Improving Power in Group Sequential, Randomized Trials by Adjusting for Prognostic Baseline Variables and Short-Term Outcomes

Tianchen Qian*, Michael Rosenblum, and Huitong Qiu

Department of Biostatistics, Johns Hopkins University

Summary

- **Setting:** delayed response randomized trials
- **Data structure:** baseline variable, short-term outcome, final outcome
- **Question:** what factors affect the precision gain from using adjusted estimators?
- **Answer:** adjusted estimators are particularly useful for estimating ATE when
  - more prognostic baseline variable
  - more prognostic short-term outcome
  - shorter delay in the short-term outcome
  - longer delay in the final outcome
  - less treatment effect heterogeneity

Introduction

In group sequential designs, adjusting for baseline variables and short-term outcomes can lead to increased power and reduced sample size.

We derive formulas for the precision gain (measured by the asymptotic relative efficiency against the unadjusted estimator) from adjusting for baseline variables and short-term outcomes using semiparametric estimators in randomized trials. Key components of the formulas include the variance explained by baseline variables, the residual variance explained by short-term outcomes, the proportion of pipeline participants, and the degree of treatment effect heterogeneity. The formulas can be used in trial planning to predict the potential precision gain from variable adjustment, which will translate to sample size reduction.

When estimating the average treatment effect, the precision gain from adjusting for baseline variables is modified by treatment effect heterogeneity. Given set prognostic value of baseline variables and short-term outcomes within each arm, the precision gain is maximal when there is no treatment effect heterogeneity. In contrast, a purely predictive baseline variable, which only explains treatment effect heterogeneity but is marginally uncorrelated with the outcome, can lead to no precision gain. The precision gain from adjusting for short-term outcomes is not modified by treatment effect heterogeneity.

Definition

For a subject, define the following:

- $\boldsymbol{A}$: binary treatment indicator;
- $\boldsymbol{W}$: baseline variables;
- $\boldsymbol{L}$: short-term outcome, observed with delay d_L;
- $\boldsymbol{Y}$: final outcome, observed with delay d_Y.

At an interim analysis, denote by:

- $\boldsymbol{p_l}$: proportion of enrollees with L observed;
- $\boldsymbol{p_y}$: proportion of enrollees with Y observed.

Define:

- $E_0(\cdot) = E(\cdot \mid A = 0)$, $\mathrm{Var}_0(\cdot) = \mathrm{Var}(\cdot \mid A = 0)$;
- Similarly for $E_1(\cdot)$ and $\mathrm{Var}_1(\cdot)$.

- $\boldsymbol{\theta}$: average treatment effect (ATE):
$$\theta = E(Y|A=1) - E(Y|A=0).$$
- $\boldsymbol{\gamma}$: treatment effect heterogeneity
$$\gamma = \frac{\mathrm{Var}\{E_1(Y \mid W) - E_0(Y \mid W)\}}{\Sigma_{a\in\{0,1\}}\mathrm{Var}_a\{Y\}}.$$
- $\boldsymbol{R^2_W}$: prognostic value in W
$$R^2_W = \frac{\Sigma_{a\in\{0,1\}}\mathrm{Var}_a\{E_a(Y \mid W)\}}{\Sigma_{a\in\{0,1\}}\mathrm{Var}_a(Y)}.$$
- $\boldsymbol{R^2_{L|W}}$: prognostic value in L after adjusting for W
$$R^2_{L|W} = \frac{\Sigma_{a\in\{0,1\}}\mathrm{Var}_a\{E_a(Y \mid L, W) - E_a(Y \mid W)\}}{\Sigma_{a\in\{0,1\}}\mathrm{Var}_a(Y)}.$$

Theoretical Limit on Precision Gain

Assume randomization ($A \perp\!\!\!\perp W$) and independent censoring on L, Y. For estimating θ, the asymptotic relative efficiency (ARE) between the efficient RAL estimator and the unadjusted estimator is

$$\mathrm{ARE}(\text{efficient, unadjusted}) = \frac{1}{(p_y/2)\gamma - [1-(p_y/p_l)]R^2_{L|W} - R^2_W}.$$

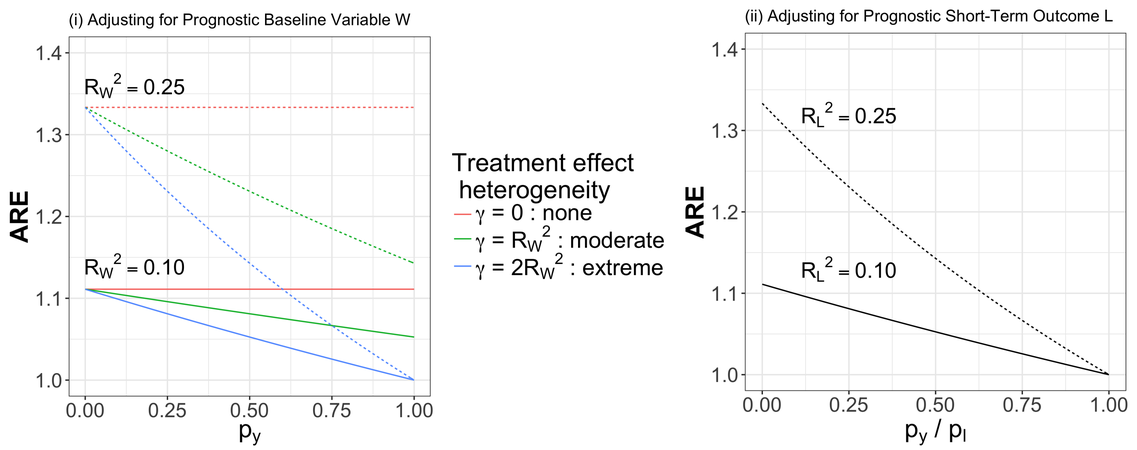
We have the following special cases.

(i) (Only W is prognostic.) If $R^2_{L|W} = 0$, then

$$\mathrm{ARE}(\text{efficient, unadjusted}) = \frac{1}{(p_y/2)\gamma + 1 - R^2_W}.$$

(ii) (Only L is prognostic.) If $R^2_W = 0$ then

$$\mathrm{ARE}(\text{efficient, unadjusted}) = \frac{1}{1 - [1-(p_y/p_l)]R^2_{L|W}}.$$



Conclusion

Precision gain from adjusting for baseline variables depends on:

- prognostic value in W
- proportion of enrollees with Y observed, which is dependent on the delay to observe Y
- degree of treatment effect heterogeneity

Precision gain from adjusting for short-term outcome depends on:

- prognostic value in L
- proportion of enrollees with Y observed, which is dependent on accrual rate and the delay d_Y
- proportion of enrollees with L observed, which is dependent on accrual rate and the delay d_L

Future Research

- Extend to incorporate model misspecification
- Deviation from theory under small sample size
- Measures other than average treatment effect

References

[1] Mark J van der Laan.
Targeted maximum likelihood based causal inference: Part I.
The International Journal of Biostatistics, 6(2), 2010.

[2] Kelly L Moore, Romain Neugebauer, Thamban Valappil, and Mark J van der Laan.
Robust extraction of covariate information to improve estimation efficiency in randomized trials.
Statistics in medicine, 30(19):2389–2408, 2011.

Contact Information

- Email: tqian2@jhu.edu
- Link to this poster:
  `tcqian.wordpress.com/presentation/`
- Link to the corresponding working paper:
  `biostats.bepress.com/jhubiostat/paper285/`