PSI Conference

A Comprehensive Self-Adaptive Mixture Prior Approach to Dynamic Borrowing from External Data

Alfredo Farjat, PhD
Ming-Dauh Wang, PhD

Statistical Innovation
Bayer Pharmaceuticals

June 17, 2026

Outline

Introduction

Motivation and challenges

Introduction

Motivation and challenges in leveraging external data for the analysis of RCTs

Randomized Control Trials (RCTs)

- Considered the “gold standard” to evaluate efficacy and safety of treatments
- Randomization aims to **balance** observed covariates across treatments
- Randomization not always feasible (recruitment, ethical, duration, costs, etc.)

External Data Sources (e.g., historical trials, RWD)

- Issues to be considered when using external data:
  - The **similarity** between trial patients and those from external data sources
  - The **amount of information** that the external data contributes to the statistical inference

Our Goal

- Develop a comprehensive, **self-adaptive** Bayesian dynamic borrowing approach that adjust for both **baseline** and **outcome** differences to determine the **amount of information** borrowing

Bayesian Dynamic Borrowing from External Data

A Brief Review

Power Prior Approach

Ibrahim and Chen (2000), Ibrahim et al (2015), Neuenschwander et al (2009)

- Bayesian inference of parameter θ, given current data D_1 and external data D_0 is through the following expression:

$$\underbrace{\pi(\theta|D_1)}_{\text{Posterior}} \propto \underbrace{L(D_1|\theta)}_{\text{Likelihood}}\underbrace{L(D_0|\theta)^{\boldsymbol{\alpha_0}}\ \pi_0(\theta)}_{\text{Power prior}}$$

Idea: construct informative prior from historical data; define power parameter to down-weight the influence of historical data on the analysis

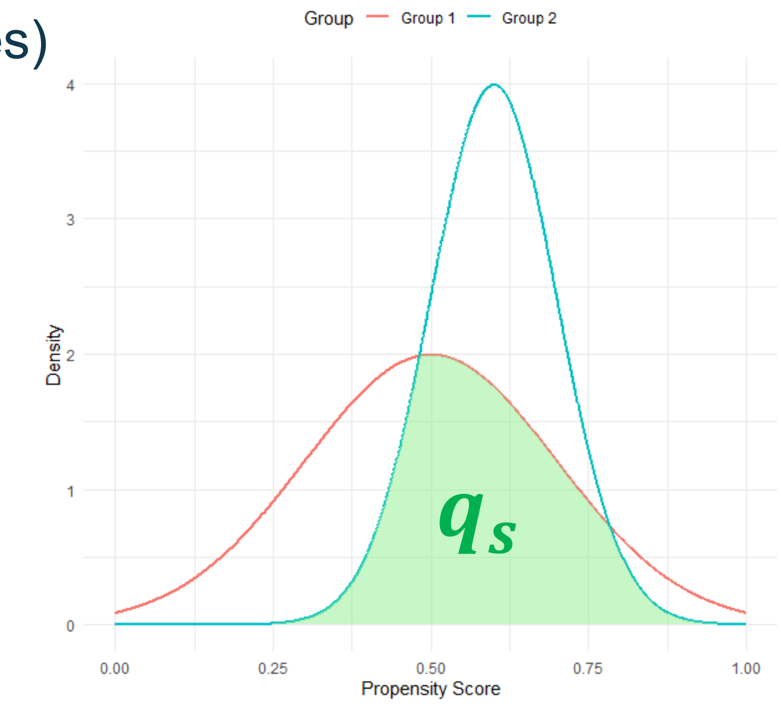
- $\boldsymbol{L(D_1|\theta)}$ is the likelihood of the **current data** given parameter θ
- $\boldsymbol{L(D_0|\theta)}$ is the likelihood of **external data** given parameter θ
- $0 \leq \boldsymbol{\alpha_0} \leq 1$ is a scalar parameter (fixed or random) that controls the weight of influence of D_0 in the analysis of D_1; with higher values of α_0, more information is borrowed from D_0
- $\pi_0(\theta)$ is an initial prior

Propensity Score-Integrated Power Prior (PSIPP)

Wang et al (2019): Adjusts borrowing based on stratified propensity score (PS) similarity

// Stratify current and external subjects into S strata (PS percentiles)

Idea: use overlap of stratified PS distributions of current and external patients to down-weight the external data in the analysis



// q_s: overlapping area of PS distributions of current and external patients

// A : number of subjects to borrow

// $n_{0,s}$: number of external subjects in strata s

// Discount power parameter: $\alpha_s = \min\left(1, \frac{A}{n_{0,s}} \frac{q_s}{\sum_{i=1}^{S} q_s}\right)$

$$q_s = \int_0^1 \min[p_{s,0}(x), p_{s,1}(x)]\, dx$$

// **Power prior**: $\pi(\theta_1, \ldots, \theta_S) \propto \prod_{s=1}^{S} \left[L(D_{s,0}|\theta_s)\right]^{\alpha_s} \pi_0(\theta_1, \ldots, \theta_S)$

Mixture Prior Approach

Schmidli et al. (2014)

To account for the possibility of prior-data conflict and improve the robustness of the inference, fixed-weight mixture prior have been proposed:

$$\pi_{mixture}(\theta) = \tilde{\omega}\pi_{informative}(\theta) + (1 - \tilde{\omega})\pi_{vague}(\theta)$$

Idea: combine and informative prior with a non-informative or vague prior into a single distribution, through a mixing weight

$\pi_{informative}(\theta)$: Informative prior constructed from external data

$\pi_{vague}(\theta)$: Non-informative or vague prior

$\tilde{\omega}$: **prespecified fixed** mixing weight (representing the prior probability of no prior-data conflict between D_1(current) and D_0(external))

Propensity Score Integrated Self Adaptive Mixture Prior

A comprehensive integrated method

Proposed Self-Adaptive Mixture Prior Approach

Define self-adaptive mixture weight

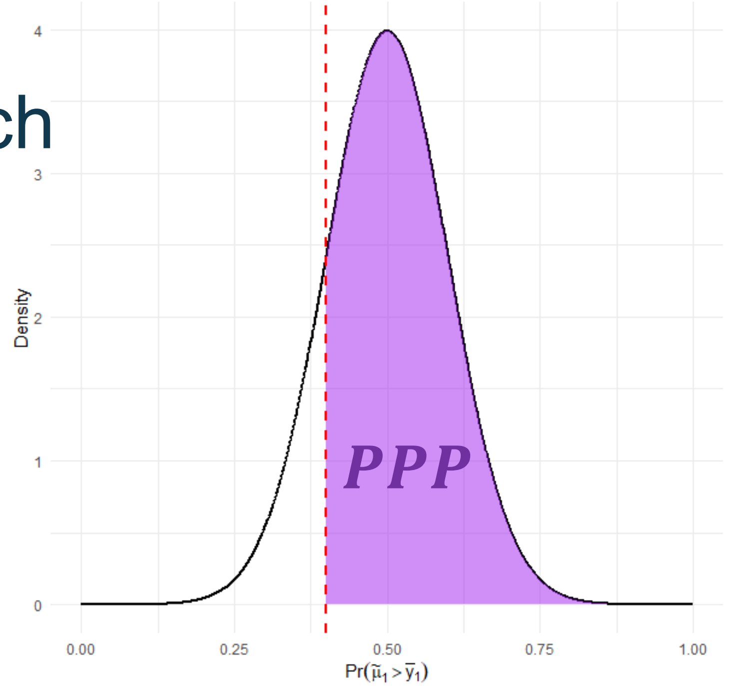
- Assess similarity between current and external data via the **predictive probability of observing the current data given the external data** via the Posterior Predictive Probability:

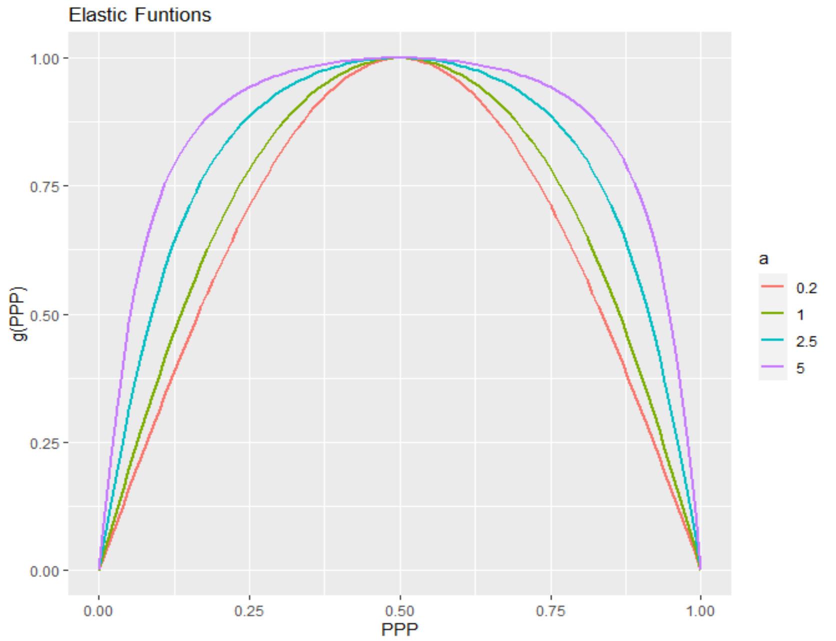
$$PPP = \Pr(\tilde{\mu}_1 > \bar{y}_1 | y_0)$$

  - ➢ $\tilde{\mu}_1$ is the predictive sample mean given external data y_0
  - ➢ $\bar{y}_1$ is the sample mean of the current data

- If external and current data are similar, $PPP \approx 1/2$
- Elastic function $g(PPP) = \dfrac{\arctan(a \times \sin(PPP * \pi)\ )}{\arctan(a)}$
- The value of a can be tuned with clinical input

Self-adaptive mixture weight: $\boldsymbol{\omega} = \boldsymbol{g(PPP)}$





Proposed Self-Adaptive Mixture Prior Approach

The self-adaptive mixture prior for stratum $s = 1,2,3 \ldots, S$, is given by:

$$\pi_{SAMP,s}(\theta) = \boldsymbol{\omega_s}(\boldsymbol{PPP_s})\boldsymbol{\pi_{informative,s}}(\boldsymbol{\theta}) + [1 - \boldsymbol{\omega_s}(\boldsymbol{PPP_s})]\boldsymbol{\pi_{vague,s}}(\boldsymbol{\theta})$$

$\boldsymbol{\pi_{informative,s}}(\boldsymbol{\theta})$: Informative prior for stratum s (derived via PSIPP)

$\boldsymbol{\pi_{vague,s}}(\boldsymbol{\theta})$: Non-informative or vague prior for stratum s

$\boldsymbol{\omega_s}(\boldsymbol{PPP_s})$: Self-adaptive mixture weight for stratum s (depends on PPPs through the elastic function $g(*)$)

Simulation Results

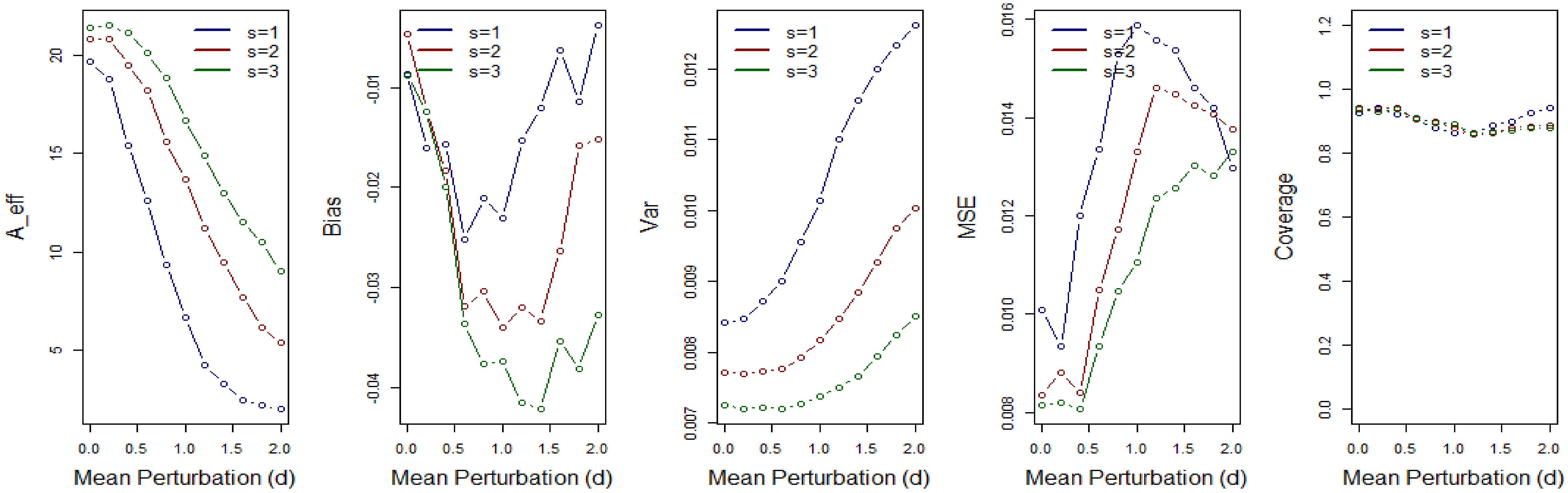
Binary Outcomes

Binary Outcome – Hybrid Control Arm

N1=75 (current), 2:1 treatment-to-control, A=25 (no. to borrow), N0=3000 (external), and s=3 (strata)

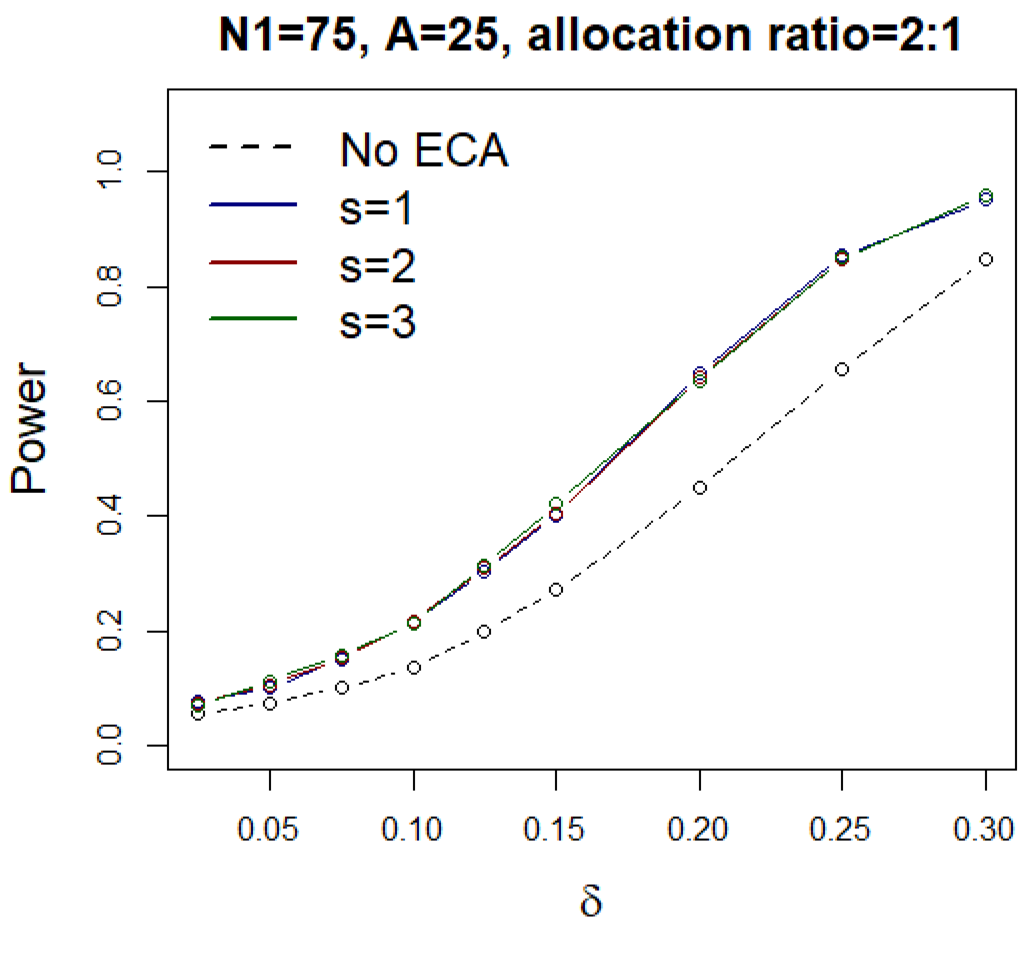
$$logit(Y_i|\boldsymbol{X_i}, Z_i) = \beta_0 + \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{X_i} + O_i\ ;\ \boldsymbol{X} = (X_1, \dots, X_{10})^{\boldsymbol{T}};\ \beta_0 = -10.48\ ;\ \boldsymbol{\beta} = (1, \dots, 1)^T;\ F_{X|Z} = MVN(\mu_z, \Sigma_z)$$

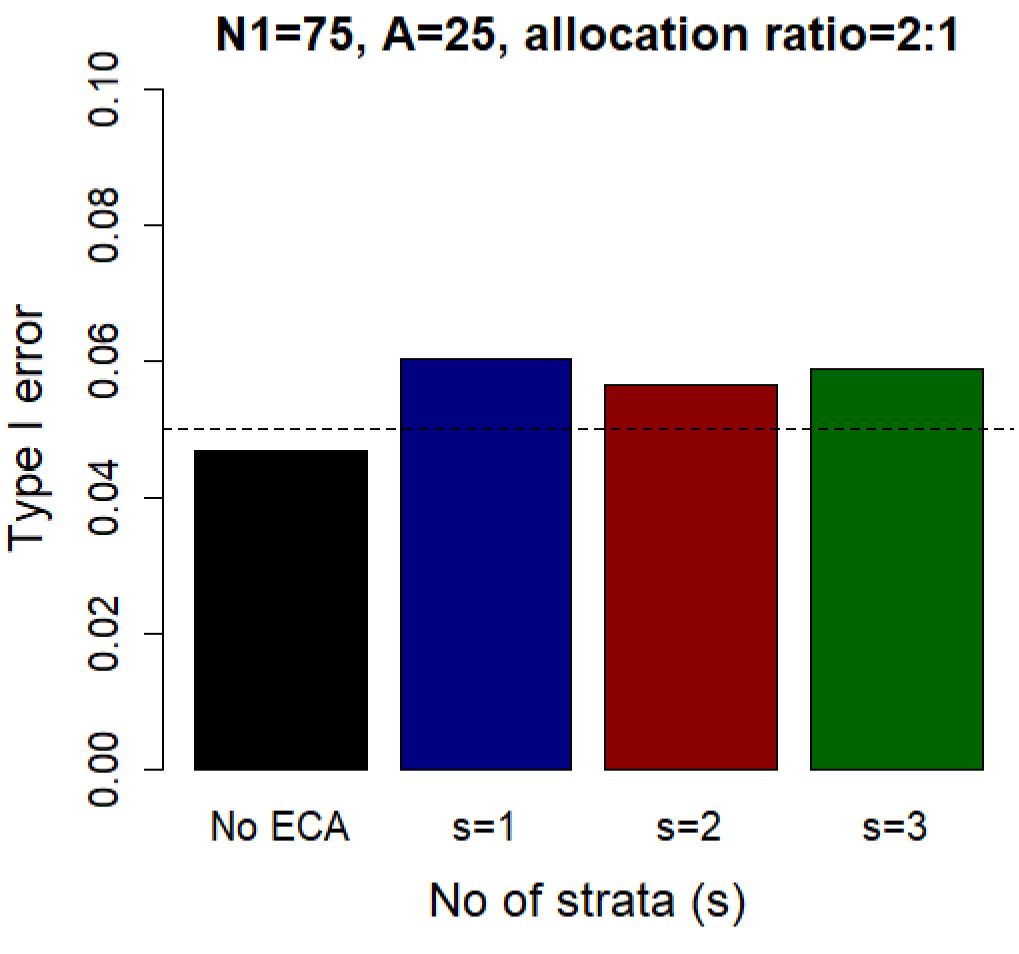
$Z = 1$ (current)	$\mu_1 = (1, \dots, 1)^T$	$\sigma_1^2 = 0.1$	$O_1 = 0$
$Z = 0$ (external)	$\mu_0 = (1.02, \dots, 1.02)^T$	$\sigma_0^2 = 0.12$	$O_0 \sim N(d, 0.25)$ (random perturbation)



Binary Outcome – Hybrid Control Arm

N1=75 (current), 2:1 treatment-to-control, A=25 (no. to borrow), N0=3000 (external), and s=3 (strata)





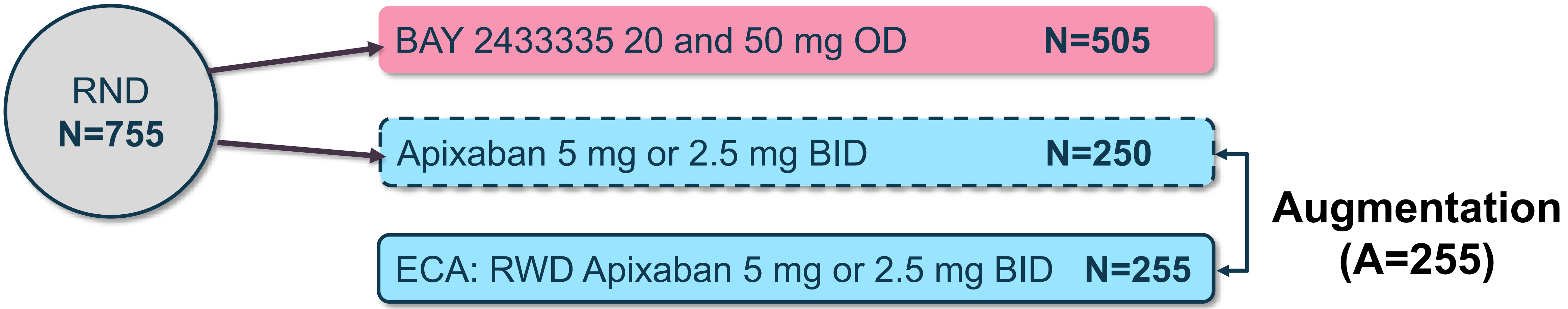
Illustrative Example

Addressing both baseline and outcome differences

Illustrative Example

PACIFIC-AF Phase II trial and External RWD

- **Trial data**: Phase II trial (PACIFIC-AF) to compare efficacy and safety of new factor XIa inhibitor against Apixaban. Two active dose groups and control arm (**N=755, ≈ 2:1 ratio**)
- **External data**: RWD with identical selection criteria as in the trial (N=3327)
- **Outcome**: Major bleeding within 90-days
- **Key Idea**: Augment control arm by borrowing information **equivalent to 255** patients from RWD to attain 1:1 ratio of treatment to control



PS Modeling, Estimation, Trimming

Similarity of baseline data – current vs. external control

- Propensity scores estimated via logistic regression (trial: 755, RWD: 3327)
- 33 baseline covariates as additive linear terms
- Trimming: 3204 of the 3327 external patients were retained (123 excluded)
- Number of strata **s=3**

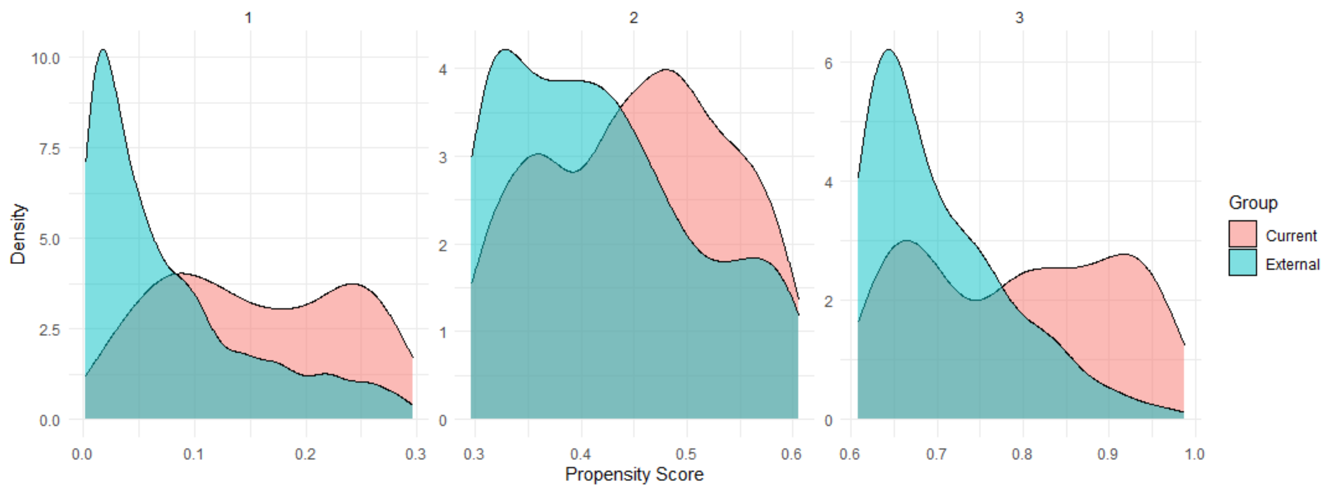


Figure. Propensity score densities by stratum

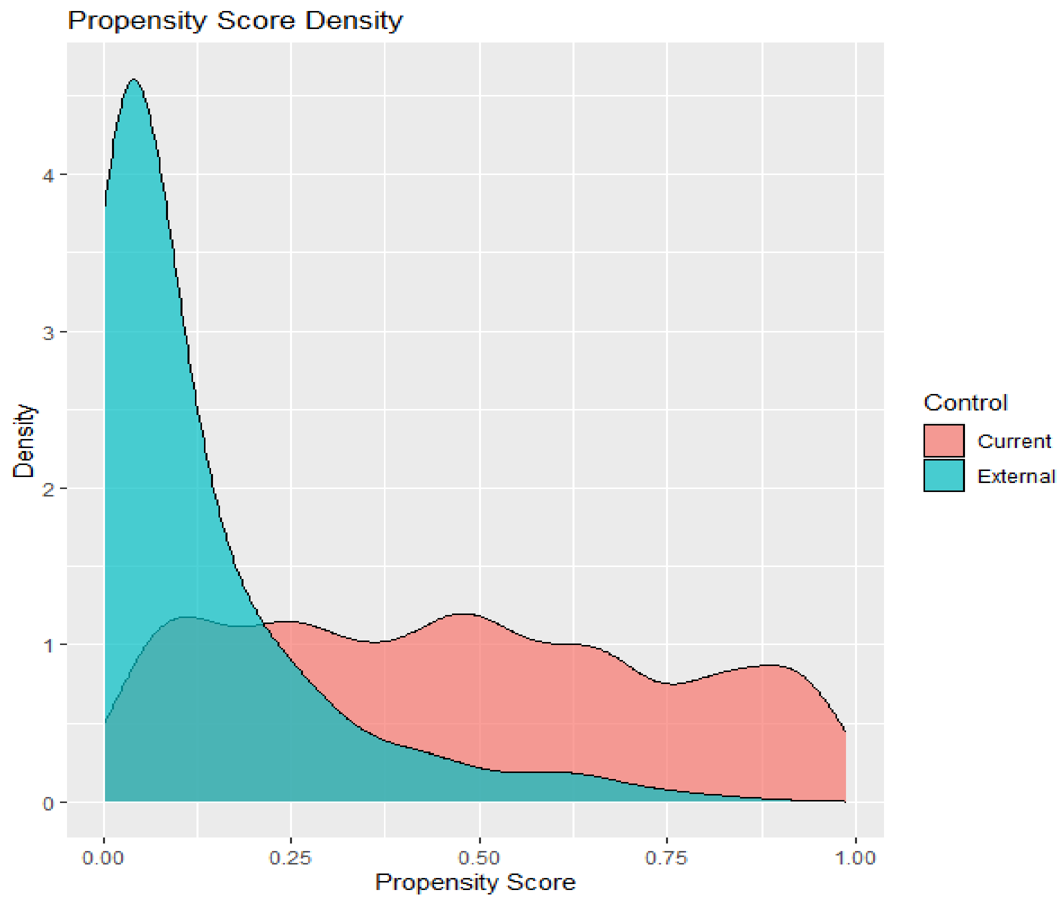


Figure. Propensity score densities for the current study and external data for the illustrative example.

Summary of Current and External Data

PACIFIC-AF Phase II trial and External RWD

Parameter		Stratum			Total
		$s = 1$	$s = 2$	$s = 3$	
Current Study (sample size)	n_{s1}	252	251	252	755
Treatment	n_{s11}	172	172	161	505
Control	n_{s10}	80	79	91	250
Current Study (no. of events)					
Treatment	ne_{s11}	1	1	2	4
Control	ne_{s10}	2	1	3	6
External data (sample size)	n_{s0}	2823	294	87	3204
External data (no. of events)	ne_{s0}	86	8	1	95

Parameters of the self-adaptive mixture prior approach

PACIFIC-AF Phase II trial and External RWD

Parameter		Stratum*			Total
		$s = 1$	$s = 2$	$s = 3$	
Overlapping coefficient	q_s	0.62	0.82	0.64	
Overlapping coefficient normalized	v_s	30%	39%	31%	100%
Discount power parameter	α_s	0.027	0.341	0.904	
Posterior Predictive Probability	PPP_s	0.451	0.626	0.129	
Self-adaptive mixture weight $(a = 0.1)$*	ω_s	0.988	0.923	0.395	
Effective sample size	A_{eff}	77	96	48	**221**

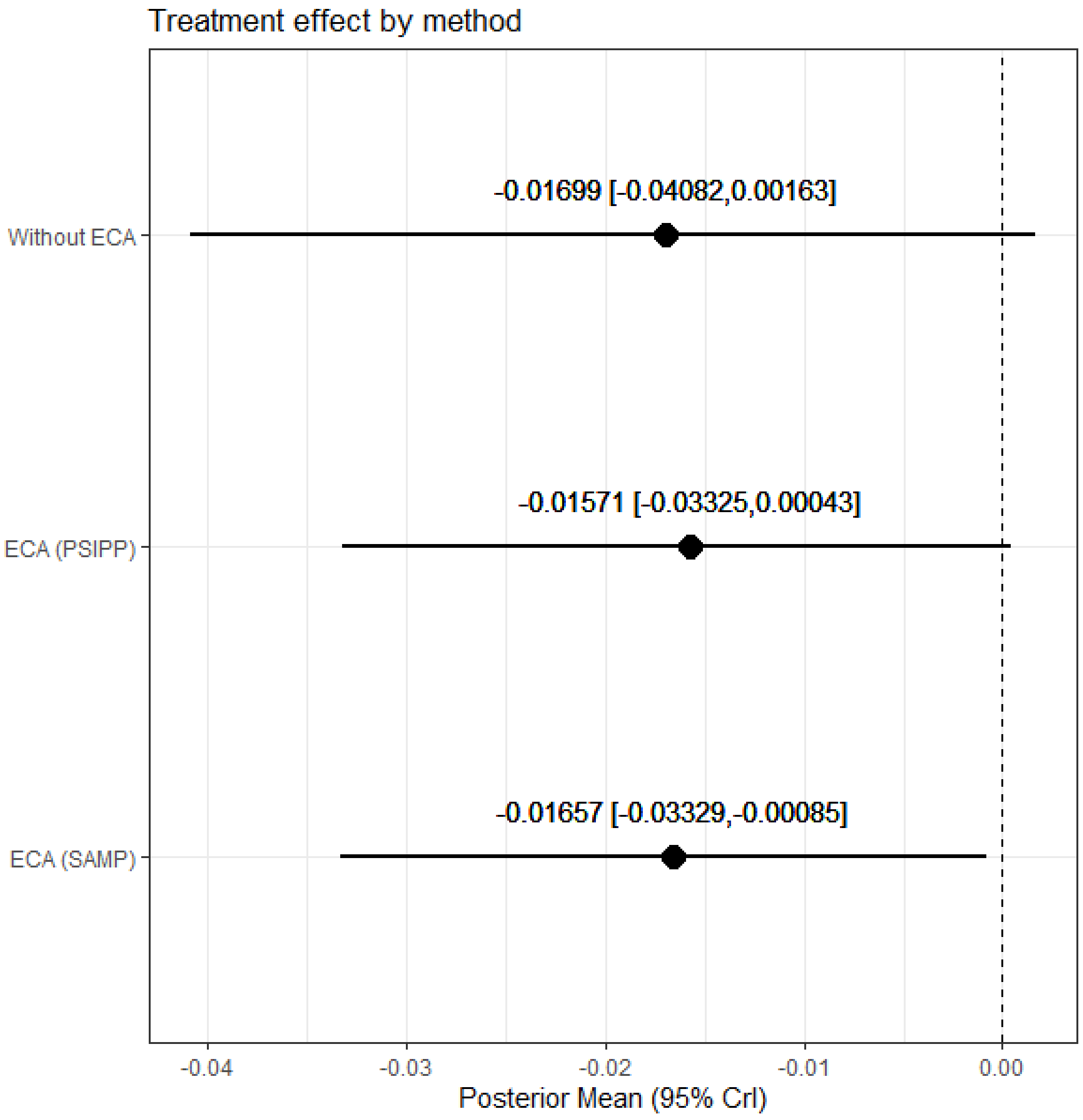
* Number of strata (s) and tuning parameter (a) can be chosen through a grid search aimed at minimizing the MSE

Comparison of Analyses

PACIFIC-AF Phase II trial and External RWD

- Bayesian analysis:
  - Without ECA (Trial data)
  - With ECA (PSIPP)
  - With ECA (PSI-SAMP)

- Adjusting for baseline only, increases precision but slightly shifts the estimate
- Analysis adjusting for baseline & outcome differences attenuates the bias resulted from borrowing from the ECA



Discussion

Summary and Conclusions

Summary and Conclusions

- We propose an integrated self-adaptive Bayesian dynamic borrowing approach
- Baseline differences are accounted for by the overlap of the stratified propensity score distributions
- Self-adaptive mixture weights are function of the predictive posterior probability of observing the current outcome given the outcome of the external data
- The target number of external subjects A is discounted according to the observed degree of difference between current and external data
- Simulations show desirable operating characteristics and adaptive information borrowing
- The proposed approach is intuitive, allows flexibility for adjusting the amount of borrowing, and it enables the incorporation of expert opinion

Key References

- Ibrahim JG and Chen M-H (2000). Power prior distributions for regression models. Stat Sci 15:46-60.
- Ibahim JG et al (2015). The power prior: theory and applications. Stat Med 34:3724-3749. https://doi.org/10.1002/sim.6728
- Jiang L et al (2021). Elastic priors to dynamically borrow information from historical data in clinical trials. Biometrics https://doi.org/10.1111/biom.13551
- Liu M et al (2021). Propensity-score-based meta-analytic predictive prior for incorporating real-world and historical data. Stat Med 40:4794-4808.
- Morita S et al (2008). Determining the effective sample size of a parametric prior. Biometrics 64:595-602. https://doi.org/10.1111/j.1541-0420.2007.00888.x
- Neuenschwander B et al (2020). Predictively consistent prior effective sample sizes. Biometrics 76:578-587.
- Ollier A et al (2020). An adaptive power prior for sequential clinical trials – Application to bridging studies. Stat Meth Med Res 29:2282-2294.
- Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D. and Neuenschwander, B., (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics, 70(4), pp.1023-1032.
- Wang C et al (2019). Propensity score-integrated power prior approach for incorporating real-world evidence in single-arm clinical studies. JBS https://doi.org/10.1080/10543406.2019.1657133

Thank you!

Alfredo Farjat, PhD
Statistical Innovation
alfredo.farjat@bayer.com

Back up

Elastic Prior Approach

Jiang et al (2021)

- **Key idea**: Use of “congruence measure” (e.g. T-test) to assess similarity of current D_1 and external data D_0.
- An “elastic function” $\boldsymbol{g(T)}$ is used to map the congruence measure to the interval (0,1)
- The mapped value is used to adjust the prior variance, controlling how much the external data influences the analysis

- Example: Normal endpoint
  - $y_{1,i} \sim N(\theta, \sigma^2)$, $y_{0,i} \sim N(\theta_0, \sigma_0^2)$, with interest in estimating θ
  - Assuming that D_1 and D_0 are incongruent
  - **Elastic prior**: $\pi^*(\theta|D_0) = N\left(\bar{y}_0, \frac{\sigma_0^2}{n_0 \boldsymbol{g(T)}}\right)$

variance is inflated with the elastic function $\boldsymbol{g(T)}$

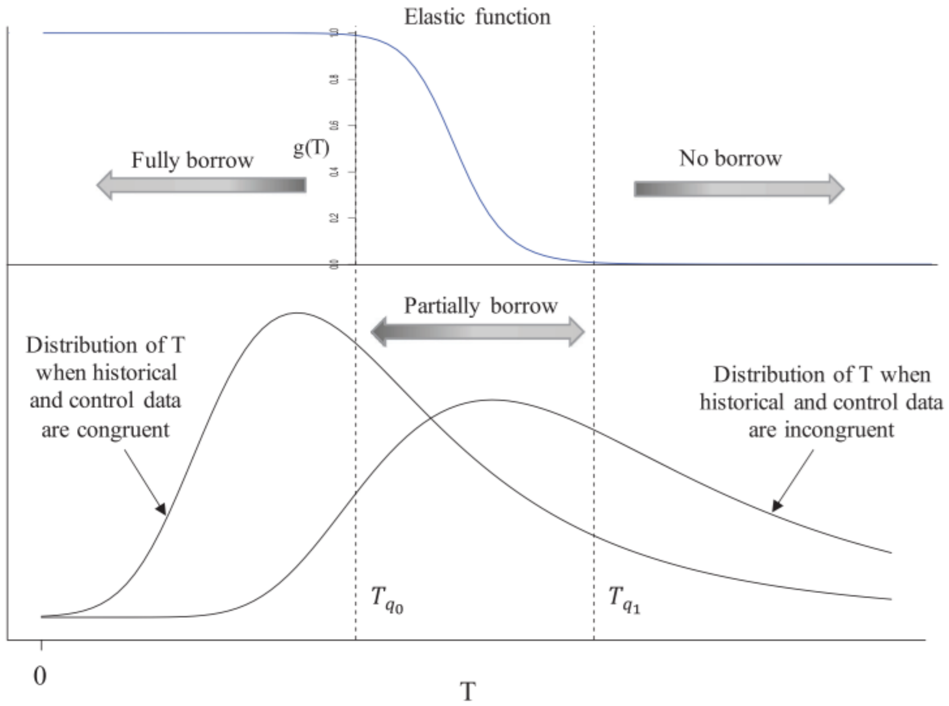


Figure (Jiang et al 2021). Dynamic information borrowing through elastic function

How to choose tuning parameter (a) and number of strata (s)?

Practical recommendations for user-specified parameters

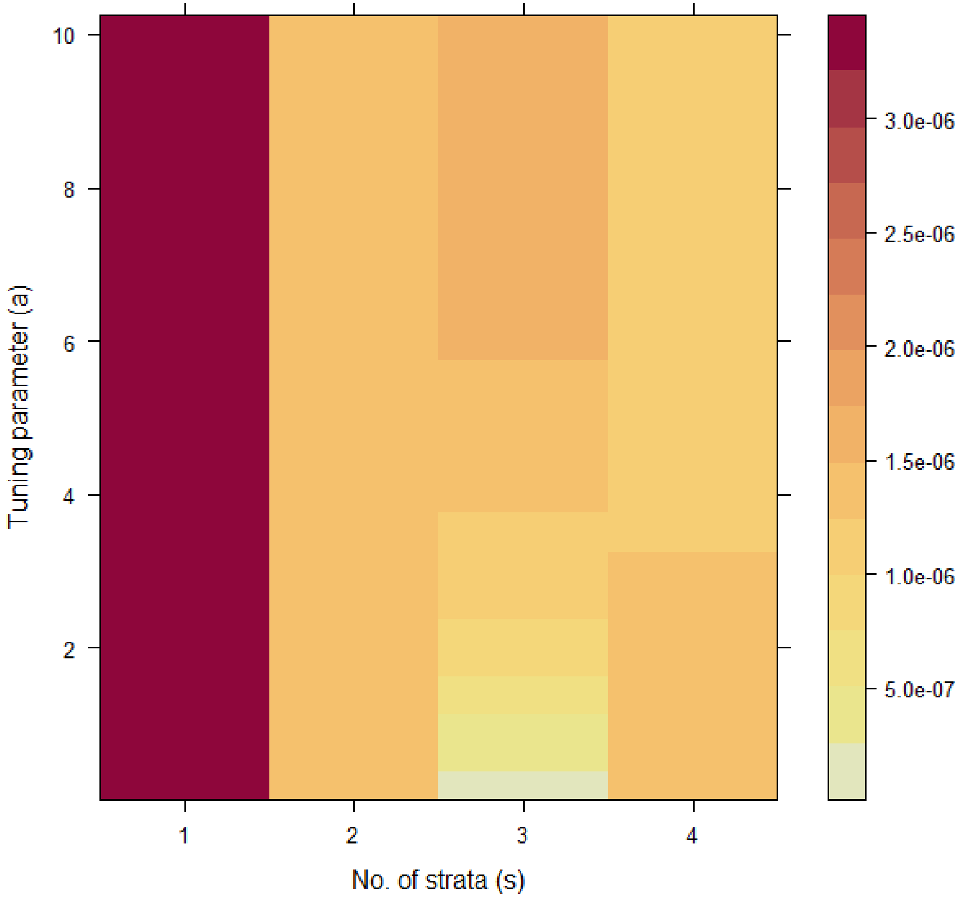
1) Use as reference θ_T (treatment effect from current trial)

2) Define:

$$MSE(s,a) = \left(\hat{\theta}(s,a) - \theta_T\right)^2 + Var\left(\hat{\theta}(s,a)\right)$$

where $\hat{\theta}(s,a)$) is the double adjustment estimator

3) Do a grid search over parameters a and s to minimize MSE



Heatmap of mean squared error *MSE(s,a)*. Minimum is attained at ***a*=0.1** and ***s*=3**

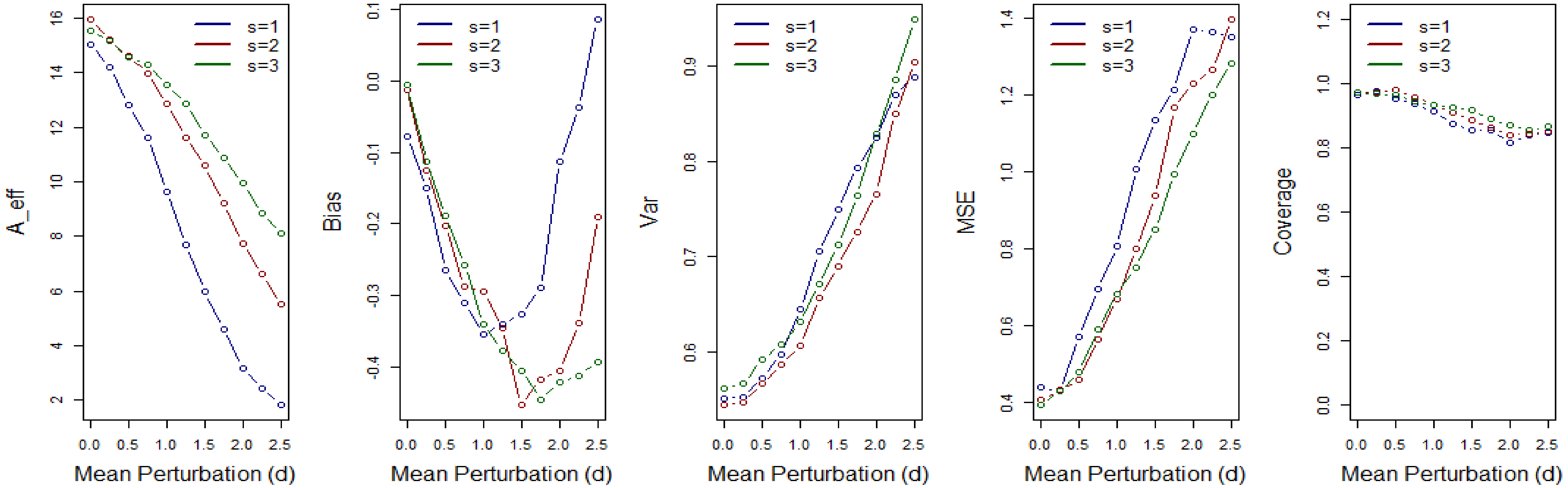
Continuous Outcome – Hybrid Control Arm

Effective SS, bias, variance, MSE and coverage for N1=75 (current), N0=3000 (external), and s=3 (strata)

$Y_i|\boldsymbol{X_i}, Z_i = \beta_0 + \boldsymbol{\beta^T X_i} + \epsilon_i + O_i$; $\boldsymbol{X} = (X_1, \dots, X_{10})^T$; $\beta_0 = 0$; $\boldsymbol{\beta} = (1, \dots, 1)^T$; $\epsilon_i \sim N(0,1)$; $F_{X|Z} = MVN(\mu_z, \Sigma_z)$

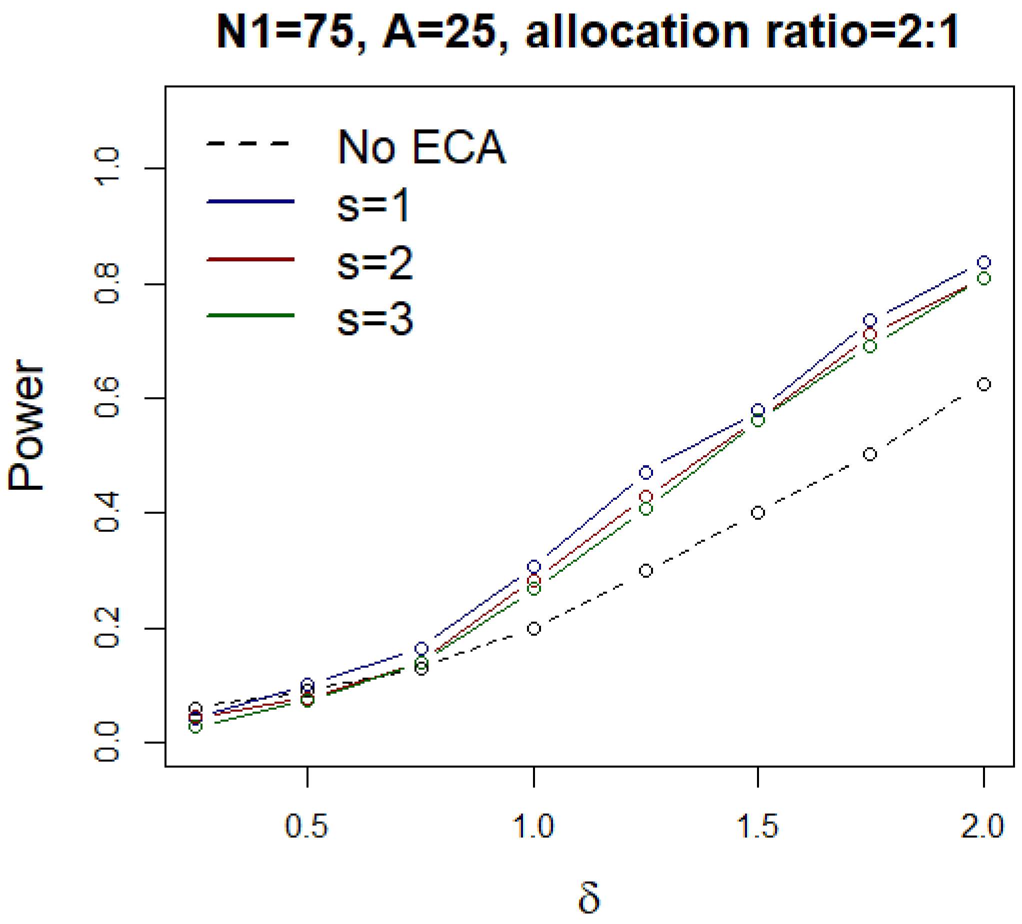
$Z = 1$ (current) $\mu_1 = (1, \dots, 1)^T$; $\sigma_1^2 = 1$; $O_1 = 0$

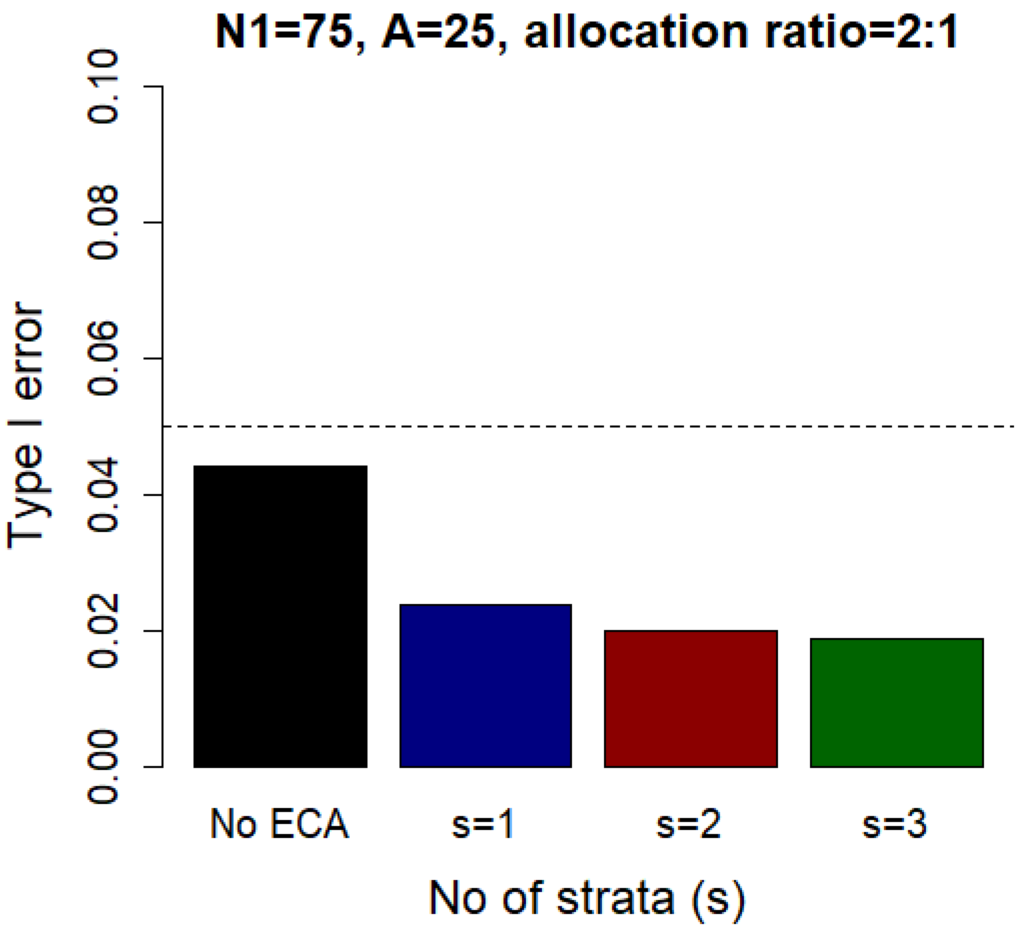
$Z = 0$ (external) $\mu_0 = (1.05, \dots, 1.05)^T$; $\sigma_0^2 = 1.5$; $O_0 \sim N(d, 1.5)$ (random perturbation)



Continuous Outcome – Hybrid Control Arm

Power and Type I error, N1=75, N0=3000, treatment-to-control allocation ratio 2:1, A=25, s=3





Event Rates and Treatment Effect Estimates by Stratum

PACIFIC-AF Phase II trial and External RWD