Modified G-estimation for Repeated Outcome Measures

Wei (Peter) Yang and Marshall Joffe

Department of Biostatistics and Epidemiology
University of Pennsylvania Perelman School of Medicine

May 24, 2012

Motivating Example

- To estimate the effect of Erythropoietin (EPO) dose on hematocrit level in incidence dialysis patients using the United States Renal Data System (USRDS) data.
- EPO is a glycoprotein hormone that controls red blood cell production and is often prescribed to treat anemia in dialysis patients.
- USRDS is the national data registry on the end-stage renal disease (ESRD) population in the U.S.
- USRDS is a claims database. Some key confounding variables, e.g., lab values other than hematocrit level, were not available.
- The data has relatively long follow-up time compared to the life span of red blood cells (100-120 days).

Notation

- $\overline{L}_K = \{L_1, \cdots, L_K\}$: time-updated covariates measures
- $\overline{A}_K = \{A_1, \cdots, A_K\}$: repeated treatment measures
- $\overline{Y}_K = \{Y_1, \cdots, Y_K\}$: repeated outcome measures
- The time ordering of these variables is L_t, A_t, Y_t.
- We assume that L_t and A_t are measured at the beginning of the time interval t and Y_t is measured at the end of the time interval t.

Notation (cont'd)

- $Y_t^{\overline{A}_s,0}, t = 1, \cdots, K$: potential outcomes if subjects receive the same treatment as was observed through time s and do not receive any treatment afterward.
- $Y_t^0, t = 1, \cdots, K$: potential outcomes if subjects do not receive any treatment.

The Blip Down Process

Observed outcomes $Y_t, t = 1, \cdots, K$

$t = K$							Y_K
$t = K - 1$						Y_{K-1}	
$t = K - 2$					Y_{K-2}		
$\vdots$				$\cdots$			
$t = 2$			Y_2				
$t = 1$		Y_1					
$t = 0$							

The Blip Down Process

$t = K$						Y_K
$t = K - 1$					Y_{K-1}	$Y_K^{\overline{A}_{K-1},0}$
$t = K - 2$				Y_{K-2}		
$\vdots$			$\cdots$			
$t = 2$		Y_2				
$t = 1$	Y_1					
$t = 0$						

- Start with Y_K and remove the effect of A_K on Y_K to get $Y_K^{\overline{A}_{K-1},0}$

The Blip Down Process

$t = K$						Y_K
$t = K-1$					Y_{K-1}	$Y_K^{\overline{A}_{K-1},0}$
$t = K-2$				Y_{K-2}	$Y_{K-1}^{\overline{A}_{K-2},0}$	$Y_K^{\overline{A}_{K-2},0}$
$\vdots$			$\dots$			
$t = 2$		Y_2				
$t = 1$	Y_1					
$t = 0$						

- Remove the effect of A_{K-1} on Y_{K-1} to get $Y_{K-1}^{\overline{A}_{K-2},0}$
- Remove the effect of A_{K-1} on $Y_K^{\overline{A}_{K-1},0}$ to get $Y_K^{\overline{A}_{K-2},0}$

The Blip Down Process

$t=K$						Y_K
$t=K-1$					Y_{K-1}	$Y_K^{\overline{A}_{K-1},0}$
$t=K-2$				Y_{K-2}	$Y_{K-1}^{\overline{A}_{K-2},0}$	$Y_K^{\overline{A}_{K-2},0}$
$\vdots$				$\vdots$	$\vdots$	$\vdots$
$t=2$		Y_2	$\cdots$	$Y_{K-2}^{\overline{A}_2,0}$	$Y_{K-1}^{\overline{A}_2,0}$	$Y_K^{\overline{A}_2,0}$
$t=1$	Y_1	$Y_2^{A_1,0}$	$\cdots$	$Y_{K-2}^{A_1,0}$	$Y_{K-1}^{A_1,0}$	$Y_K^{A_1,0}$
$t=0$	Y_1^0	Y_2^0	$\cdots$	Y_{K-2}^0	Y_{K-1}^0	Y_K^0

- Remove the effect of $\overline{A}_K$ on $\overline{Y}_K$ to get $\overline{Y}^{\overline{0}}$.
- Similar ideas available for non-rank preserving structural nested distribution models and structural nested mean models.

The Blip Down Process (cont'd)

For example, when $K = 6$:

$t = 6$						Y_6
$t = 5$					Y_5	$Y_6^{\overline{A}_5,0}(\psi)$
$t = 4$				Y_4	$Y_5^{\overline{A}_4,0}(\psi)$	$Y_6^{\overline{A}_4,0}(\psi)$
$t = 3$			Y_3	$Y_4^{\overline{A}_3,0}(\psi)$	$Y_5^{\overline{A}_3,0}(\psi)$	$Y_6^{\overline{A}_3,0}(\psi)$
$t = 2$		Y_2	$Y_3^{\overline{A}_2,0}(\psi)$	$Y_4^{\overline{A}_2,0}(\psi)$	$Y_5^{\overline{A}_2,0}(\psi)$	$Y_6^{\overline{A}_2,0}(\psi)$
$t = 1$	Y_1	$Y_2^{A_1,0}(\psi)$	$Y_3^{A_1,0}(\psi)$	$Y_4^{A_1,0}(\psi)$	$Y_5^{A_1,0}(\psi)$	$Y_6^{A_1,0}(\psi)$
$t = 0$	$Y_1^0(\psi)$	$Y_2^0(\psi)$	$Y_3^0(\psi)$	$Y_4^0(\psi)$	$Y_5^0(\psi)$	$Y_6^0(\psi)$

- SNMs used to parametrize the blip-down process. ψ represents the finite dimensional causal parameter.

The Likelihood Function

Using fully blipped down potential outcomes, the joint density of the data can be written as:

$$
\begin{aligned}
& f(\overline{Y}_K, \overline{A}_K, \overline{L}_K) \\
& = \frac{\partial \overline{Y}_K^0}{\partial \overline{Y}_K} \times f(\overline{Y}_K^0, \overline{A}_K, \overline{L}_K) \\
& = f(\overline{Y}_K^0) \times \prod_{t=1}^{K} \left\{ f(L_t | \overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}_t^{\overline{A}_{t-1},0}) \times f(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}_t^{\overline{A}_{t-1},0}) \times \frac{\partial \underline{Y}_t^{\overline{A}_{t-1},0}}{\partial \underline{Y}_t^{\overline{A}_t,0}} \right\}
\end{aligned}
$$

where $\underline{Y}_t^{\overline{A}_{t-1},0} = \left\{ Y_t^{\overline{A}_{t-1},0}, Y_{t+1}^{\overline{A}_{t-1},0}, \cdots, Y_K^{\overline{A}_{t-1},0} \right\}$

- Start with the whole vector of potential outcomes $\overline{Y}_K^0$
- For $t = 1, \cdots, K$
    - Generate L_t
    - Generate A_t
    - Blip up all potential outcomes adding the effect of A_t

Sequential Ignorability Assumption

- $A_t \coprod \underline{Y}_t^{\overline{A}_{t-1},0} | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, t = 1, \cdots, K$
- Under sequential ignorability assumptions, the likelihood function is:

$$f(\overline{Y}_K, \overline{A}_K, \overline{L}_K) \\ = f(\overline{Y}_K^0) \times \prod_{t=1}^{K} \left\{ f(L_t | \overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}_t^{\overline{A}_{t-1},0}) \times f(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}) \times \frac{\partial \underline{Y}_t^{\overline{A}_{t-1},0}}{\partial \underline{Y}_t^{\overline{A}_t,0}} \right\}$$

An Estimating Equation

A choice of the estimating equation is:

$$S_\psi = \sum_{t=1}^{K} \left[\left\{ A_t - E\left(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}\right) \right\} \left(\sum_{m=t}^{K} Y_m^{\overline{A}_{t-1},0}(\psi) \right) \right]$$

where $E\left(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}\right)$ is the propensity score.

Partial Blip Down Process

- Motivation
  - Specification of the causal model is a concern especially when blipping down too many periods.
  - Scientifically may be only interested in the effect of treatments in short periods.
- Rather than defining all potential outcomes, only a subset of potential outcomes are modeled, i.e., $\left\{ Y_t^{\overline{A}_{t-1},0}, \cdots, Y_t^{\overline{A}_{t-\delta},0} \right\}$, $t = 1, \cdots, K$, where δ is the number of blip down periods.
- Similar to the idea of history-adjusted marginal structural models.
- Use finite dimensional parameter ψ for potential outcomes within δ periods; infinite dimensional parameter beyond that.

Partial Blip Down Process (cont'd)

For example, when $K = 6, \delta = 3$:

$t = 6$						Y_6
$t = 5$					Y_5	$Y_6^{\overline{A}_5,0}(\psi)$
$t = 4$				Y_4	$Y_5^{\overline{A}_4,0}(\psi)$	$Y_6^{\overline{A}_4,0}(\psi)$
$t = 3$			Y_3	$Y_4^{\overline{A}_3,0}(\psi)$	$Y_5^{\overline{A}_3,0}(\psi)$	$Y_6^{\overline{A}_3,0}(\psi)$
$t = 2$		Y_2	$Y_3^{\overline{A}_2,0}(\psi)$	$Y_4^{\overline{A}_2,0}(\psi)$	$Y_5^{\overline{A}_2,0}(\psi)$	$Y_6^{\overline{A}_2,0}$
$t = 1$	Y_1	$Y_2^{A_1,0}(\psi)$	$Y_3^{A_1,0}(\psi)$	$Y_4^{A_1,0}(\psi)$	$Y_5^{A_1,0}$	$Y_6^{A_1,0}$
$t = 0$	$Y_1^0(\psi)$	$Y_2^0(\psi)$	$Y_3^0(\psi)$	Y_4^0	Y_5^0	Y_6^0

- Models for potential outcomes in red are parametrized with finite dimensional parameter ψ.
- Models for potential outcomes in gray are left unspecified.

Fully Vs. Partially Blipped Down Potential Outcomes

- Partially blipped down potential outcomes make less assumption about the structural models.
- Examples: consider two models with identical parametrization
  - Model for fully blipped down potential outcomes: $Y^{0}_{t+\delta} = Y_{t+\delta} - \left(\sum_{j=t+1}^{t+\delta} A_j\right)\psi$
  - Model for partially blipped down potential outcomes: $Y^{\overline{A}_t,0}_{t+\delta} = Y_{t+\delta} - \left(\sum_{j=t+1}^{t+\delta} A_j\right)\psi$
- First model implicitly assumes that treatment before time t, i.e., $\overline{A}_t$, has no direct effect on outcome $Y_{t+\delta}$.
- Second model does not make any restrictions on the effect of $\overline{A}_t$ on $Y_{t+\delta}$.

The Revised Likelihood Function

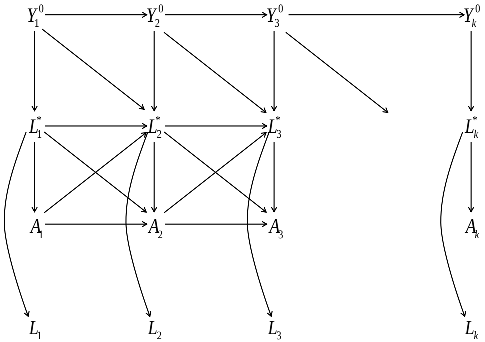
$$
\begin{aligned}
& f(\overline{Y}_K, \overline{A}_K, \overline{L}_K) \\
&= \left\{ f(\overline{Y}^0_\delta) \times f(L_1|\overline{Y}^0_\delta) \times f(A_1|L_1, \overline{Y}^0_\delta) \times \frac{\partial \overline{Y}^0_\delta}{\partial \overline{Y}^{A_1,0}_\delta} \right\} \\
&\times \prod_{t=2}^{K-\delta+1} \left\{ \begin{array}{l} f(Y^{\overline{A}_{t-1},0}_{t+\delta-1}|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}) \\ \times f(L_t|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)}) \times f(A_t|\overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)}) \times \frac{\partial Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)}}{\partial Y^{\overline{A}_t,0}_{t:(t+\delta-1)}} \end{array} \right\} \\
&\times \prod_{t=K-\delta+2}^{K} \left\{ f(L_t|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}^{\overline{A}_{t-1},0}_t) \times f(A_t|\overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}^{\overline{A}_{t-1},0}_t) \times \frac{\partial \underline{Y}^{\overline{A}_{t-1},0}_t}{\partial \underline{Y}^{\overline{A}_t,0}_t} \right\}
\end{aligned}
$$

- When $t = 1$, start with the set of potential outcomes that are fully blipped down to time 1, i.e., $\overline{Y}^0_\delta$.
- When $t = 2, \cdots, K - \delta + 1$, add $Y^{\overline{A}_{t-1},0}_{t+\delta-1}$, which is fully blipped down to time t, at each step.
- No additional potential outcomes is added after $t = K - \delta + 2$.

Relaxing Sequential Ignorability Assumption

- Motivation: insufficient measured covariates to control for confounding in observational studies.
- Outcomes measured after treatment may contain information that allows control of confounding.
- Control for observed outcomes leads to bias in general.
- Potential outcomes can be viewed as pretreat variables and can be used to control for confounding in principle.

Relaxing Sequential Ignorability Assumption



- L_t^* denotes the complete set of covariates to achieve ignorability.
- L_t denotes the observed set of covariates.
- A_t is not ignorable conditioning on L_t alone.
- Ignorability can be achieved by conditioning on future potential outcomes, e.g., $A_1 \coprod Y_2^0 | Y_1^0$.

Relaxing Sequential Ignorability Assumption (cont'd)

Assume all outcomes are blipped down δ periods:

- Sequential ignorability assumption:
  $A_t \coprod Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)} | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, t = 1, \cdots, K$
- Relaxed ignorability assumption:
  $A_t \coprod Y^{\overline{A}_{t-1},0}_{(t+\tau):(t+\delta-1)} | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\tau-1)}, t = 1, \cdots, K-\tau$
  - Requires $\delta > \tau$.

The Revised Likelihood Function (cont'd)

Under the relaxed ignorability assumption, the likelihood function for the data is:

$$
\begin{aligned}
& f(\overline{Y}_K, \overline{A}_K, \overline{L}_K) \\
&= \prod_{t=1} \left\{ f(\overline{Y}^0_\delta) \times f(L_1|\overline{Y}^0_\delta) \times f(A_1|L_1, \overline{Y}^0_\tau) \times \frac{\partial \overline{Y}^0_\delta}{\partial \overline{Y}^{A_1,0}_\delta} \right\} \\
&\times \prod_{t=2}^{K-\delta+1} \left\{ \begin{array}{l} f(Y^{\overline{A}_{t-1},0}_{t-\delta+1}|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}) \\ \times f(L_t|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)}) \times f(A_t|\overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\tau-1)}) \times \dfrac{\partial Y^{\overline{A}_{t-1},0}_{t:(t+\delta-1)}}{\partial Y^{\overline{A}_t,0}_{t:(t+\delta-1)}} \end{array} \right\} \\
&\times \prod_{t=K-\delta+2}^{K-\tau} \left\{ f(L_t|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}^{\overline{A}_{t-1},0}_t) \times f(A_t|\overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y^{\overline{A}_{t-1},0}_{t:(t+\tau-1)}) \times \frac{\partial \underline{Y}^{\overline{A}_{t-1},0}_t}{\partial \underline{Y}^{\overline{A}_t,0}_t} \right\} \\
&\times \prod_{t=K-\tau+1}^{K} \left\{ f(L_t|\overline{L}_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}^{\overline{A}_{t-1},0}_t) \times f(A_t|\overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, \underline{Y}^{\overline{A}_{t-1},0}_t) \times \frac{\partial \underline{Y}^{\overline{A}_{t-1},0}_t}{\partial \underline{Y}^{\overline{A}_t,0}_t} \right\}
\end{aligned}
$$

The Estimating Equation

- Under sequential ignorability assumption, a practical estimating equation is:

$$S_\psi = \sum_{t=1}^{K} \left[\left\{ A_t - E\left(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}\right) \right\} \left(\sum_{m=t}^{min(t+\delta-1,K)} Y_m^{\overline{A}_{t-1},0}(\psi) \right) \right]$$

- Under revised assumption, a practical estimating equation is:

$$S_\psi = \sum_{t=1}^{K-\tau} \left[\left\{ A_t - E\left(A_t | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y_{t:(t+\tau-1)}^{\overline{A}_{t-1},0}(\psi)\right) \right\} \left(\sum_{m=t+\tau}^{min(t+\delta-1,K)} Y_m^{\overline{A}_{t-1},0}(\psi) \right) \right]$$

The Estimating Procedure

- Estimation procedure:
  - Start with an arbitrary value for the causal parameter ψ and calculate the putative potential outcomes
  - Update the parameters in the treatment model and the propensity score
  - Update the causal parameter
  - Iterate until convergence criterion is met
- Empirically works better than simultaneously updating all parameters
- Variance covariance matrix can be estimated using sandwich estimator

Simulation Setup

- Step 1: simulate $Y_1^0, Y_2^0, \cdots, Y_J^0$:

$$\begin{pmatrix} Y_1^0 \\ Y_2^0 \\ \vdots \\ Y_J^0 \end{pmatrix} \sim N \left\{ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho & \cdots & \rho^{J-1} \\ \rho & 1 & \cdots & \rho^{J-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{J-1} & \rho^{J-2} & \cdots & 1 \end{pmatrix} \right\}$$

  in which $\rho = 0.7, J = 9$.
- Step 2: set $L_0 = 0, A_0 = 0, Y_0 = 0$.

Simulation Setup (cont'd)

- Step 3: for $j = 1, 2, \cdots, J$:
  - 3a: $L_j \sim N(0.8L_{j-1} + 0.6A_{j-1} + 0.5Y_{j-1} + 0.4Y_j^0, 1)$
  - 3b: $logit\{E(A_j)\} = 0.6A_{j-1} + 0.1L_{j-1} + 0.3L_j + 0.2Y_{j-1} + \gamma Y_j^0$
  - 3c: $Y_j = Y_j^0 + \left(\sum_{t=\max(1, j-\delta+1)}^{j} A_t\right)\psi$, in which $\psi = 1$.
- γ determines whether treatment assignment depends on immediate future potential outcome.
- δ determines the time period during which the treatment has cumulative effect on the outcome.
- Simulated four scenarios:
  - $\gamma = 0, \delta = 9$
  - $\gamma = 0, \delta = 6$
  - $\gamma = 0.4, \delta = 9$
  - $\gamma = 0.4, \delta = 6$
- Sample size is 1000 with 1000 replicates.

Simulation I

Table : $\gamma = 0, \delta = 9$: ignorable treatment assignment; treatment effect cumulative during follow-up.

Type	δ	PE	Model-based SE	Empirical SE	Coverage Rate
Standard g-estimation	9	1.00	0.016	0.016	94.1%
	6	1.00	0.018	0.019	94.6%
	3	1.00	0.023	0.024	94.3%
Modified g-estimation	9	1.00	0.019	0.019	94.6%
	6	1.00	0.023	0.024	94.6%
	3	1.00	0.039	0.041	94.0%

Simulation II

Table : $\gamma = 0, \delta = 6$: ignorable treatment assignment; treatment effect cumulative over last six months only.

Type	δ	PE	Model-based SE	Empirical SE	Coverage Rate
Standard g-estimation	9	0.84	0.018	0.018	0.0%
	6	1.00	0.019	0.019	94.7%
	3	1.00	0.023	0.023	94.9%
Modified g-estimation	9	0.77	0.029	0.026	0.0%
	6	1.00	0.024	0.024	95.5%
	3	1.00	0.038	0.038	96.0%

Simulation III

Table : $\gamma = 0.4, \delta = 9$: nonignorable treatment assignment; treatment effect cumulative over follow-up.

Type	δ	PE	Model-based SE	Empirical SE	Coverage Rate
Standard g-estimation	9	1.07	0.014	0.015	0.4%
	6	1.08	0.017	0.017	0.3%
	3	1.13	0.023	0.022	0.0%
Modified g-estimation	9	1.00	0.020	0.020	94.5%
	6	1.00	0.024	0.025	94.2%
	3	0.99	0.042	0.042	95.3%

Simulation IV

Table : $\gamma = 0.4, \delta = 6$: nonignorable treatment assignment; treatment effect cumulative over last six months only.

Type	δ	PE	Model-based SE	Empirical SE	Coverage Rate
Standard g-estimation	9	0.92	0.015	0.016	0.1%
	6	1.09	0.017	0.017	0.1%
	3	1.14	0.023	0.023	0.0%
Modified g-estimation	9	0.77	0.033	0.030	0.0%
	6	1.00	0.025	0.025	94.7%
	3	1.00	0.041	0.041	94.4%

USRDS Data

- Included $N = 24,687$ incident dialysis patients from USRDS 2004 data
- A total of 134,595 months follow-up (average 5.5 months per patient)
- Baseline covariates: hemoglobin level before initiation of dialysis, dialysis chain ID, type of dialysis chain
- Time-updated covariates: monthly EPO dose, hematocrit level and number of days of hospitalization

Results

Table : Cumulative EPO effect; fixed blip down periods $\delta = 6$

Model	Standard g-estimation	Modified g-estimation
$\tau = 0$	0.19 (0.011)	
$\tau = 1$	0.23 (0.011)	0.31 (0.013)
$\tau = 2$	0.22 (0.013)	0.27 (0.014)
$\tau = 3$	0.20 (0.014)	0.25 (0.015)
$\tau = 4$	0.18 (0.016)	0.22 (0.018)
$\tau = 5$	0.14 (0.021)	0.18 (0.023)

- EPO effect estimated from modified g-estimation is consistently higher than from standard g-estimation.
  - Better control for confounding.
- The estimated EPO effect becomes smaller when τ increases.
  - Better control for confounding.
  - Misspecified causal model.

Results (cont'd)

Table : Cumulative EPO effect; $\tau = 1$ for modified g-estimation

Model	Standard g-estimation	Modified g-estimation
$\delta = 1$	-0.15 (0.007)	
$\delta = 4$	0.19 (0.010)	0.37 (0.014)
$\delta = 6$	0.23 (0.011)	0.31 (0.013)
$\delta = 8$	0.19 (0.010)	0.31 (0.009)
$\delta = 12$	0.19 (0.009)	0.29 (0.007)

- The estimated EPO effect becomes smaller when δ increases.
- May indicate misspecification of the causal model.

Summary

- Modified g-estimation for repeated outcomes:
  - Partially blipped down potential outcomes
  - Relaxed sequential ignorability assumption by conditioning on future potential outcomes
- Increased EPO dose is associated with increased hematocrit level
  - The effect was larger using modified g-estimation.
- Assumed linear dose response relationship and constant treatment effect over time.
  - Will relax both assumptions in future analyses.

Thank you!

Reserved Slides...

The Optimal Estimating Equation

The optimal estimating equation for location shift model, i.e., $\frac{\partial y_t^{\bar{0}}}{\partial y_t} = 1, t = 1, \cdots, K$

$$S_{\psi,eff} = \sum_{t=1}^{K} \left[\left\{ \frac{\partial \underline{Y}_t^{\overline{A}_{t-1},0}}{\partial \psi} - E\left(\frac{\partial \underline{Y}_t^{\overline{A}_{t-1},0}}{\partial \psi} | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1} \right) \right\} \frac{\partial \log f\left(\underline{Y}_t^{\overline{A}_{t-1},0} | \overline{L}_t, \overline{A}_{t-1}, \overline{Y}_{t-1} \right)}{\partial \underline{Y}_t^{\overline{A}_{t-1},0}} \right]$$

which depends on the joint distribution of potential outcomes

Results

- $M1$: Standard g-estimation w/o covariate adjustment
- $M2$: Standard g-estimation w covariate adjustment
- $M3$: Modified g-estimation w/o covariate adjustment
- $M4$: Modified g-estimation w covariate adjustment

Table : Average EPO effect; fixed blip down periods $\delta = 6$

Model	M1	M2	M3	M4
$\tau = 0$	0.51 (0.043)	0.73 (0.043)		
$\tau = 1$	0.79 (0.054)	1.06 (0.053)	1.33 (0.077)	1.57 (0.070)
$\tau = 2$	0.87 (0.067)	1.11 (0.064)	1.24 (0.085)	1.41 (0.076)
$\tau = 3$	0.92 (0.080)	1.10 (0.077)	1.24 (0.096)	1.34 (0.085)
$\tau = 4$	0.97 (0.095)	1.04 (0.095)	1.29 (0.11)	1.25 (0.10)
$\tau = 5$	1.02 (0.11)	0.85 (0.13)	1.31 (0.13)	1.07 (0.13)

Results (cont'd)

- $M1$: Standard g-estimation w/o covariate adjustment
- $M2$: Standard g-estimation w covariate adjustment
- $M3$: Modified g-estimation w/o covariate adjustment
- $M4$: Modified g-estimation w covariate adjustment

Table : Average EPO effect; different blip down periods. $\tau = 0$ for M1 and M2. $\tau = 1$ for M3 and M4

Model	M1	M2	M3	M4
$\delta = 1$	-0.18 (0.008)	-0.15 (0.007)		
$\delta = 4$	0.37 (0.029)	0.59 (0.031)	1.23 (0.061)	1.48 (0.057)
$\delta = 8$	0.65 (0.050)	0.86 (0.049)	1.50 (0.087)	1.74 (0.078)
$\delta = 12$	0.79 (0.054)	0.99 (0.052)	1.67 (0.091)	1.98 (0.079)

Results

- $M1$: Standard g-estimation w/o covariate adjustment
- $M2$: Standard g-estimation w covariate adjustment
- $M3$: Modified g-estimation w/o covariate adjustment
- $M4$: Modified g-estimation w covariate adjustment

Table : Cumulative EPO effect; fixed blip down periods $\delta = 6$

Model	M1	M2	M3	M4
$\tau = 0$	0.13 (0.011)	0.19 (0.011)		
$\tau = 1$	0.17 (0.012)	0.23 (0.011)	0.35 (0.010)	0.31 (0.013)
$\tau = 2$	0.17 (0.013)	0.22 (0.013)	0.32 (0.010)	0.27 (0.014)
$\tau = 3$	0.17 (0.014)	0.20 (0.014)	0.31 (0.010)	0.25 (0.015)
$\tau = 4$	0.17 (0.016)	0.18 (0.016)	0.31 (0.010)	0.22 (0.018)
$\tau = 5$	0.17 (0.019)	0.14 (0.021)	0.32 (0.010)	0.18 (0.023)

Results (cont'd)

- $M1$: Standard g-estimation w/o covariate adjustment
- $M2$: Standard g-estimation w covariate adjustment
- $M3$: Modified g-estimation w/o covariate adjustment
- $M4$: Modified g-estimation w covariate adjustment

Table : Cumulative EPO effect, different blip down periods. $\tau = 0$ for M1 and M2. $\tau = 1$ for M3 and M4

Model	M1	M2	M3	M4
$\delta = 1$	-0.18 (0.008)	-0.15 (0.007)		
$\delta = 4$	0.12 (0.009)	0.19 (0.010)	0.37 (0.012)	0.37 (0.014)
$\delta = 6$	0.17 (0.012)	0.23 (0.011)	0.35 (0.010)	0.31 (0.013)
$\delta = 8$	0.14 (0.011)	0.19 (0.010)	0.34 (0.008)	0.31 (0.009)
$\delta = 12$	0.16 (0.010)	0.19 (0.009)	0.32 (0.006)	0.29 (0.007)