Improving Power through Adjustment for Prognostic Variables in Group Sequential Trials: the impact of baseline variables, short-term outcomes, and treatment effect heterogeneity

Tianchen Qian, Michael Rosenblum, Huitong Qiu

Department of Biostatistics
Johns Hopkins University

JSM 2017, Baltimore

Motivation

- Covariate adjustment is often used to analyze randomized trials.
- Adjusting for baseline variables and short-term outcomes can lead to increased power and reduced sample size, compared to the standard unadjusted estimator.

- **Research Question:** How much precision gain can be obtained by properly adjusting for baseline variables and short-term outcomes? What factors impact this precision gain?

Outline of talk

- A Trial Example
- Notation and Assumption
- Main Result: Formula for Precision Gain
- Simulations
- Summary

A Trial Example: MISTIE-II

- MISTIE Phase II trial for evaluating a surgical treatment for intracerabral hemorrhage.
- **Primary outcome**: modified Rankin Scale (mRS, a score indicating the degree of disability) at 180 days after enrollment.
- **Short-term outcome**: mRS at 30 days after enrollment.
- **Baseline variable**: age, baseline NIHSS (NIH Stroke Scale total score), ICH (intracerebral hemorrhage volume), GCS (Glasgow Coma Scale).

How much can we gain by properly adjusting for baseline variables and short-term outcomes?

Notation

- W: baseline variables
- A: binary treatment indicator
- L: short-term outcome, observed at time d_L after enrollment
- Y: primary outcome, observed at time d_Y after enrollment, with $d_Y > d_L$

Assumption

1. **Randomization**: study arm A is assigned independent of the baseline variable W .
2. **Independent censoring**: missingness of L and Y is independent of (W, A) and the potential outcomes of (L, Y).
   (This holds if administrative censoring is the only source of missingness.)
3. **Monotone censoring**: for a participant, if L is missing, then Y is also missing.
4. For simplicity, in this talk we assume $P(A = 1) = 0.5$.

Definition

- We consider testing for positive **average treatment effect**: $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$, with θ defined as:

$$\theta := E(Y \mid A = 1) - E(Y \mid A = 0).$$

- The **unadjusted estimator** is defined as the difference in sample means of the observed primary outcome Y between the two arms.
- An estimator $\hat{\theta}$ of the parameter θ is consistent and asymptotically normal if $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2)$, where $\xrightarrow{d}$ denotes convergence in distribution.
- The **asymptotic variance** of $\hat{\theta}$ is $\text{AVar}(\hat{\theta}) := \sigma^2$.

- The **asymptotic relative efficiency** (ARE) between two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ is the inverse of the ratio of their asymptotic variances:

$$\text{ARE}(\hat{\theta}_1, \hat{\theta}_2) := \frac{\text{AVar}(\hat{\theta}_2)}{\text{AVar}(\hat{\theta}_1)}.$$

- Asymptotically it equals the inverse of the ratio of the sample sizes required for the two estimators to achieve the same power.
- For example, $\text{ARE}(\hat{\theta}_1, \hat{\theta}_2) = 1.2$ means that if $\hat{\theta}_1$ requires 1000 sample size to achieve a certain power, then $\hat{\theta}_2$ requires 1200 sample size to achieve the same power.

Characterizing Prognostic Value in W and L

- $E_a(\cdot) := E(\cdot \mid A = a)$ and $\mathrm{Var}_a(\cdot) := \mathrm{Var}(\cdot \mid A = a)$ for $a \in \{0, 1\}$.
- The **proportion of variance in Y explained by W**:

$$R^2_W := \frac{\sum_{a\in\{0,1\}} \mathrm{Var}_a\{E_a(Y \mid W)\}}{\sum_{a\in\{0,1\}} \mathrm{Var}_a(Y)} \quad (= 0.28 \text{ for MISTIE})$$

- The **proportion of variance in Y explained by L after accounting for W** as

$$R^2_{L|W} := \frac{\sum_{a\in\{0,1\}} \mathrm{Var}_a\{E_a(Y \mid L, W) - E_a(Y \mid W)\}}{\sum_{a\in\{0,1\}} \mathrm{Var}_a(Y)} \quad (= 0.11)$$

- The **treatment effect heterogeneity**

$$\gamma = \frac{\mathrm{Var}\{E_1(Y \mid W) - E_0(Y \mid W)\}}{\sum_{a\in\{0,1\}} \mathrm{Var}_a\{Y\}} \quad (= 0.06),$$

with $0 \le \gamma \le 2R^2_W$.

Main Result: Formula for Precision Gain

- Denote by p_l the probability for a participant to have L observed, and p_y the probability to have Y observed (at analysis time).
  - For example, if a trial enrolls 1000 participants, with 800 having L observed and 600 having Y observed, then $p_l = 0.8$ and $p_y = 0.6$.
- The following extends results of Moore and van der Laan (2009).

Theorem

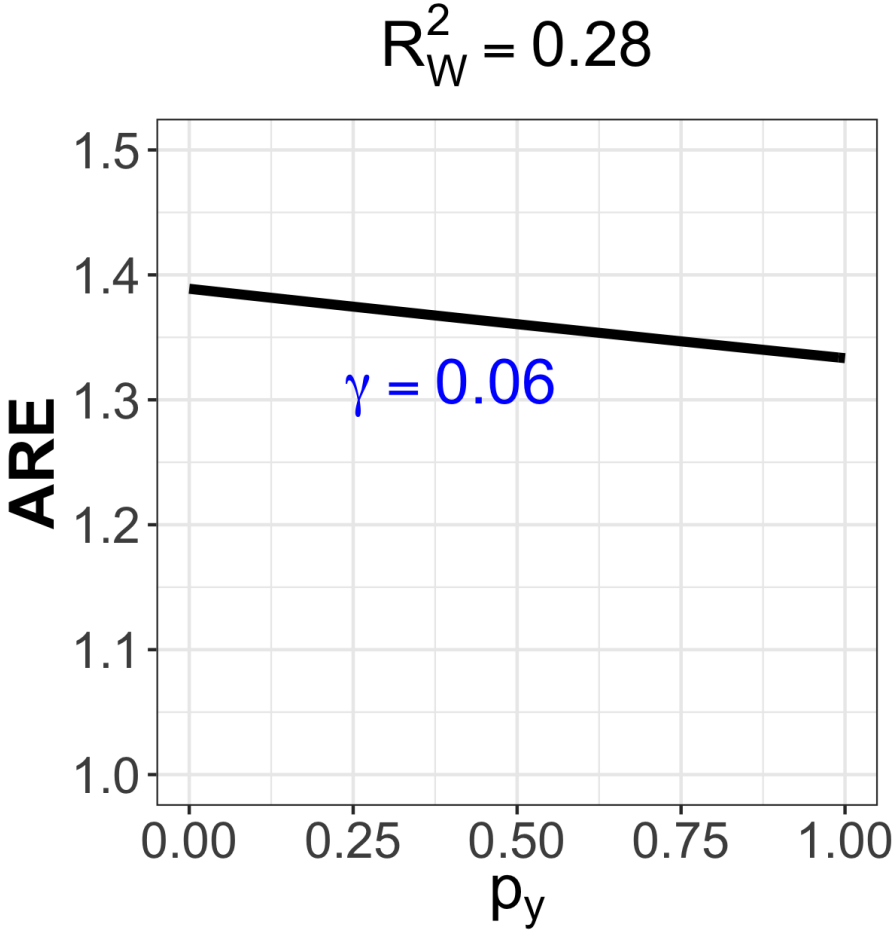
The asymptotic relative efficiency between an efficient RAL estimator and the unadjusted estimator when estimating θ is

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

- **Practical guide:** when using locally efficient estimators [e.g., Targeted Maximum Likelihood Estimator (van der Laan and Gruber, 2012)], we can estimate R^2_W and $R^2_{L|W}$ using model fits, and the formula still holds approximately.

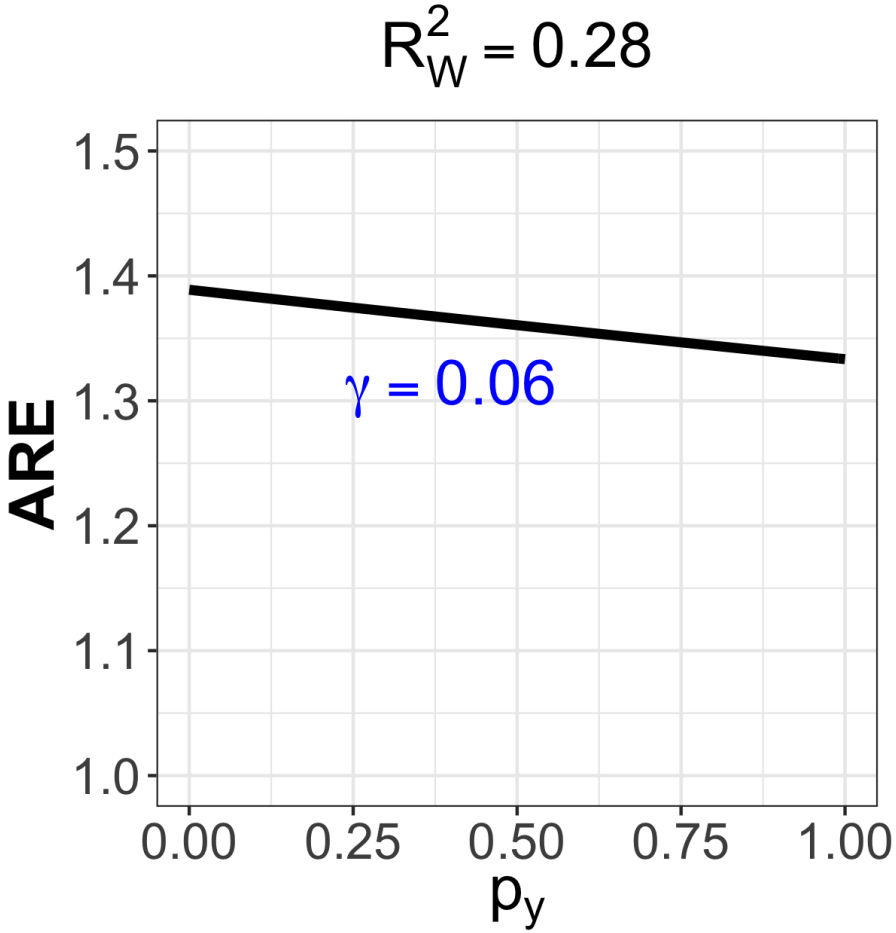
Precision Gain from Adjustment When Only Baseline Variable W is Prognostic:

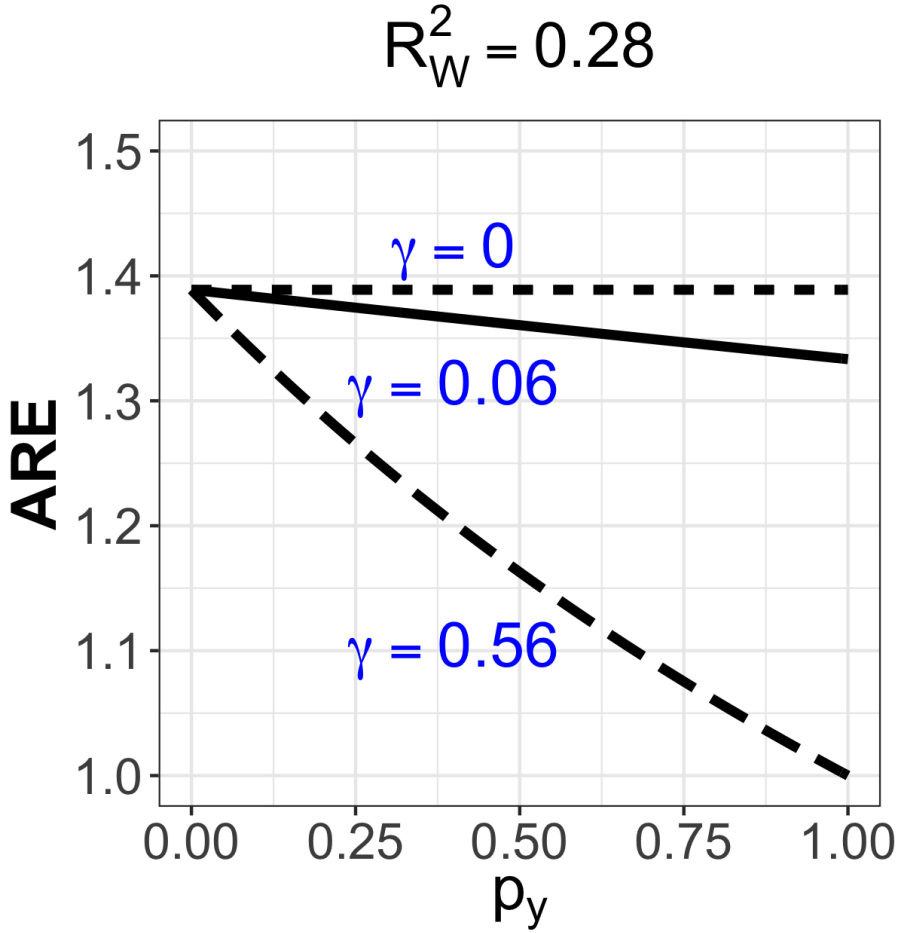
$$\text{ARE}(\text{efficient}, \text{unadjusted}) = \{1 + (p_y/2)\gamma - R_W^2\}^{-1}$$



Precision Gain from Adjustment When Only Baseline Variable W is Prognostic:

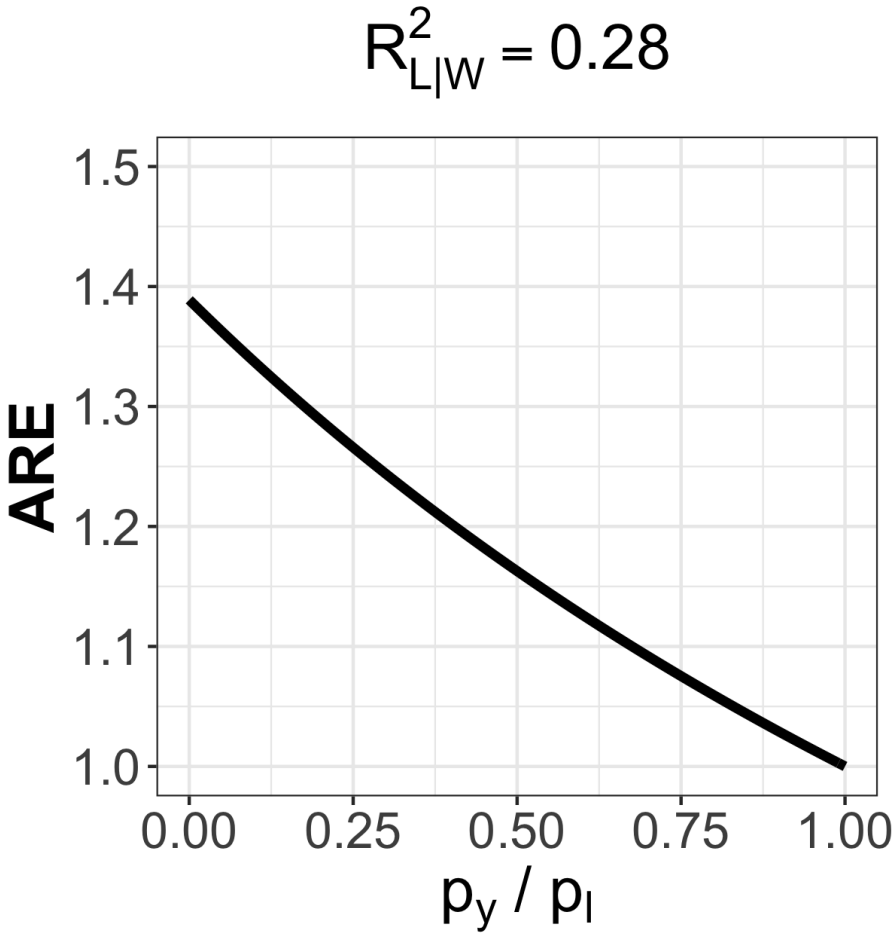
$$\text{ARE}(\text{efficient}, \text{unadjusted}) = \{1 + (p_y/2)\gamma - R_W^2\}^{-1}$$





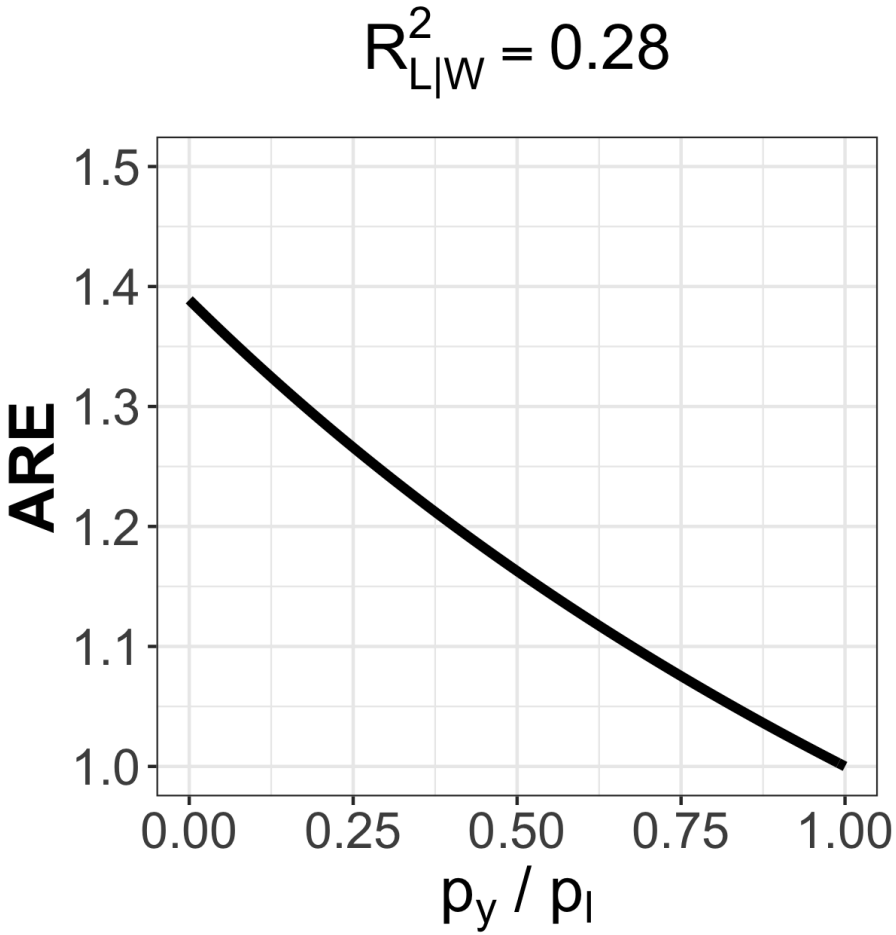
Precision Gain from Adjustment When Only Short-term Outcome L is Prognostic:

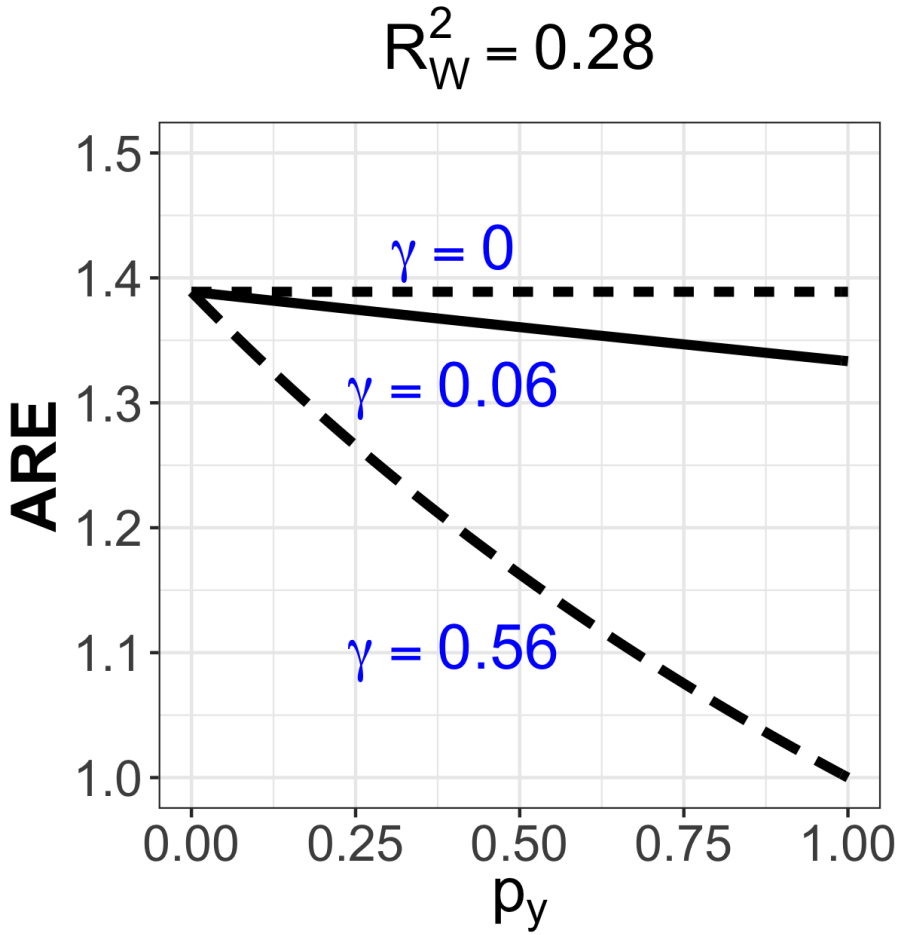
$$\text{ARE}(\text{efficient}, \text{unadjusted}) = \{1 - (1 - p_y/p_l)R^2_{L|W}\}^{-1}$$



Precision Gain from Adjustment When Only Short-term Outcome L is Prognostic:

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





Simulation Setup

- Patient data generate from MISTIE-II data set.
- Use TMLE as the adjusted estimator. As an alternative, one could also use the ANCOVA estimator.
- Use group sequential designs from Hampson and Jennison (2013).
- Consider four prognostic settings:
  - $\text{progn}_{W,L}$: both W and L are prognostic
  - progn_{W}: only W is prognostic
  - progn_{L}: only L is prognostic
  - $\text{progn}_{\emptyset}$: none of W, L is prognostic

Simulation Result

- In the simulations, we assume we know R^2_W and $R^2_{L|W}$, and set analysis timing to get 80% power for each setting.

Estimator	Progn. set.	$n_{\max}$	Type I error	Power	ESS H_0	ESS H_1
unadjusted	all	480	0.025	0.81	318	383
adjusted	$\text{progn}_{W,L}$	300	0.025	0.80	224	258
	progn_{W}	300	0.025	0.80	228	261
	progn_{L}	480	0.025	0.80	309	374
	$\text{progn}_{\emptyset}$	480	0.025	0.81	321	384

Table: The maximum sample size ($n_{\max}$), Type I error, Power, and expected sample size (ESS) under H_0 and H_1 for each estimator to achieve 80% power under each prognostic setting (Progn. set.). Columns (except for $n_{\max}$) are averaged from 10,000 simulated trials.

Summary

- We derived formulas for the precision gain from adjusting for baseline variables and short-term outcomes.
- The value added from adjusting for a prognostic short-term outcome can be much less than an equally prognostic baseline variable.
- Impact of treatment effect heterogeneity is complex:
  - Everything else held equal, more heterogeneity decreases the value added from covariate adjustment;
  - In the case of maximal heterogeneity, no precision gain from adjustment if all Y observed;
  - However, pipeline W contributes to precision gain only when there is heterogeneity.

Reference

This talk is based on:

- Qian T., Rosenblum M., *and* Qiu H. Improving Power through Adjustment for Prognostic Variables in Group Sequential Trial Designs: Impact of Baseline Variables, Short-Term Outcomes, and Treatment Effect Heterogeneity. Manuscript in progress.

Related work:

- Hampson, L. V. *and* Jennison, C. (2013). Group sequential tests for delayed responses (with discussion). *JRSSB*, 75, 354.
- Hanley, D. F., Thompson, R. E., Muschelli, J., Rosenblum, M., McBee, N., Lane, K., Bistran- Hall, A. J., Mayo, S. W., et al. (2016). Safety and efficacy of minimally invasive surgery plus recombinant tissue plasminogen activator in intracerebral haemorrhage evacuation (MISTIE): a randomised phase 2 trial. *Lancet Neurology*.
- Moore K. L. *and* van der Laan M.J. (2009) Covariate adjustment in randomized trials with binary outcomes: Targeted maximum likelihood estimation. *Statistics in medicine*, 28, 3964.
- Rosenblum, M. *and* van der Laan, M. J. (2011). Simple examples of estimating causal effects using targeted maximum likelihood estimation. Unpublished paper, available at `http://works.bepress.com/mark_van_der_laan/223`.
- van der Laan, M. J. and Gruber, S. (2012). Targeted minimum loss based estimation of causal effects of multiple time point interventions. *The International Journal of Biostatistics*, 8.

Thank you!

tqian2@jhu.edu

This work was supported by the Patient-Centered Outcomes Research Institute (ME-1306- 03198) and the US Food and Drug Administration (HHSF223201400113C); this work is solely the responsibility of the authors and does not represent the views of the above agencies.
We thank Dr. Dan Hanley for providing the data.