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
1 Motivation

2 Methods
   - Review of standard g-estimation
   - Extensions of g-estimation

3 Illustration
   - Simulation
   - Data Analyses

4 Summary
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).
\( \overline{L}_K = \{L_1, \cdots, L_K\} \): 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 $Y_t, t = 1, \cdots, K$

\begin{align*}
t = K & \quad Y_K \\
t = K - 1 & \quad Y_{K-1} \\
t = K - 2 & \quad Y_{K-2} \\
\vdots & \quad \vdots \\
t = 2 & \quad Y_2 \\
t = 1 & \quad Y_1 \\
t = 0 & 
\end{align*}
The Blip Down Process

- Start with $Y_K$ and remove the effect of $A_K$ on $Y_K$ to get $Y_K^{A_{K-1},0}$
## The Blip Down Process

<table>
<thead>
<tr>
<th>$t$</th>
<th>$Y_t$</th>
<th>$\bar{A}_{t-1,0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = K$</td>
<td>$Y_K$</td>
<td>$\bar{A}_{K-1,0}$</td>
</tr>
<tr>
<td>$t = K - 1$</td>
<td>$Y_{K-1}$</td>
<td>$\bar{A}_{K-2,0}$</td>
</tr>
<tr>
<td>$t = K - 2$</td>
<td>$Y_{K-2}$</td>
<td>$\bar{A}_{K-2,0}$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td></td>
</tr>
<tr>
<td>$t = 2$</td>
<td>$Y_2$</td>
<td></td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$Y_1$</td>
<td></td>
</tr>
<tr>
<td>$t = 0$</td>
<td>$Y_0$</td>
<td></td>
</tr>
</tbody>
</table>

- Remove the effect of $A_{K-1}$ on $Y_{K-1}$ to get $\bar{A}_{K-2,0}$
- Remove the effect of $A_{K-1}$ on $Y_{K}^{\bar{A}_{K-1},0}$ to get $Y_{K}^{\bar{A}_{K-2},0}$
The Blip Down Process

<table>
<thead>
<tr>
<th>$t = K$</th>
<th>$Y_K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = K - 1$</td>
<td>$Y_{K-1}$</td>
</tr>
<tr>
<td>$t = K - 2$</td>
<td>$Y_{K-2}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$t = 2$</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>$t = 0$</td>
<td>$Y_0$</td>
</tr>
</tbody>
</table>

- Remove the effect of $\bar{A}_K$ on $\bar{Y}_K$ to get $\bar{Y}_0$.
- 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 $K = 6$:

<table>
<thead>
<tr>
<th>$t$</th>
<th>$Y_1$</th>
<th>$Y_2$</th>
<th>$Y_3$</th>
<th>$Y_4$</th>
<th>$Y_5$</th>
<th>$Y_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 6$</td>
<td>$Y_0^1(\psi)$</td>
<td>$Y_0^2(\psi)$</td>
<td>$Y_0^3(\psi)$</td>
<td>$Y_0^4(\psi)$</td>
<td>$Y_0^5(\psi)$</td>
<td>$Y_0^6(\psi)$</td>
</tr>
<tr>
<td>$t = 5$</td>
<td>$Y_1$</td>
<td>$Y_2^A(\psi)$</td>
<td>$Y_3^A(\psi)$</td>
<td>$Y_4^A(\psi)$</td>
<td>$Y_5^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
<tr>
<td>$t = 4$</td>
<td>$Y_2$</td>
<td>$Y_3^A(\psi)$</td>
<td>$Y_4^A(\psi)$</td>
<td>$Y_5^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
<tr>
<td>$t = 3$</td>
<td>$Y_3$</td>
<td>$Y_4^A(\psi)$</td>
<td>$Y_5^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
<tr>
<td>$t = 2$</td>
<td>$Y_4$</td>
<td>$Y_5^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$Y_5$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
<tr>
<td>$t = 0$</td>
<td>$Y_6$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
<td>$Y_6^A(\psi)$</td>
</tr>
</tbody>
</table>

- SNMs used to parametrize the blip-down process. $\psi$ 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:

\[
f(\overline{Y}_K, \overline{A}_K, L_K) = \frac{\partial Y^0_K}{\partial Y^K} \times f(\overline{Y}_K, \overline{A}_K, L_K)
= f(\overline{Y}^0_K) \times \prod_{t=1}^{K} \left\{ f(L_t | L_{t-1}, \overline{A}_{t-1}, \overline{Y}_{t-1}, \overline{Y}_{A_{t-1}, 0}) \times f(A_t | L_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, \overline{Y}_{A_{t-1}, 0}) \times \frac{\partial Y_{A_{t-1}, 0}}{\partial Y_{A_{t}, 0}} \right\}
\]

where \( Y_{A_{t-1}, 0} = \left\{ Y_t^{A_{t-1}, 0}, Y_{t+1}^{A_{t-1}, 0}, \ldots, Y_K^{A_{t-1}, 0} \right\} \)

- Start with the whole vector of potential outcomes \( \overline{Y}_K \)
- For \( t = 1, \ldots, K \)
  - Generate \( L_t \)
  - Generate \( A_t \)
  - Blip up all potential outcomes adding the effect of \( A_t \)
Sequential Ignorability Assumption

- \( A_t \prod Y_t^{\bar{A}_{t-1},0} | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, t = 1, \ldots, K \)
- Under sequential ignorability assumptions, the likelihood function is:

\[
f(\bar{Y}_K, \bar{A}_K, \bar{L}_K) = f(\bar{Y}_K^0) \times \prod_{t=1}^{K} \left\{ f(L_t | L_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times f(A_t | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}) \times \frac{\partial Y_t^{\bar{A}_{t-1},0}}{\partial Y_t^{\bar{A}_{t-1},0}} \right\}
\]
An Estimating Equation

A choice of the estimating equation is:

\[
S_\psi = \sum_{t=1}^{K} \left[ \left\{ A_t - E \left( A_t | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1} \right) \right\} \left( \sum_{m=t}^{K} Y_{m|A_{t-1},0}(\psi) \right) \right]
\]

where \( E \left( A_t | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1} \right) \) is the propensity score.
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.,
\[
\left\{ Y_{t-t-1,0}, \ldots, Y_{t-t-\delta,0} \right\}, \ t = 1, \ldots, K,
\]
where \( \delta \) is the number of blip down periods.

- Similar to the idea of history-adjusted marginal structural models.
- Use finite dimensional parameter \( \psi \) for potential outcomes within \( \delta \) periods; infinite dimensional parameter beyond that.
Partial Blip Down Process (cont’d)

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

<table>
<thead>
<tr>
<th>$t = 6$</th>
<th>$t = 5$</th>
<th>$t = 4$</th>
<th>$t = 3$</th>
<th>$t = 2$</th>
<th>$t = 1$</th>
<th>$t = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_6$</td>
<td>$Y_5$</td>
<td>$Y_4$</td>
<td>$Y_3$</td>
<td>$Y_2$</td>
<td>$Y_1$</td>
<td>$Y_0$</td>
</tr>
</tbody>
</table>

- Models for potential outcomes in red are parametrized with finite dimensional parameter $\psi$.
- Models for potential outcomes in gray are left unspecified.
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:
  \[ Y_{t+\delta}^0 = Y_{t+\delta} - \left( \sum_{j=t+1}^{t+\delta} A_j \right) \psi \]

- Model for partially blipped down potential outcomes:
  \[ Y_{t+\delta}^{A_{t,0}} = Y_{t+\delta} - \left( \sum_{j=t+1}^{t+\delta} A_j \right) \psi \]

- 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

\[
f(\bar{Y}_K, \bar{A}_K, \bar{L}_K) \\
= \left\{ f(\bar{Y}_\delta^0) \times f(L_1|Y_0^\delta) \times f(A_1|L_1, \bar{Y}_0^\delta) \times \frac{\partial \bar{Y}_0^\delta}{\partial Y_1^{A_1,0}} \right\} \\
\times \prod_{t=2}^{K-\delta+1} \left\{ f(Y_{t+\delta-1}^{A_{t-1},0}|L_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}) \\
\times f(L_t|L_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_{t}^{A_{t-1},0}) \times f(A_t|L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_{t}^{A_{t-1},0}) \times \frac{\partial Y_{t}^{A_{t-1},0}}{\partial Y_{t}^{A_{t},0}} \right\} \\
\times \prod_{t=K-\delta+2}^{K} \left\{ f(L_t|L_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_{t}^{A_{t-1},0}) \times f(A_t|L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_{t}^{A_{t-1},0}) \times \frac{\partial Y_{t}^{A_{t-1},0}}{\partial Y_{t}^{A_{t},0}} \right\}
\]

- When \( t = 1 \), start with the set of potential outcomes that are fully blipped down to time 1, i.e., \( \bar{Y}_\delta^0 \).

- When \( t = 2, \ldots, K - \delta + 1 \), add \( Y_{t+\delta-1}^{A_{t-1},0} \), which is fully blipped down to time \( t \), at each step.

- No additional potential outcomes is added after \( t = K - \delta + 2 \).
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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ 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., $A_1 \perp Y^0_2 \mid Y^0_1$. 

$L^*$ denotes the complete set of covariates to achieve ignorability.

$L_t$ denotes the observed set of covariates.
Assume all outcomes are blipped down $\delta$ periods:

- **Sequential ignorability assumption:**
  \[ A_t \perp Y_{t:(t+\delta-1)}^{\overline{A}_{t-1},0} | L_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, t = 1, \cdots, K \]

- **Relaxed ignorability assumption:**
  \[ A_t \perp Y_{(t+\tau):(t+\delta-1)}^{\overline{A}_{t-1},0} | L_t, \overline{A}_{t-1}, \overline{Y}_{t-1}, Y_{t:(t+\tau-1)}^{\overline{A}_{t-1},0}, t = 1, \cdots, K - \tau \]
  
  - Requires $\delta > \tau$. 

...
The Revised Likelihood Function (cont’d)

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

\[
f(\bar{Y}_K, \bar{A}_K, \bar{L}_K) = \prod_{t=1}^{K-\delta+1} \left\{ f(\bar{Y}_0^\delta) \times f(L_1|\bar{Y}_\delta^0) \times f(A_1|L_1, \bar{Y}_\tau^0) \times \frac{\partial \bar{Y}_0^0}{\partial \bar{Y}_A^{1,0}} \right\}
\]

\[
\times \prod_{t=2}^{K-\delta+1} \left\{ f(Y_{t-\delta+1}^{\delta-1,0}|\bar{L}_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}) \right. \\
\left. \times f(L_t|\bar{L}_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times f(A_t|\bar{L}_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times \frac{\partial Y_t^{\bar{A}_{t-1},0}}{\partial Y_t^{A_t,0}} \right\}
\]

\[
\times \prod_{t=K-\delta+2}^{K-\tau} \left\{ f(L_t|\bar{L}_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times f(A_t|\bar{L}_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times \frac{\partial Y_t^{\bar{A}_{t-1},0}}{\partial Y_t^{A_t,0}} \right\}
\]

\[
\times \prod_{t=K-\tau+1}^{K} \left\{ f(L_t|\bar{L}_{t-1}, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times f(A_t|\bar{L}_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_t^{\bar{A}_{t-1},0}) \times \frac{\partial Y_t^{\bar{A}_{t-1},0}}{\partial Y_t^{A_t,0}} \right\}
\]
The Estimating Equation

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

\[
S_\psi = \sum_{t=1}^{K} \left[ \{ A_t - E (A_t | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}) \} \left( \sum_{m=t}^{\min(t+\delta-1, K)} Y_{m-1,0}^{\bar{A}_{t-1}}(\psi) \right) \right]
\]

- Under revised assumption, a practical estimating equation is:

\[
S_\psi = \sum_{t=1}^{K-\tau} \left[ \{ A_t - E (A_t | L_t, \bar{A}_{t-1}, \bar{Y}_{t-1}, Y_{t:(t+\tau-1)}^{\bar{A}_{t-1},0}(\psi)) \} \left( \sum_{m=t+\tau}^{\min(t+\delta-1, K)} Y_{m-1,0}^{\bar{A}_{t-1}}(\psi) \right) \right]
\]
The Estimating Procedure

- Estimation procedure:
  - Start with an arbitrary value for the causal parameter $\psi$ 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 $Y_1^0, Y_2^0, \ldots, Y_J^0$:

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

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 $j = 1, 2, \ldots, J$:

- 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 $\psi = 1$.

- $\gamma$ determines whether treatment assignment depends on immediate future potential outcome.
- $\delta$ 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.
Table: $\gamma = 0, \delta = 9$: ignorable treatment assignment; treatment effect cumulative during follow-up.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\delta$</th>
<th>PE</th>
<th>Model-based SE</th>
<th>Empirical SE</th>
<th>Coverage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard g-estimation</td>
<td>9</td>
<td>1.00</td>
<td>0.016</td>
<td>0.016</td>
<td>94.1%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.018</td>
<td>0.019</td>
<td>94.6%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.023</td>
<td>0.024</td>
<td>94.3%</td>
</tr>
<tr>
<td>Modified g-estimation</td>
<td>9</td>
<td>1.00</td>
<td>0.019</td>
<td>0.019</td>
<td>94.6%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.023</td>
<td>0.024</td>
<td>94.6%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.039</td>
<td>0.041</td>
<td>94.0%</td>
</tr>
</tbody>
</table>
Simulation II

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

<table>
<thead>
<tr>
<th>Type</th>
<th>( \delta )</th>
<th>PE</th>
<th>Model-based SE</th>
<th>Empirical SE</th>
<th>Coverage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard g-estimation</td>
<td>9</td>
<td>0.84</td>
<td>0.018</td>
<td>0.018</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.019</td>
<td>0.019</td>
<td>94.7%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.023</td>
<td>0.023</td>
<td>94.9%</td>
</tr>
<tr>
<td>Modified g-estimation</td>
<td>9</td>
<td>0.77</td>
<td>0.029</td>
<td>0.026</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.024</td>
<td>0.024</td>
<td>95.5%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.038</td>
<td>0.038</td>
<td>96.0%</td>
</tr>
</tbody>
</table>
Table: $\gamma = 0.4, \delta = 9$: nonignorable treatment assignment; treatment effect cumulative over follow-up.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\delta$</th>
<th>PE</th>
<th>Model-based SE</th>
<th>Empirical SE</th>
<th>Coverage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard g-estimation</td>
<td>9</td>
<td>1.07</td>
<td>0.014</td>
<td>0.015</td>
<td>0.4%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.08</td>
<td>0.017</td>
<td>0.017</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.13</td>
<td>0.023</td>
<td>0.022</td>
<td>0.0%</td>
</tr>
<tr>
<td>Modified g-estimation</td>
<td>9</td>
<td>1.00</td>
<td>0.020</td>
<td>0.020</td>
<td>94.5%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.024</td>
<td>0.025</td>
<td>94.2%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.99</td>
<td>0.042</td>
<td>0.042</td>
<td>95.3%</td>
</tr>
</tbody>
</table>
Table: $\gamma = 0.4, \delta = 6$: nonignorable treatment assignment; treatment effect cumulative over last six months only.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\delta$</th>
<th>PE</th>
<th>Model-based SE</th>
<th>Empirical SE</th>
<th>Coverage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard g-estimation</td>
<td>9</td>
<td>0.92</td>
<td>0.015</td>
<td>0.016</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.09</td>
<td>0.017</td>
<td>0.017</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.14</td>
<td>0.023</td>
<td>0.023</td>
<td>0.0%</td>
</tr>
<tr>
<td>Modified g-estimation</td>
<td>9</td>
<td>0.77</td>
<td>0.033</td>
<td>0.030</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.00</td>
<td>0.025</td>
<td>0.025</td>
<td>94.7%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.00</td>
<td>0.041</td>
<td>0.041</td>
<td>94.4%</td>
</tr>
</tbody>
</table>
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
Table: Cumulative EPO effect; fixed blip down periods $\delta = 6$

<table>
<thead>
<tr>
<th>Model $\tau$</th>
<th>Standard g-estimation</th>
<th>Modified g-estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 0$</td>
<td>0.19 (0.011)</td>
<td></td>
</tr>
<tr>
<td>$\tau = 1$</td>
<td>0.23 (0.011)</td>
<td>0.31 (0.013)</td>
</tr>
<tr>
<td>$\tau = 2$</td>
<td>0.22 (0.013)</td>
<td>0.27 (0.014)</td>
</tr>
<tr>
<td>$\tau = 3$</td>
<td>0.20 (0.014)</td>
<td>0.25 (0.015)</td>
</tr>
<tr>
<td>$\tau = 4$</td>
<td>0.18 (0.016)</td>
<td>0.22 (0.018)</td>
</tr>
<tr>
<td>$\tau = 5$</td>
<td>0.14 (0.021)</td>
<td>0.18 (0.023)</td>
</tr>
</tbody>
</table>

- 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 $\tau$ increases.
  - Better control for confounding.
  - Misspecified causal model.
Table: Cumulative EPO effect; $\tau = 1$ for modified g-estimation

<table>
<thead>
<tr>
<th>Model</th>
<th>Standard g-estimation</th>
<th>Modified g-estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta = 1$</td>
<td>-0.15 (0.007)</td>
<td></td>
</tr>
<tr>
<td>$\delta = 4$</td>
<td>0.19 (0.010)</td>
<td>0.37 (0.014)</td>
</tr>
<tr>
<td>$\delta = 6$</td>
<td>0.23 (0.011)</td>
<td>0.31 (0.013)</td>
</tr>
<tr>
<td>$\delta = 8$</td>
<td>0.19 (0.010)</td>
<td>0.31 (0.009)</td>
</tr>
<tr>
<td>$\delta = 12$</td>
<td>0.19 (0.009)</td>
<td>0.29 (0.007)</td>
</tr>
</tbody>
</table>

- The estimated EPO effect becomes smaller when $\delta$ 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 for location shift model, i.e.,

\[ \frac{\partial y^0_t}{\partial y_t} = 1, \; t = 1, \cdots, K \]

\[
S_{\psi, \text{eff}} = \sum_{t=1}^{K} \left\{ \frac{\partial Y_{t}^{\bar{A}_{t-1},0}}{\partial \psi} - E \left( \frac{\partial Y_{t}^{\bar{A}_{t-1},0}}{\partial \psi} \middle| L_t, \bar{A}_{t-1}, \bar{Y}_{t-1} \right) \right\} \frac{\partial \log f \left( Y_{t}^{\bar{A}_{t-1},0} \middle| L_t, \bar{A}_{t-1}, \bar{Y}_{t-1} \right)}{\partial Y_{t}^{\bar{A}_{t-1},0}}
\]

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

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

<table>
<thead>
<tr>
<th>Model</th>
<th>( \delta = 1 )</th>
<th>( \delta = 4 )</th>
<th>( \delta = 8 )</th>
<th>( \delta = 12 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M1</td>
<td>M2</td>
<td>M3</td>
<td>M4</td>
</tr>
<tr>
<td>( \delta = 1 )</td>
<td>-0.18 (0.008)</td>
<td>-0.15 (0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta = 4 )</td>
<td>0.37 (0.029)</td>
<td>0.59 (0.031)</td>
<td>1.23 (0.061)</td>
<td>1.48 (0.057)</td>
</tr>
<tr>
<td>( \delta = 8 )</td>
<td>0.65 (0.050)</td>
<td>0.86 (0.049)</td>
<td>1.50 (0.087)</td>
<td>1.74 (0.078)</td>
</tr>
<tr>
<td>( \delta = 12 )</td>
<td>0.79 (0.054)</td>
<td>0.99 (0.052)</td>
<td>1.67 (0.091)</td>
<td>1.98 (0.079)</td>
</tr>
</tbody>
</table>
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>
<thead>
<tr>
<th>Model</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau = 0)</td>
<td>0.13 (0.011)</td>
<td>0.19 (0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\tau = 1)</td>
<td>0.17 (0.012)</td>
<td>0.23 (0.011)</td>
<td>0.35 (0.010)</td>
<td>0.31 (0.013)</td>
</tr>
<tr>
<td>(\tau = 2)</td>
<td>0.17 (0.013)</td>
<td>0.22 (0.013)</td>
<td>0.32 (0.010)</td>
<td>0.27 (0.014)</td>
</tr>
<tr>
<td>(\tau = 3)</td>
<td>0.17 (0.014)</td>
<td>0.20 (0.014)</td>
<td>0.31 (0.010)</td>
<td>0.25 (0.015)</td>
</tr>
<tr>
<td>(\tau = 4)</td>
<td>0.17 (0.016)</td>
<td>0.18 (0.016)</td>
<td>0.31 (0.010)</td>
<td>0.22 (0.018)</td>
</tr>
<tr>
<td>(\tau = 5)</td>
<td>0.17 (0.019)</td>
<td>0.14 (0.021)</td>
<td>0.32 (0.010)</td>
<td>0.18 (0.023)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Model</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta = 1 )</td>
<td>-0.18 (0.008)</td>
<td>-0.15 (0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta = 4 )</td>
<td>0.12 (0.009)</td>
<td>0.19 (0.010)</td>
<td>0.37 (0.012)</td>
<td>0.37 (0.014)</td>
</tr>
<tr>
<td>( \delta = 6 )</td>
<td>0.17 (0.012)</td>
<td>0.23 (0.011)</td>
<td>0.35 (0.010)</td>
<td>0.31 (0.013)</td>
</tr>
<tr>
<td>( \delta = 8 )</td>
<td>0.14 (0.011)</td>
<td>0.19 (0.010)</td>
<td>0.34 (0.008)</td>
<td>0.31 (0.009)</td>
</tr>
<tr>
<td>( \delta = 12 )</td>
<td>0.16 (0.010)</td>
<td>0.19 (0.009)</td>
<td>0.32 (0.006)</td>
<td>0.29 (0.007)</td>
</tr>
</tbody>
</table>