Augmenting treatment arms with external data through propensity-score weighted power-priors: an application in expanded access

Tobias Polak[1,2,3,4], Jeremy Labrecque[2], Carin Uyl-de Groot[4], Joost van Rosmalen[1,2]

[1]Department of Biostatistics, Erasmus MC, Rotterdam, the Netherlands

[2]Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands

[3]RWD Department, myTomorrows, Amsterdam, The Netherlands

[4]Erasmus School of Health & Policy Management, Rotterdam, the Netherlands

Bayes Pharma 2023

Erasmus MC
University Medical Center Rotterdam

Introduction

- Leveraging information from external controls can improve power and precision of a clinical trial.
- Classically, methods have focused on using historical controls (Pocock 1976).
- In recent years, increasing interest in using real-world data (RWD).
- Methods focused mostly on external controls, not on external treatment arms.

Introduction

Pocock's criteria for suitability of historical controls:

- Same treatment
- Historical trial is a recent clinical study
- Same requirements for patient eligibility
- Same methods of treatment evaluation
- Comparable distributions of important patient characteristics
- Trials performed in the same organization
- No other reasons to expect different results

Introduction

- Two sources of potential bias when including external data in RCT:
  1. Measured confounding
  2. Unmeasured confounding
- Bayesian dynamic borrowing methods mitigate unmeasured confounding by downweighting external data.
- Main methods: power prior, meta-analytic-predictive prior, commensurate prior (Viele et al. Pharm Stat 2014)
- Informative prior for analysis of RCT, based on external controls
- Covariate adjustment or propensity score methods to account for measured confounding

Introduction

- When borrowing historical controls, patient characteristics should be comparable (Pocock's criteria).
- This assumption is not plausible when borrowing RWD.
- Few hybrid methods that can adjust for both sources of bias.
- Goal of new method: provide a double safeguard against measured and unmeasured confounding when combining RWD with trial data

Hybrid methods

Hybrid methods developed so far:

- Wang et al. (J Biopharm Stat 2019): power prior with propensity score stratification
- Liu et al. (Stat Med 2021): meta-analytic predictive prior with propensity score stratification
- Wang et al. (J Biopharm Stat 2022) reviewed several methods and proposed propensity weighting with fixed and commensurate priors

Disadvantages of most of these methods:

- Fixed number of external patients to be borrowed
- Many parameters/settings to be chosen for implementation

Vemurafenib application

- Expanded access: pathway of non-trial access to investigational medicine for patients with seriously debilitating or life-threatening illnesses
- Expanded access data are a source of RWD.
- Motivating example: Vemurafenib is a drug approved for late-stage melanoma harboring a V600E BRAF mutation.
- Data: single-arm phase II study (n=132) and expanded access program (n=241)
- Primary endpoint: objective response rate

Erasmus MC
University Medical Center Rotterdam

Vemurafenib application

Patient characteristics of vemurafenib case study

Characteristic	Clinical Program Expanded access, N = 241[1]	Trial, N = 132[1]
Age at enrolment	53 (13)	50 (15)
Gender assigned at birth		
Female	95 (39%)	51 (39%)
Male	146 (61%)	81 (61%)
Melanoma stage		
M1a	22 (9.1%)	33 (25%)
M1b	26 (11%)	18 (14%)
M1c	182 (76%)	80 (61%)
Unresectable Stage III	11 (4.6%)	0 (0%)
ECOG performance status		
Grade 0	112 (46%)	61 (46%)
Grade 1	98 (41%)	71 (54%)
Grade 2	30 (12%)	0 (0%)
Grade 3	1 (0.4%)	0 (0%)
Objective response rate	129 (54%)	75 (57%)

[1] Mean (SD); n (%)

Erasmus MC
University Medical Center Rotterdam

Power prior

- Power prior: Bayesian method for borrowing external data
- Formulation: $p(\theta \mid \mathcal{Y}, \delta) \propto \mathcal{L}(\theta|\mathcal{Y}_0)\mathcal{L}(\theta|\mathcal{Y}_e)^{\delta}\pi(\theta)$, where δ is weight of external data, with $0 \leq \delta \leq 1$.
- Modified power prior (MPP): estimate weight based on available data (Duan et al. 2005):

$$p(\theta, \delta \mid \mathcal{Y}) \propto \mathcal{L}(\theta|\mathcal{Y}_0)\mathcal{L}(\theta|\mathcal{Y}_e)^{\delta}\frac{1}{C(\delta)}\pi(\delta)\pi(\theta)$$

- $C(\delta) = \int_{\theta} \mathcal{L}(\theta \mid \mathcal{D}_e)^{\delta}\pi(\theta)d\theta$ is a scaling constant to ensure the MPP abides by the likelihood principle.

Propensity scores

- Propensity score methods typically used to address biases due to confounding in non-randomized experimental settings.
- Here we use propensity scores to model the allocation between current and external data (Z), with $\lambda_i = \Pr(Z = 1 \mid \mathcal{X} = x_i)$, following Lin et al. (Pharm Stat 2018).

Integrating propensity scores and power prior

- Main idea of proposed method (ProPP): combine propensity score weights with likelihood of MPP.
- ProPP method uses RWD likelihood with:

$$\mathcal{L}(\theta \mid \mathcal{Y}_e)^{\delta} = \left(\prod_i f(y_i \mid \theta)^{w_i} \right)^{\delta}$$

- Weight w_i chosen as function of the propensity score λ_i.
- Effective weight for patient i: $\delta \times w_i$.

Erasmus MC
University Medical Center Rotterdam

Integrating propensity scores and power prior

- Estimand is average treatment effect of the trial.
- Weights > 1 are not desirable:
  1. Inflated sample size in Bayesian analysis would lead to overestimation of precision.
  2. Weighting a non-trial participant higher than trial participants may not be acceptable to regulators.
- We set $w_i = 1$ for trial population and $w_i = \min\left(1, \frac{\lambda(x)}{1-\lambda(x)}\right)$ for external data.

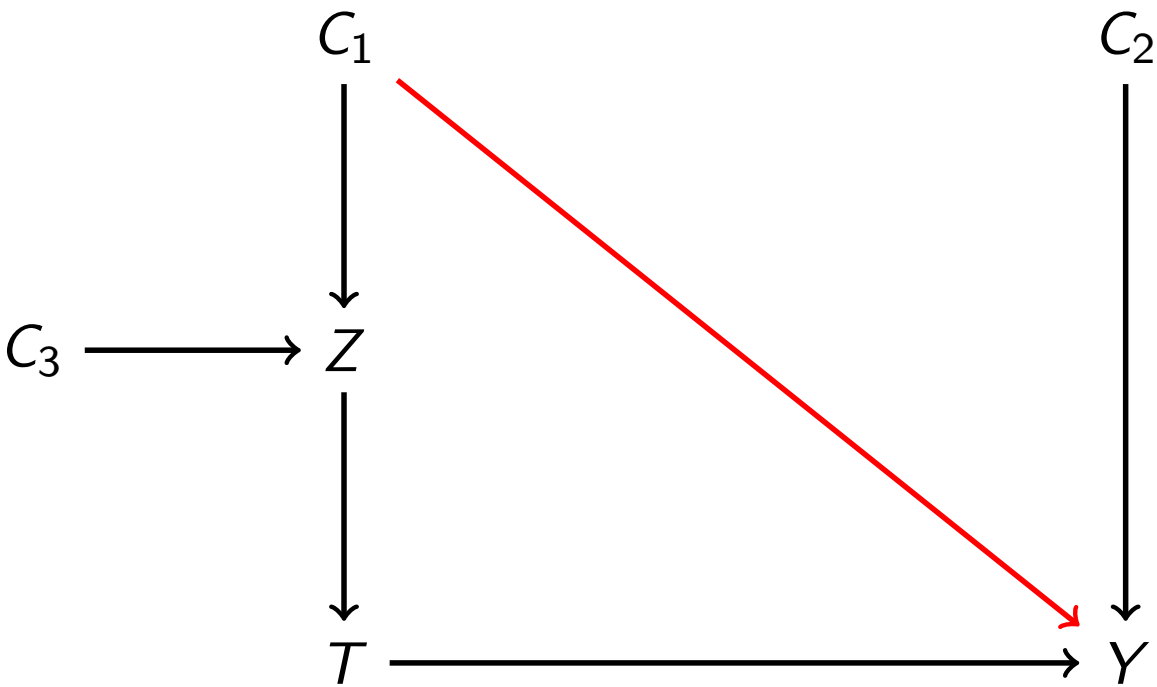
Propensity score weighting schemes under different populations of interest

Trial	External	Population of interest
$\frac{1}{\lambda(x)}$	$\frac{1}{1-\lambda(x)}$	Average treatment effect
1	$\frac{\lambda(x)}{1-\lambda(x)}$	Average treatment effect of the trial
$\frac{1-\lambda(x)}{\lambda(x)}$	1	Average treatment effect of the external

Erasmus MC
University Medical Center Rotterdam

Causal interpretation

Directed acyclic graph to explore the causal implications of the ProPP

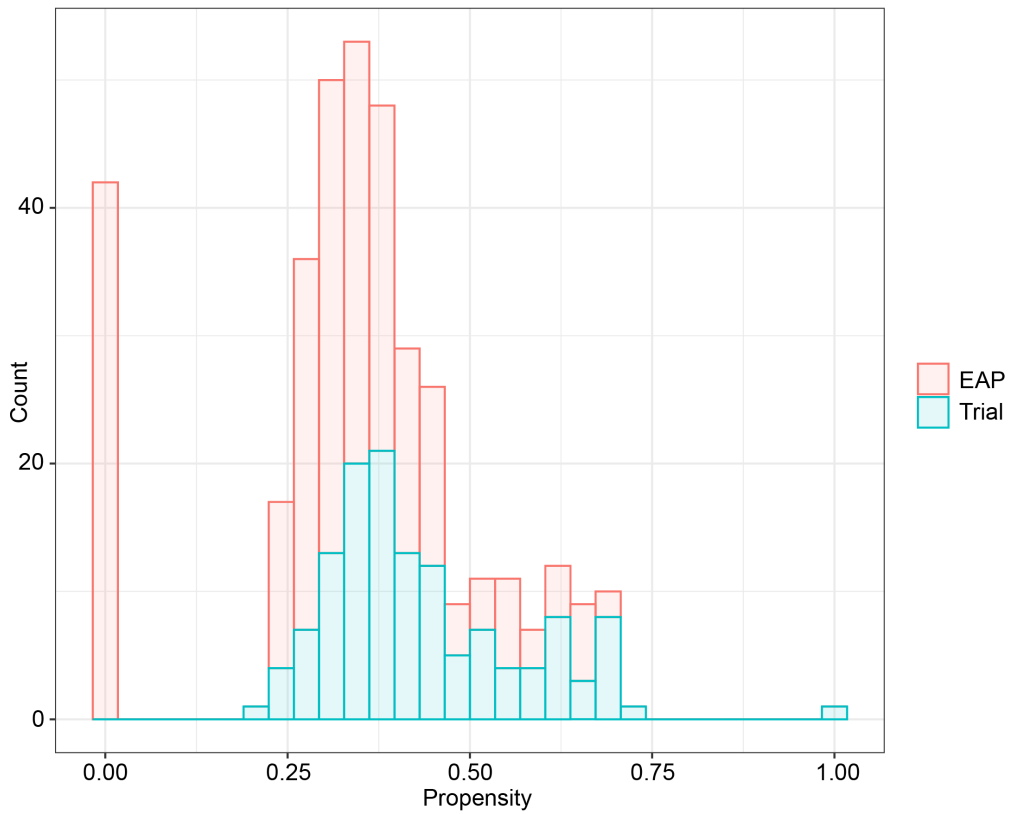


Implementation

- ProPP method applied for dichotomous outcomes
- Closed-form expressions for marginal posterior of δ and conditional posterior of θ given δ
- Easy posterior sampling using rejection sampling (no need for MCMC)

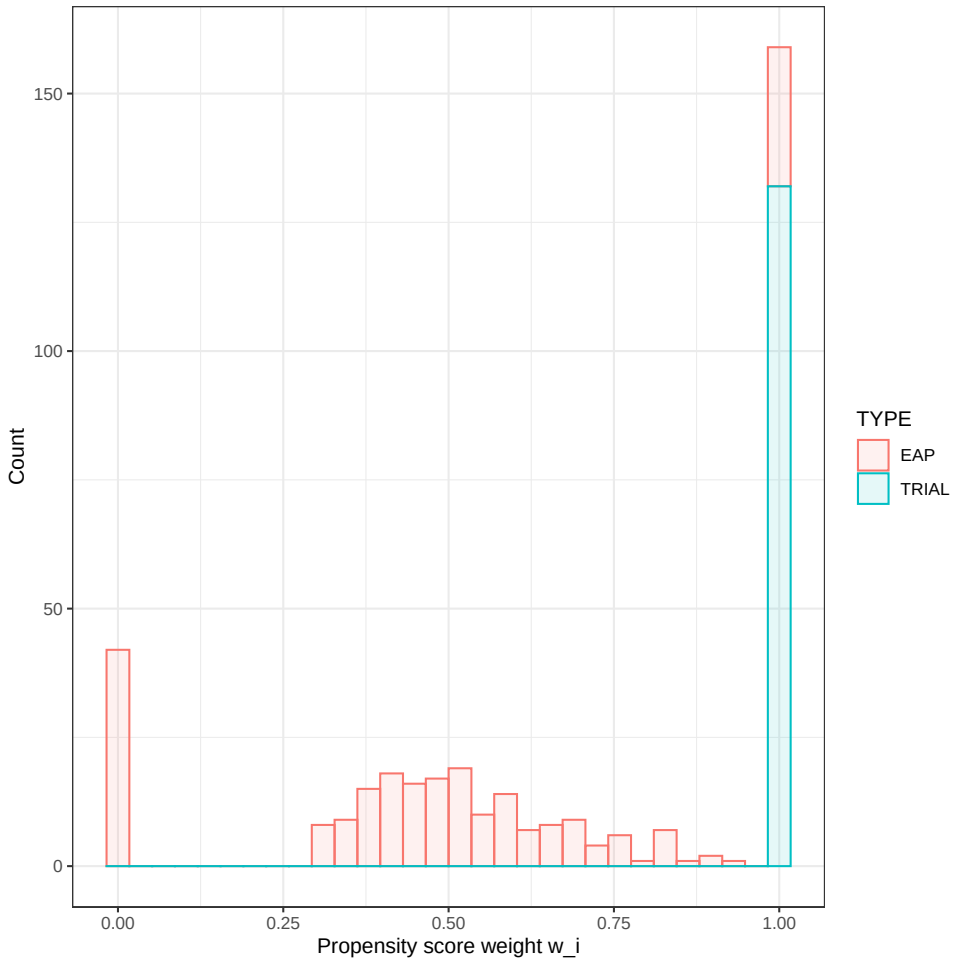
Vemurafenib application

Distribution of propensity scores in trial and expanded access program



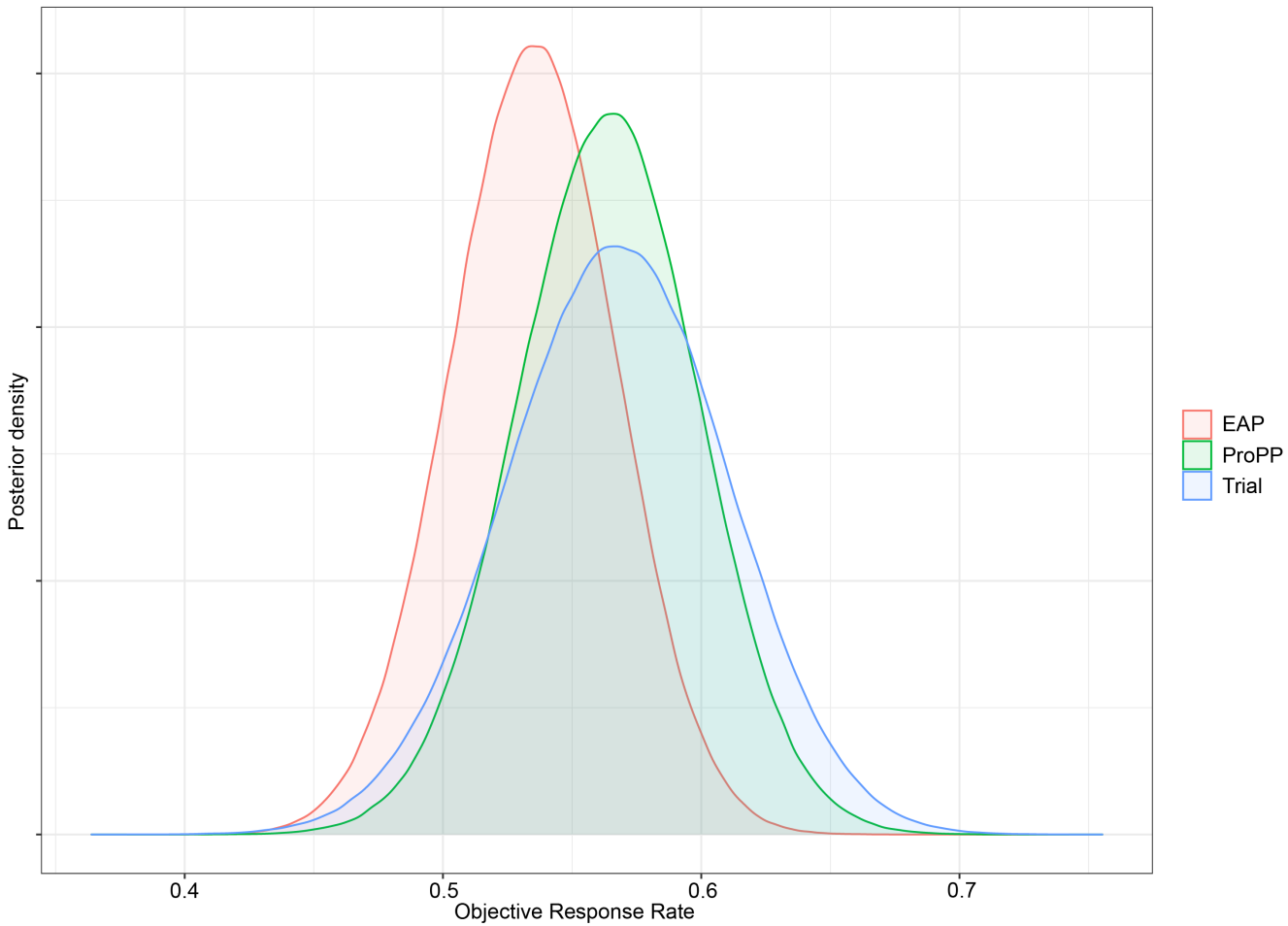
Vemurafenib application

Distribution of weights w_i in trial and expanded access program



Vemurafenib application

Inclusion of expanded access data increases precision of posterior estimates



Simulation study

- Simulation study to assess performance of ProPP in terms of type I error rate and RMSE
- Data of dichotomous outcome simulated using

$$\text{logit } y_i|x_i, z_i = \beta_0 + \boldsymbol{\beta} x_i + \eta \times I(z_i = 1)$$

- β_0 is intercept, $\boldsymbol{\beta}$ is row vector of coefficients and η is drift term.

Simulation study

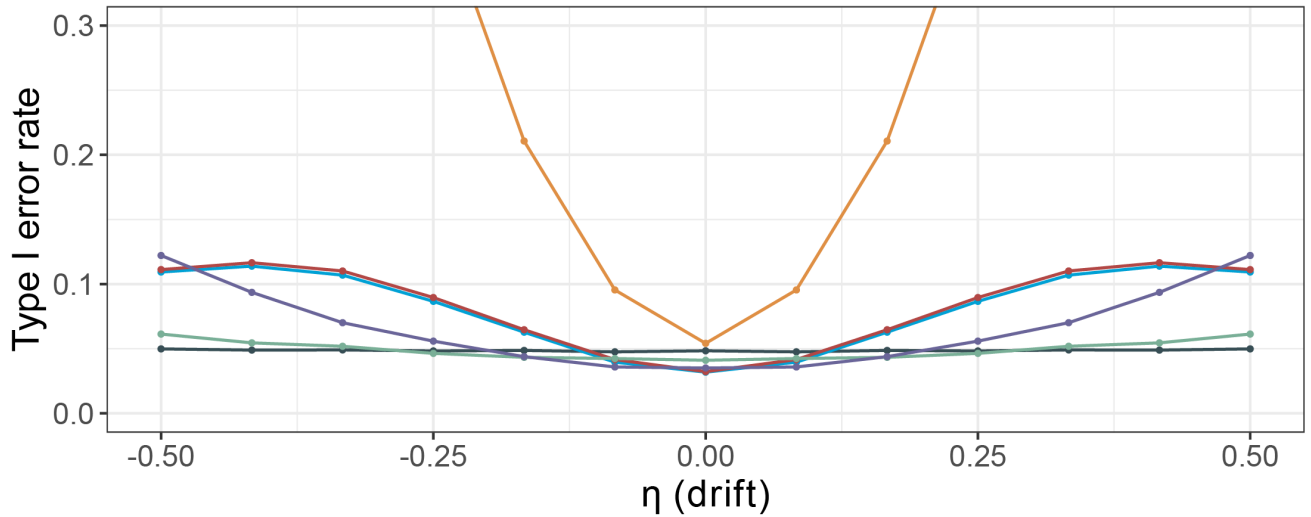
- Base case setting: $N_0 = 400$ in the trial, $N_e = 400$ in the external data, $K = 5$ different continuous covariates X with $\beta_j = 0.1$, $j = 1, \ldots, 5$ and $\beta_0 = 0$.
- Patient characteristics in trial: $X_0 \sim \mathcal{N}(\mu_0, \sigma_0^2)$
- Patient characteristics in external data: $X_e \sim (1 - \psi)\mathcal{N}(\mu_e, \sigma_e^2) + \psi\mathcal{N}(\mu_0, \sigma_0^2)$
- Degree of overlap (ψ) varied in simulations

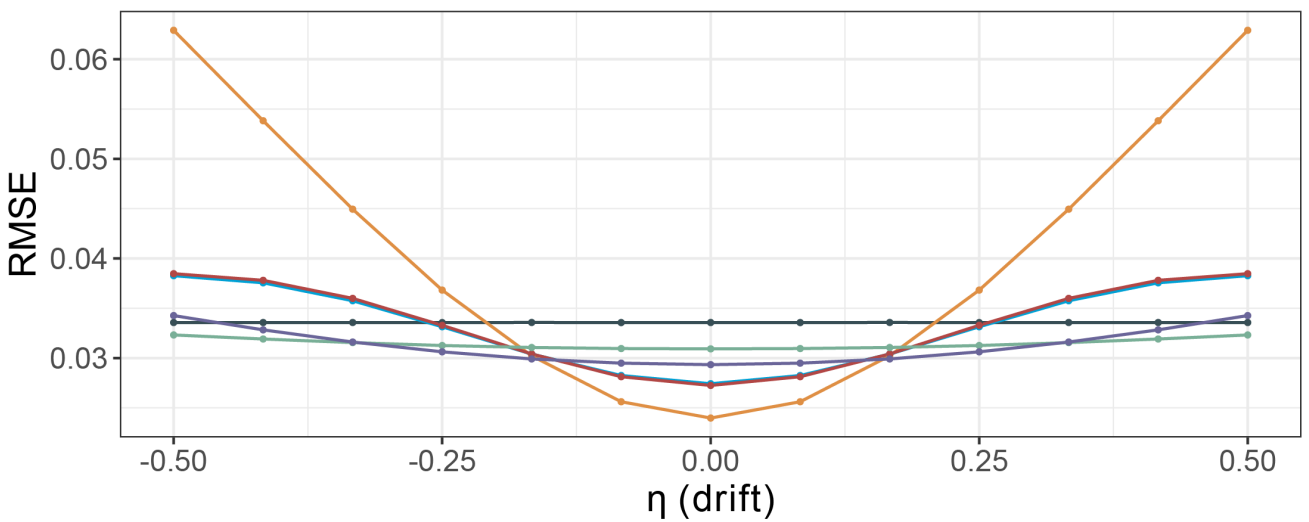
Simulation study

- Two scenarios:
  - Drift: bias due to variation in η, with $\eta \in [-0.5, 0.5]$ and $X_e, X_0 \sim N(0,1)$ and $\beta = 0$ (no effect of covariates).
  - Mixture: bias due to variation in covariate distributions (patient characteristics), with $\beta = 0.1$.
    The covariates come from a mixture distribution with $\psi = 0.5$.
    We assume $X_0 \sim \mathcal{N}(0,1)$ and we vary mean of RWD $\mu_e \in [-0.5, 0.5]$.
    There is no drift, $\eta = 0$.
- Parameter of interest: baseline rate in trial, β_0.

- ProPP method compared with:
  1. Ignore (no RWD)
  2. Pooling (combining without adjustment)
  3. Modified power prior
  4. Stratification + power-prior method suggested by Wang et al. (2019) with 10% or 20% borrowing of RWD

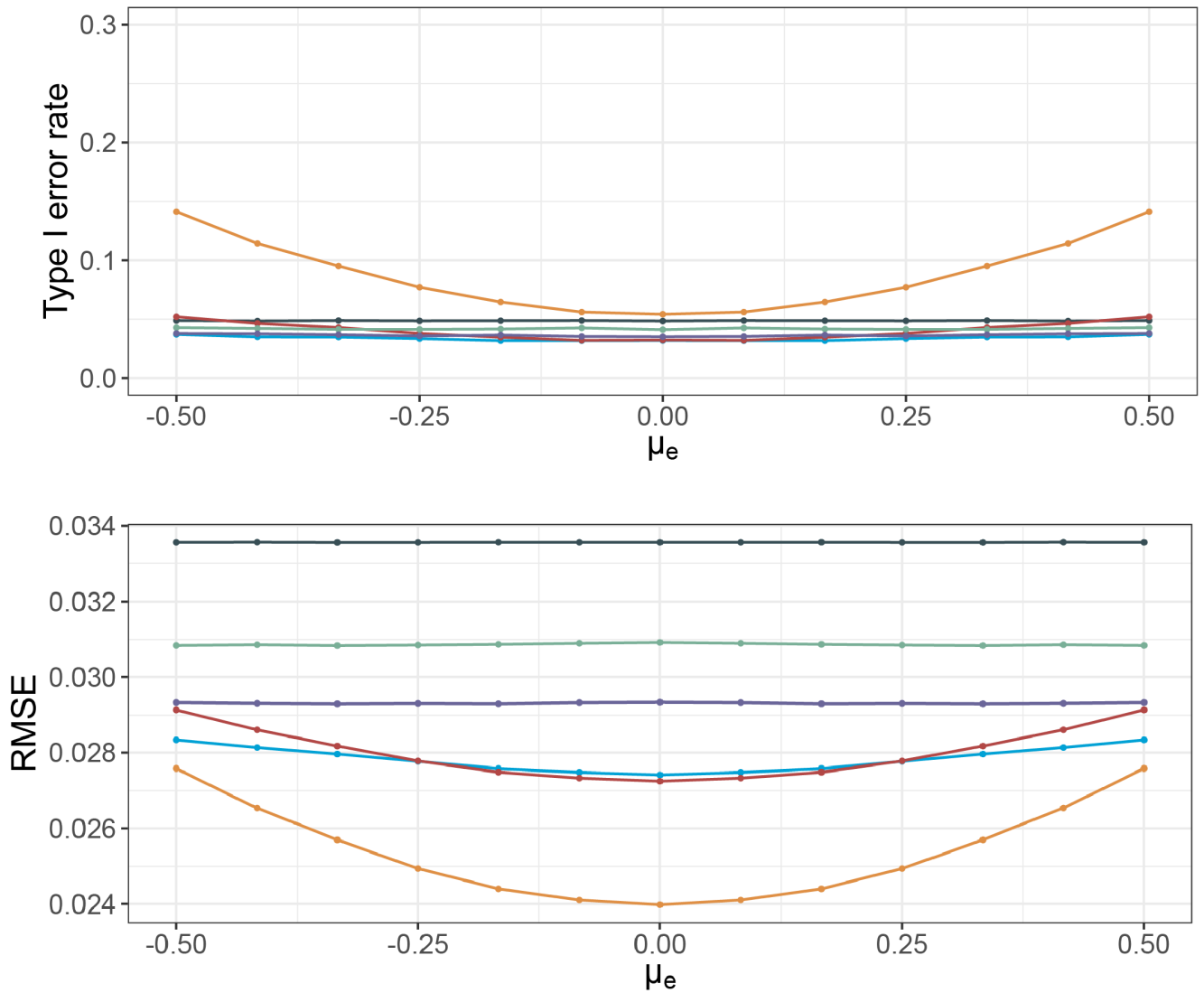
Simulation study: results of Drift scenario





Method Ignore Pooling ProPP MPP Wang 10% Wang 20%

Simulation study: results of Mixture scenario



Discussion

- ProPP compares favorably with traditional and novel methods in terms of RMSE and type-I error rate.
- Expanded access can be an important source of RWD.
- ProPP uses frequentist and Bayesian methodology to combine propensity score and dynamic borrowing methods.
- No pre-elicitation of a fixed power parameter or a fixed amount of external patients to be borrowed, but also no 'outcome-free' design.

Discussion

- Limiting the weights to a maximum of 1 may not capture all measured confounding.
- ProPP can be extended to time-to-event or normally distributed outcomes.
- Inclusion of RWD sources remains a trade-off between bias due to including non-trial data and increased precision due to increased sample size.
- ProPP allows to augment treatment arms with expanded access data in a prudent manner.