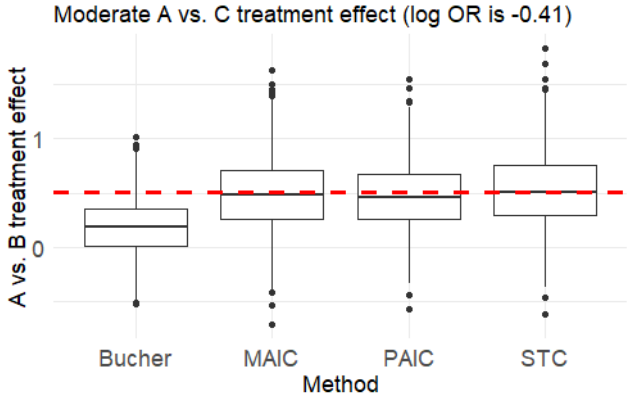


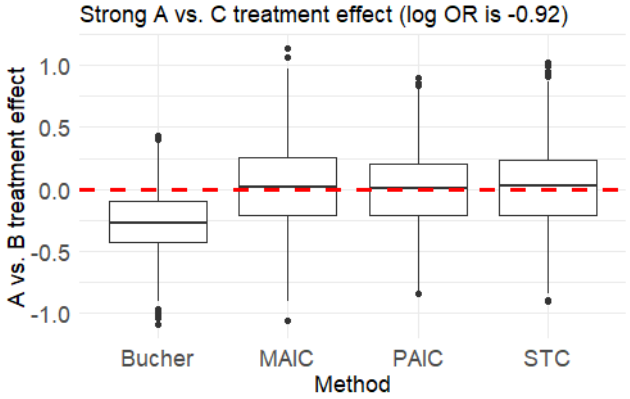


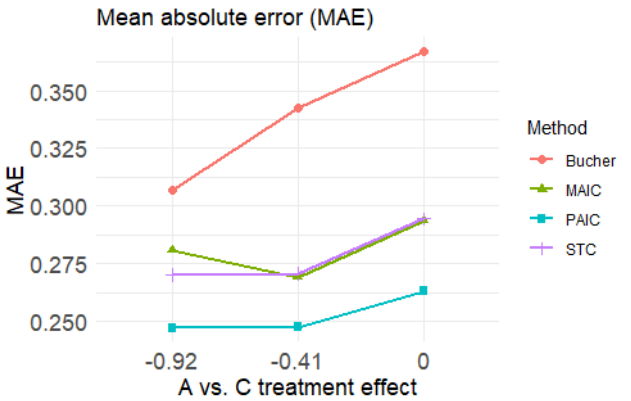


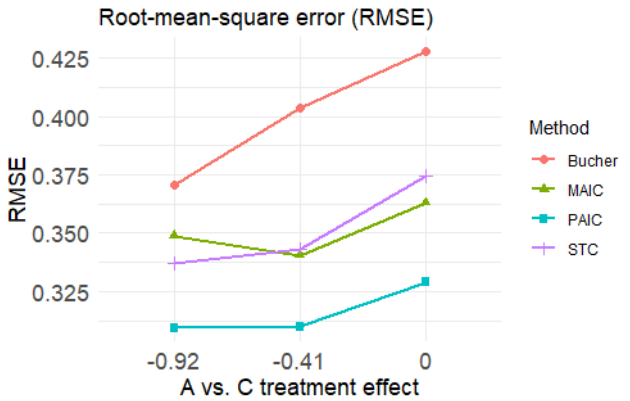
Strength of A vs. C treatment effect

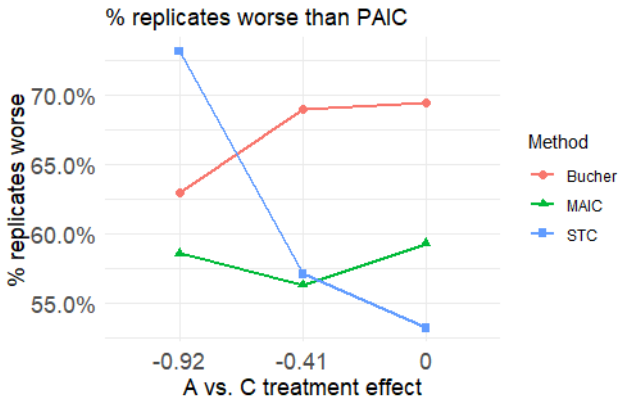




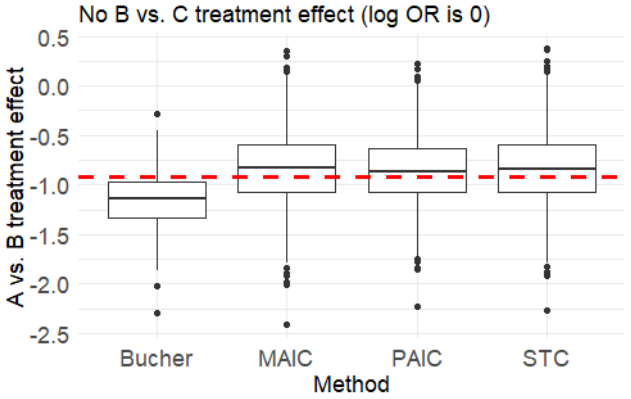


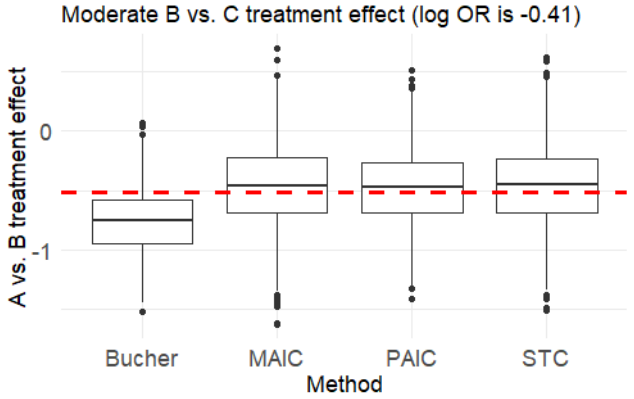


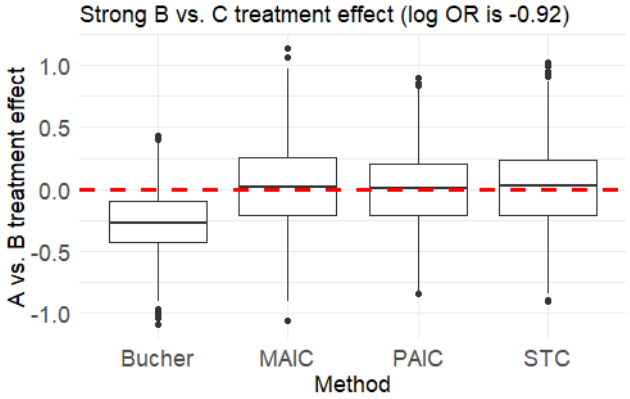


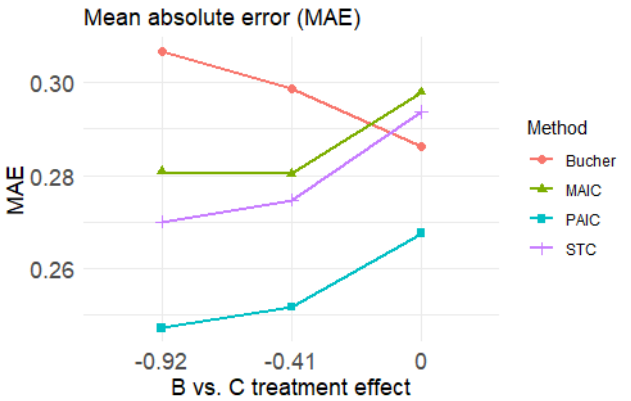


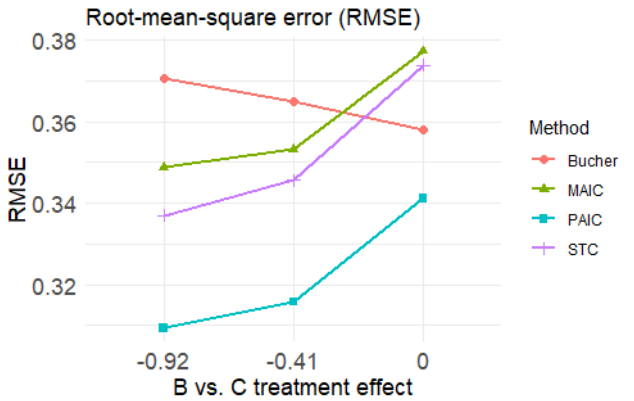
Strength of B vs. C treatment effect

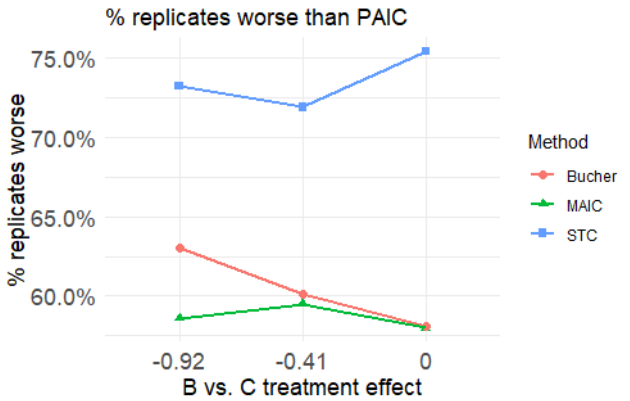




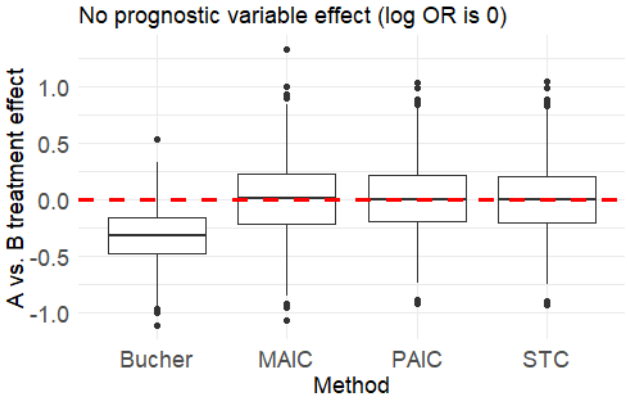


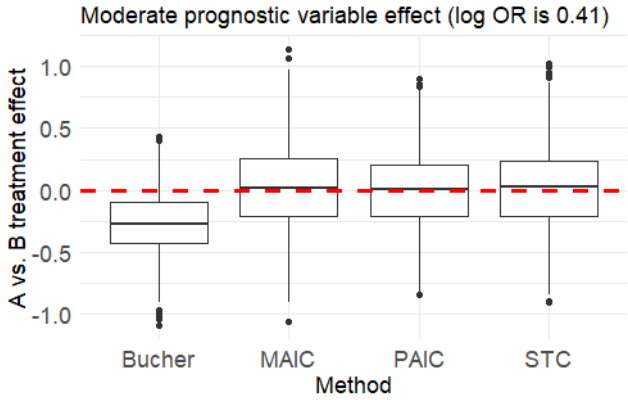


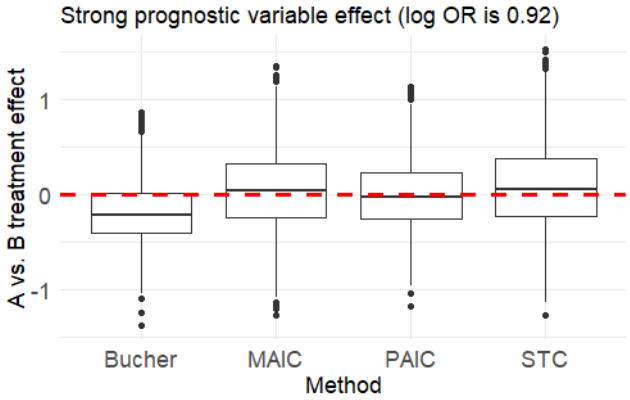


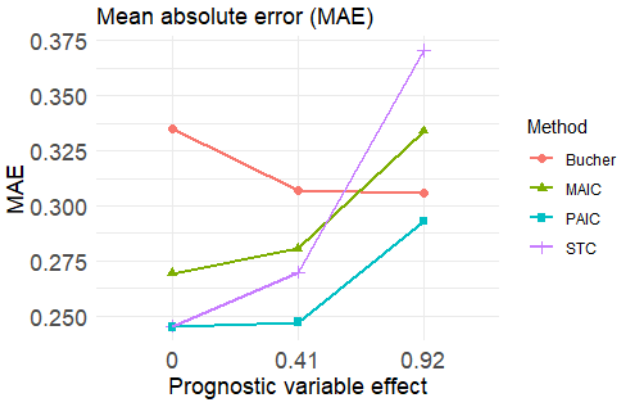


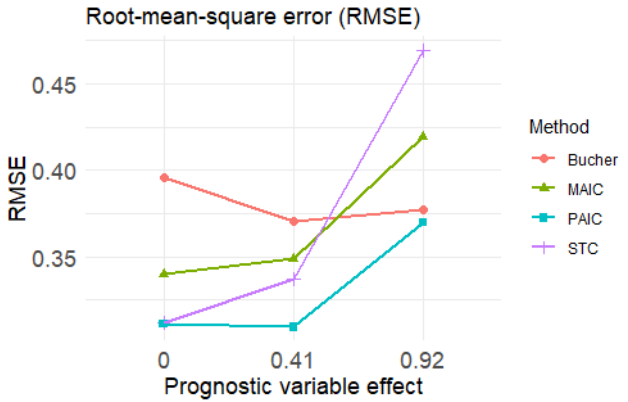
Explanatory power of prognostic variables

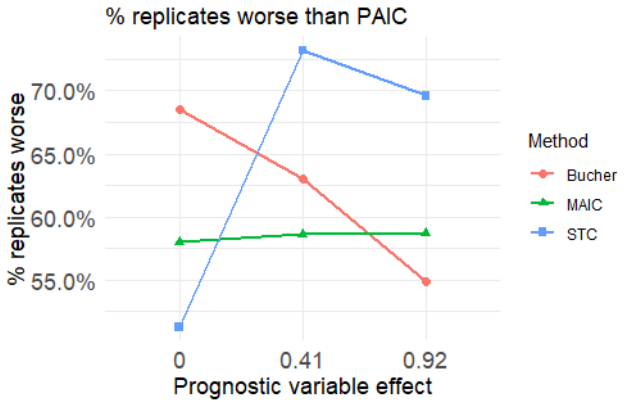




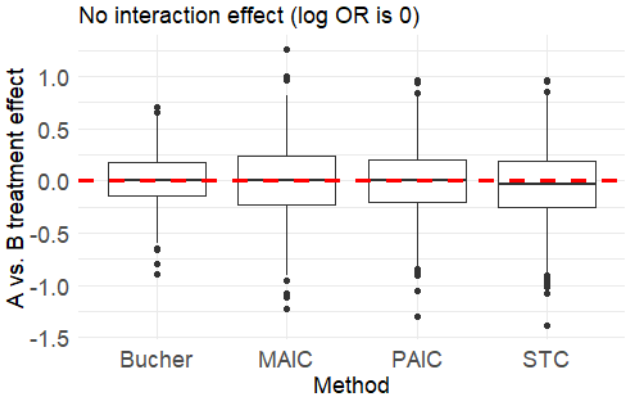


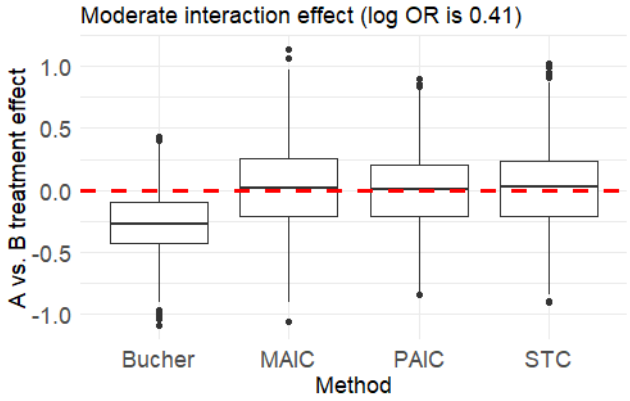


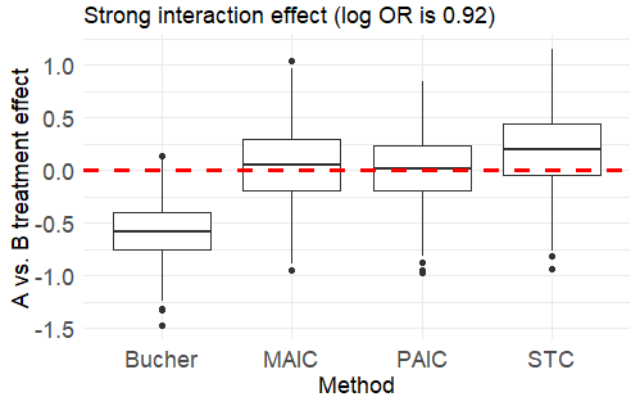


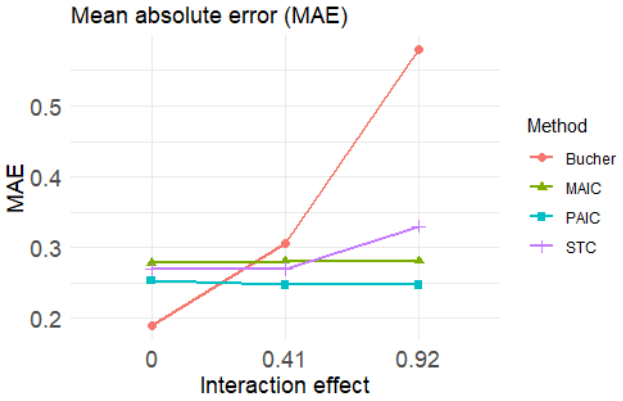


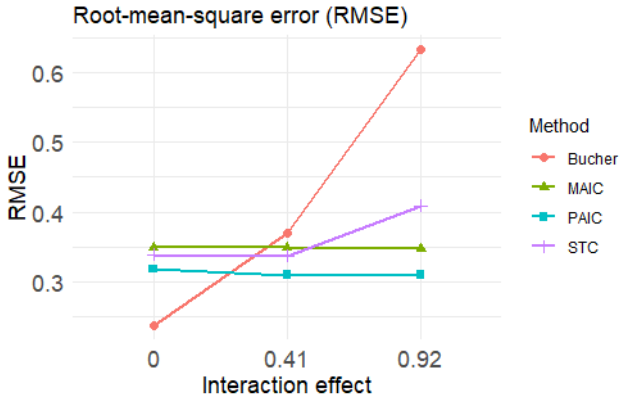
Strength of effect-modification (interaction term)

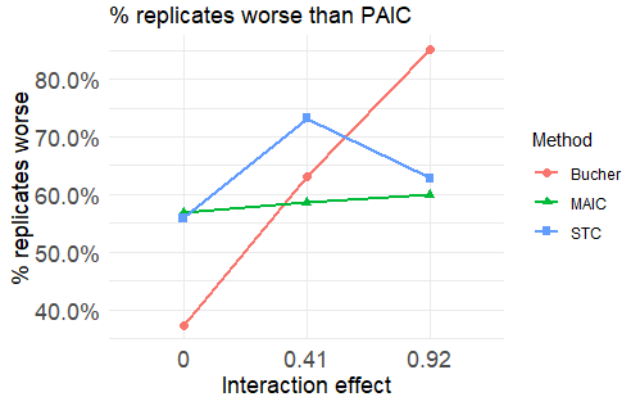




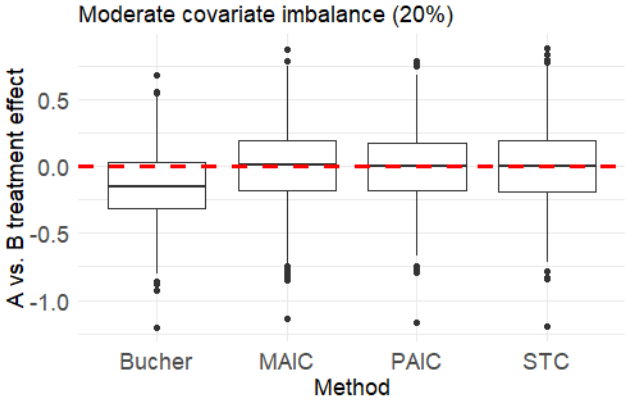


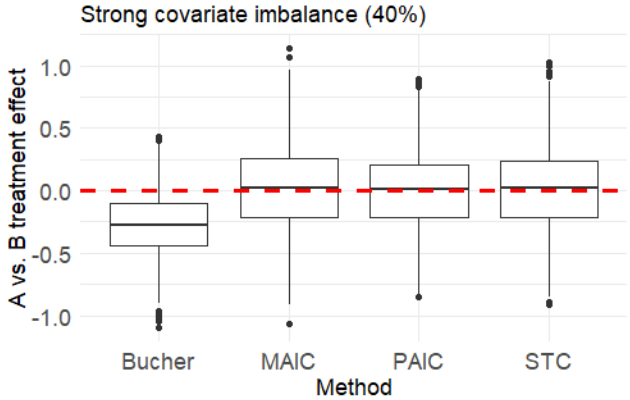


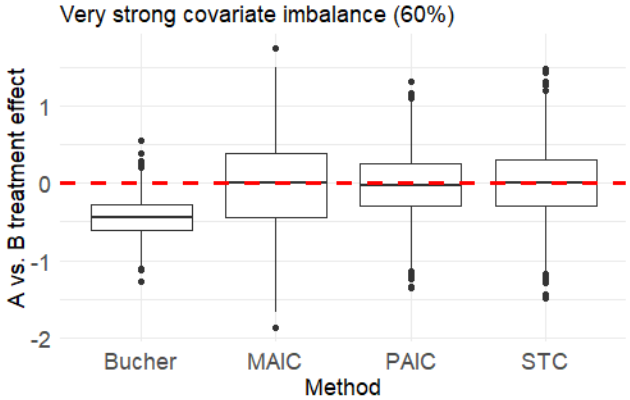


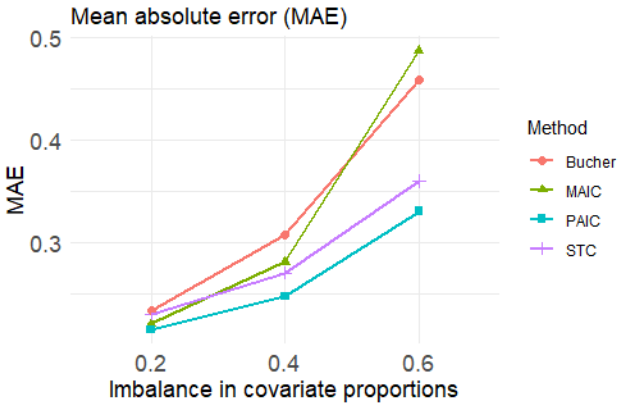


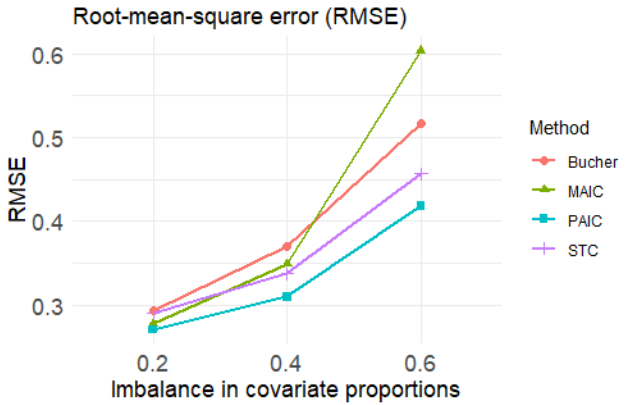
Degree of covariate imbalance

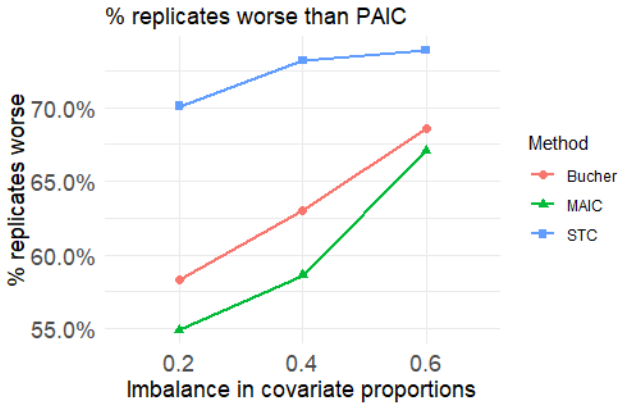




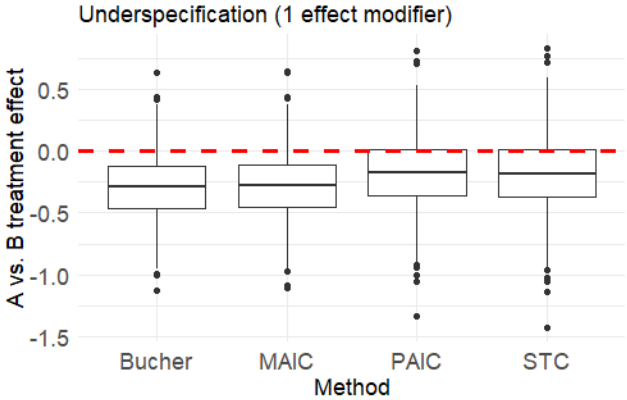


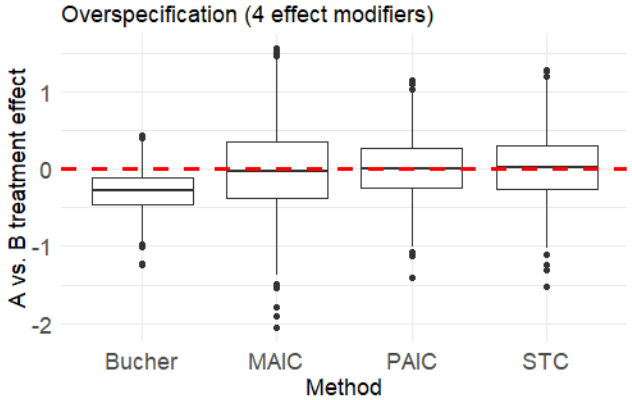


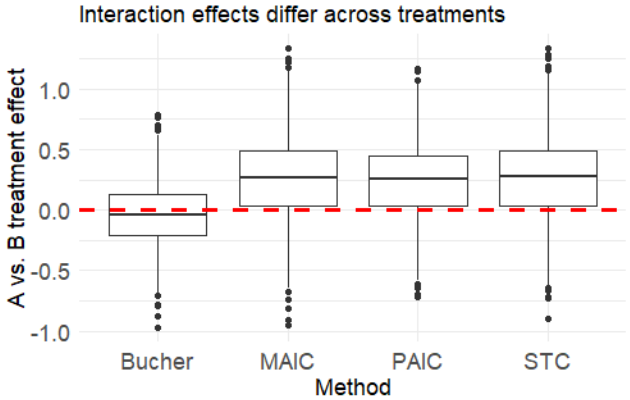


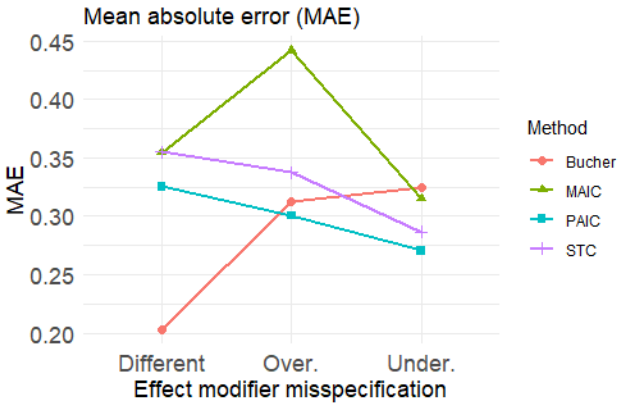


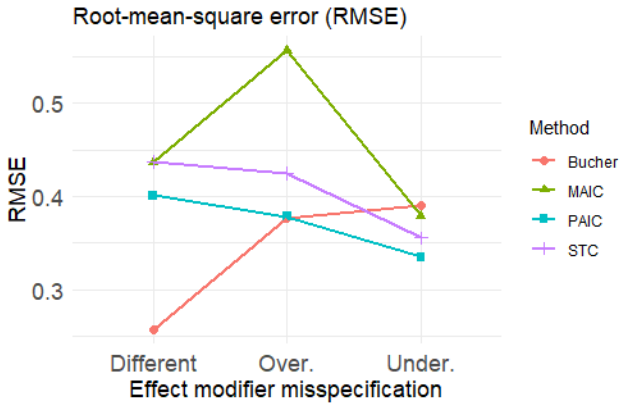
Effect modifier misspecification

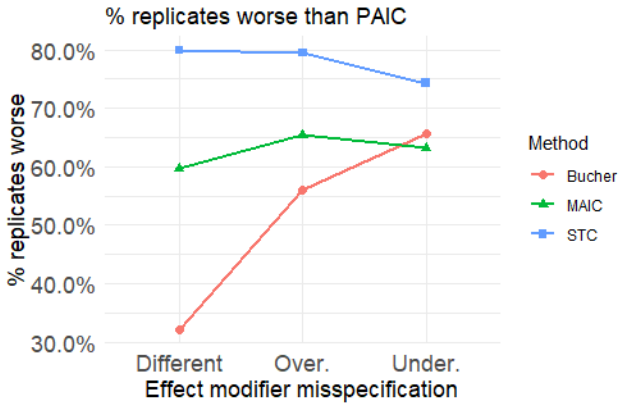




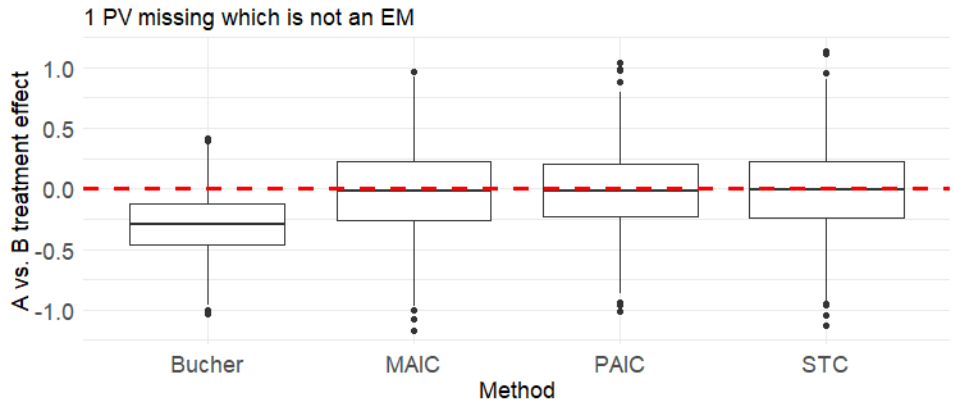


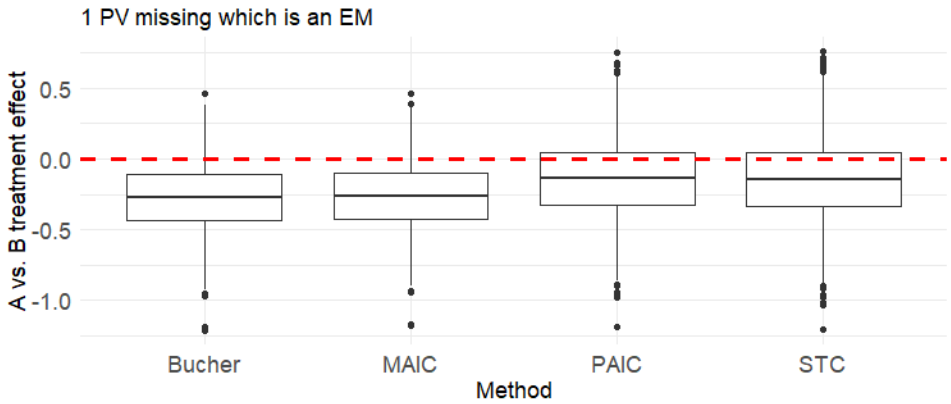


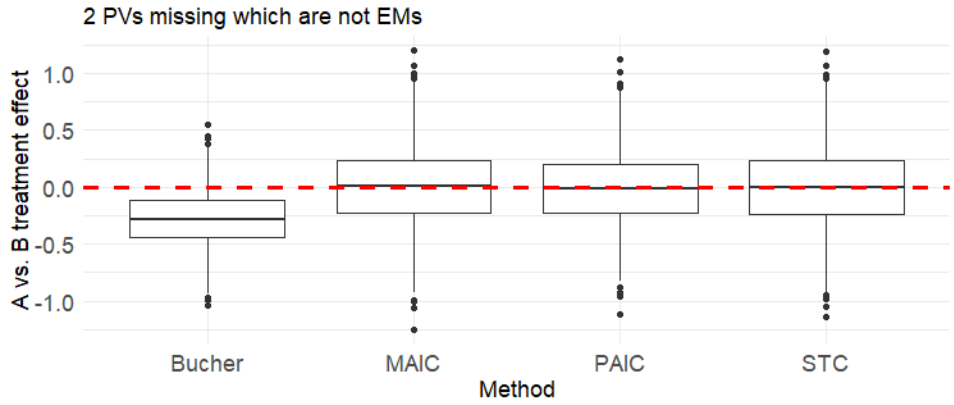


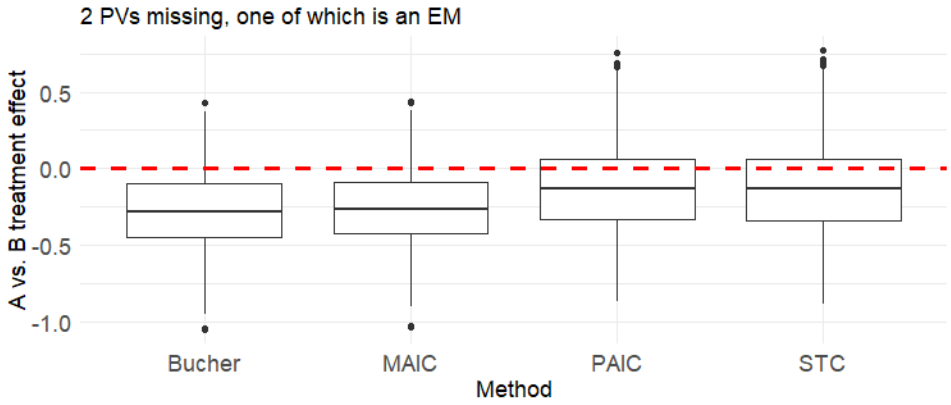


Prognostic variable misspecification









Conclusion I

- PAIC, MAIC and STC are unbiased when all effect modifiers are accounted for and when these interact with treatment in the same way in both trials. These methods cannot adjust for unobserved confounders

- These assumptions seem hard to be met in practice, particularly given the limited data/knowledge on what constitutes an EM for new interventions

- PAIC and STC were less biased than MAIC when effect modifiers were misspecified

- Standard indirect comparison methods such as the Bucher method are the gold standard when covariates are balanced but are clearly biased and inappropriate when they are not

Conclusion II

- PAIC consistently provided less biased and more accurate estimates (in terms of MAE and RMSE) across scenarios and within scenarios, on a per-replicate basis
- MAIC displays large variance over replicates when covariate imbalance is strong, as the weighting considerably reduces the effective sample size
- STC provided biased estimates when the interaction effect was strong
- Indirect comparison methods should be applied with caution under the current setup: logistic regression (binary outcome), small sample sizes and strong covariate imbalance. Such methods appear underpowered to consistently recover the true treatment effect in this setting

References

- Bucher, H. C., Guyatt, G. H., Griffith, L. E., & Walter, S. D. (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of clinical epidemiology, 50(6), 683-691

- Caro, J. J., & Ishak, K. J. (2010). No head-to-head trial? Simulate the missing arms. Pharmacoeconomics, 28(10), 957-967.

- Signorovitch, J. E., Wu, E. Q., Andrew, P. Y., Gerrits, C. M., Kantor, E., Bao, Y., & Mulani, P. M. (2010). Comparative effectiveness without head-to-head trials. Pharmacoeconomics, 28(10), 935-945.