Biometrics 00, 1–12 DOI: 10.1111/j.0000-0000.0000.00000.x
December 2000

Supplementary material for: Modeling the impact of hepatitis C viral clearance on end-stage liver disease in an HIV co-infected cohort with Targeted Maximum Likelihood Estimation

Mireille E Schnitzer[1,*], Erica EM Moodie[2], Mark J van der Laan[3], Robert W Platt[2], and Marina B Klein[4]

[1]Department of Biostatistics, Harvard School of Public Health, Boston, MA, 02115, USA

[2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, H3A 1A2, Canada

[3]Division of Biostatistics, School of Public Health, University of California, Berkeley, CA, 94720, USA

[4]Department of Medicine, McGill University, Royal Victoria Hospital, Montréal, Québec, H3A 1A1, Canada

email: mschnitz@hsph.harvard.edu

This paper has been submitted for consideration for publication in *Biometrics*

1. Background of Targeted Maximum Likelihood Estimation

Estimation with Targeted Maximum Likelihood Estimation (TMLE) requires that the target parameter be identified as a differentiable function of a component of the underlying data density. Let Ψ be a differentiable function that takes an argument in a model space $\mathcal{M}$ and value in the space of real numbers (or vectors). Then, our parameter can be defined as $\psi \equiv \Psi(Q)$, for some Ψ as described, and where $Q \in \mathcal{M}$ is some component of the underlying data density, P. Treating the function Ψ as a substitution estimator, we can estimate Q using a dataset, O, and plug it into the function Ψ so that the estimate of the target parameter becomes $\hat{\psi} = \Psi(\hat{Q})$.

Loosely, the influence curve of a regular, asymptotically linear (RAL) estimator is the component of the estimator that determines its asymptotic properties. Suppose one observes n sets of independent, identically distributed subject-specific data, $O = \{O_i, i = 1, .., n\}$. The influence curve of the estimator $\Psi(\hat{Q})$ is a function of the data and denoted $D(P)(O)$ (with subject-specific components $D(P)(O_i)$). The influence curve can be defined as

$$n^{1/2}[\Psi(\hat{Q}) - \psi] = n^{-1/2} \sum_{i=1}^{n} D(P)(O_i) + o_P(1)$$

where $o_P(i)$ is a random term than converges to zero in probability (van der Laan and Robins, 2003; Tsiatis, 2006). By an application of the Central Limit Theorem, this implies that

$$n^{1/2}\{\Psi(\hat{Q}) - \psi\} \rightarrow^{\mathcal{D}} N[0, E\{D(P)D(P)^T\}]$$

so that the influence curve provides a large-sample approximation for the variance of the estimator. Specifically, $Var\{\Psi(\hat{Q})\} \approx 1/nVar\{D(P)\}$. There is a lower variance bound in the class of influence curves of RAL estimators, and the unique influence curve that attains this bound is called the efficient influence curve (Tsiatis, 2006).

The general TMLE procedure is described in van der Laan and Rubin (2006). TMLE is defined by the procedure of updating the estimate, $\hat{Q}$, of the component of the data density used in the substitution estimator $\Psi(\hat{Q})$ in order to produce inference using the efficient

influence curve. For example, TMLE is often implemented using the G-computation formula with carefully constructed updates for the density estimates (Rosenblum and van der Laan, 2010).

Estimators based on the efficient influence curve are favoured for having low variance. For many causal parameters, estimation with the efficient influence curve is also doubly-robust (Kang and Schafer, 2007). This means that, while two components of the underlying density P must be fit, the estimator is asymptotically unbiased if either of the components is correctly specified. In the examples of the longitudinal TMLEs presented in this paper, only the exposure mechanism or the outcome models must be correctly specified in order for the method to be asymptotically unbiased.

2. Target parameter of the IPTW

Here we show that the inverse probability of treatment weighted (IPTW) estimator for the hazard model used in the simulation study and the example targets the same parameter of interest as the TMLE.

Let $\lambda_\beta(\bar{a}, t) = \text{expit}(X_{l,t}^T \boldsymbol{\beta})$ be the logit-linear model for the hazard. The parameter estimated by the TMLE was defined as

$$\text{argmax}_{\boldsymbol{\beta}} \sum_{l,t} S_{\bar{a}^l}(t-1) \left[\lambda_{\bar{a}^l}(t) \log(\lambda_\beta) + \{1 - \lambda_{\bar{a}^l}(t)\} \log(1-\lambda_\beta)\right].$$

Factoring out $\lambda_{\bar{a}^l}(t)$ gives

$$\begin{aligned}
&\text{argmax}_{\boldsymbol{\beta}} \sum_{l,t} S_{\bar{a}^l}(t-1) \lambda_{\bar{a}^l}(t) \left[\{\log \lambda_\beta - \log(1-\lambda_\beta)\} + \log(1-\lambda_\beta)\right] \\
&= \text{argmax}_{\boldsymbol{\beta}} \sum_{l,t} S_{\bar{a}^l}(t-1) \lambda_{\bar{a}^l}(t) \left[\log\left\{\frac{\lambda_\beta}{(1-\lambda_\beta)}\right\} + \log(1-\lambda_\beta)\right].
\end{aligned}$$

To maximize this expression, the derivative with respect to $\boldsymbol{\beta}$ can be taken and the resulting

expression set to zero. This results in

$$\begin{aligned}
&\sum_{l,t} S_{\bar{a}^l}(t-1)\lambda_{\bar{a}^l}(t)\left[\frac{d}{d\boldsymbol{\beta}}\log\left\{\frac{\lambda_\beta}{(1-\lambda_\beta)}\right\}+\frac{d}{d\boldsymbol{\beta}}\log(1-\lambda_\beta)\right]\\
&=\frac{d}{d\boldsymbol{\beta}}(\lambda_\beta)\frac{1}{\lambda_\beta(1-\lambda_\beta)}\left\{\lambda_{\bar{a}^l}(t)-\lambda_\beta\right\}\\
&=\frac{d}{d\boldsymbol{\beta}}\text{logit}(\lambda_\beta)\left\{\lambda_{\bar{a}^l}(t)-\lambda_\beta\right\}=0.
\end{aligned}$$

Noting that $\text{logit}(\lambda_\beta) = X_{l,t}^T\boldsymbol{\beta}$ is the linear specification so that the derivative with respect to $\boldsymbol{\beta}$ is the vector of variables in the marginal structural model. The above score equation is therefore the logistic regression defined on the counterfactuals that is solved by the IPTW-MSM.

3. Efficient influence function for the MSM for the hazard function

We are interested in estimation of $\boldsymbol{\beta}$, which can be implicitly written as a function of the parameters $\mathbf{S} = (S_{\bar{a}^l}(t),\ \text{for all unique values of } \bar{a}_t^l)$ through the score equation:

$$0 = U(\mathbf{S},\boldsymbol{\beta}) \equiv \sum_{l,t} S_{\bar{a}^l}(t-1)X_{l,t}\left\{\frac{S_{\bar{a}^l}(t)-S_{\bar{a}^l}(t-1)}{S_{\bar{a}^l}(t-1)} - \text{expit}(X_{l,t}^T\boldsymbol{\beta})\right\}.$$

In order to derive the efficient influence function for $\boldsymbol{\beta}$, we will use the functional delta method (van der Vaart and Wellner, 1996). In this context, it states that for a parameter $\boldsymbol{\beta} = \boldsymbol{\beta}(\mathbf{S})$ that can be written as a function of other parameters whose efficient influence functions, $D_{\bar{a}^l,t}$ are already known, the efficient influence function for $\boldsymbol{\beta}$ is equal to

$$D_{\boldsymbol{\beta}} = \sum_{l,t}\frac{d\boldsymbol{\beta}(\mathbf{S})}{dS_{\bar{a}^l}(t)}D_{\bar{a}^l,t}. \tag{1}$$

By the implicit function theorem, the derivative in Equation (1) can be obtained using

$$\frac{d\boldsymbol{\beta}(\mathbf{S})}{dS_{\bar{a}^l}(t)} = -\left\{\frac{dU(\mathbf{S},\boldsymbol{\beta})}{d\boldsymbol{\beta}}\right\}^{-1}\frac{dU(\mathbf{S},\boldsymbol{\beta})}{dS_{\bar{a}^l}(t)}. \tag{2}$$

We then obtain

$$\frac{dU(\mathbf{S},\boldsymbol{\beta})}{d\boldsymbol{\beta}} = -\sum_{l,t} S_{\bar{a}^l}(t-1)X_{l,t}X_{l,t}^T\frac{\exp(X_{l,t}^T\boldsymbol{\beta})}{\{1+\exp(X_{l,t}^T\boldsymbol{\beta})\}^2},$$

a matrix with dimension $r \times R$. For each unique exposure pattern $\bar{a}_t^l$,

$$\frac{dU(\mathbf{S}, \boldsymbol{\beta})}{dS_{\bar{a}^l}(t)} = X_{l,t} - \sum_{m:\{\bar{a}_t^l \subset \bar{a}_{t+1}^m\}} X_{m,t+1}\{1 + \text{expit}(X_{m,t+1}^T \boldsymbol{\beta})\},$$

a column vector of length R. The above summation is taken over all m for which the truncated exposure pattern $\bar{a}_t^l$ is a subset of the pattern $\bar{a}_{t+1}^m$ (or, equivalently, $\bar{a}_t^m = \bar{a}_t^l$) so that in particular, $S_{\bar{a}^m}(t) = S_{\bar{a}^l}(t)$. The two above components can be numerically evaluated and combined to form a column vector of length R using Equation (2).

Substituting these expressions into Equation (2) gives a form for $d\boldsymbol{\beta}(\mathbf{S})/d\boldsymbol{\beta}$ which can then be substituted into Equation (1) to produce the form of the efficient influence function for the parameters of the MSM:

$$D_{\boldsymbol{\beta}} = \left[\sum_{l,t} S_{\bar{a}^l}(t-1) X_{l,t} X_{l,t}^T \frac{\exp(X_{l,t}^T \boldsymbol{\beta})}{\{1 + \exp(X_{l,t}^T \boldsymbol{\beta})\}^2}\right]^{-1}$$
$$\sum_{l,t} \left[X_{l,t} - \sum_{m:\{\bar{a}_t^l \subset \bar{a}_{t+1}^m\}} X_{m,t+1}\{1 + \text{expit}(X_{m,t+1}^T \boldsymbol{\beta})\}\right] D_{\bar{a}^l,t}.$$

The efficient influence function components can be numerically evaluated for each of the n subjects, producing an influence matrix of dimension $n \times R$, representing the joint influence components for $\boldsymbol{\beta}$.

4. MSM for the log-odds of survival

A model for the log-odds of survival can be described as $\log[S_{\bar{a}^l}(t)/\{1 - S_{\bar{a}^l}(t)\}] = X_{l,t}^T \boldsymbol{\beta}$ for all unique patterns $\bar{a}_t^l$, where $X_{l,t}^T \boldsymbol{\beta}$ represents the form of the linear specification of the model. Let $\mathbf{X}$ be the design matrix, potentially including functions of $\bar{a}^l$ and t. Let $X_{l,t}$ represent the R-dimensional row of the design matrix corresponding with exposure $\bar{a}^l$ and time t, represented as a column vector. For example, if the MSM was a linear model with an intercept and a linear term for time, then for each unique pattern $\bar{a}_{t^*}^l$ for the time point t^*, $X_{l,t^*} = (1, t^*)^T$. The design matrix can also contain subgroups if $\mathbf{S}$ was calculated separately for the components of a categorical variable, V, and although we do not include conditioning

in our notation for simplicity, the following development easily extends to such a case. Finally, let $\boldsymbol{\beta}$ denote the vector of coefficients corresponding with the columns of the design matrix. Therefore, since there are M estimates for the survival function, the dimension of the matrix $\mathbf{X}$ is R by M, corresponding with a $\boldsymbol{\beta}$-vector of length R.

The parameter $\boldsymbol{\beta}$ can be defined as

$$\text{argmax}_{\boldsymbol{\beta}} E \sum_{l,t} \log \left[\{\text{expit}(X_{l,t}^T \boldsymbol{\beta})\}^{I(T^{\bar{a}^l} > t)} \{1 - \text{expit}(X_{l,t}^T \boldsymbol{\beta})\}^{I(T^{\bar{a}^l} \leqslant t)} \right],$$

i.e. the maximum log-likelihood for the logistic model with marginal mean specification $\text{expit}(X_{l,t}^T \boldsymbol{\beta})$.

We are interested in estimation of $\boldsymbol{\beta}$, which can be implicitly written as a function of the parameters $\mathbf{S} = (S_{\bar{a}^l}(t), \text{ for all unique values of } \bar{a}_t^l)$ through the score equation:

$$0 = U(\mathbf{S}, \boldsymbol{\beta}) \equiv \sum_{l,t} X_{l,t} \left\{ S_{\bar{a}^l}(t) - \text{expit}(X_{l,t}^T \boldsymbol{\beta}) \right\}; \quad S_{\bar{a}^l}(0) = 1.$$

The functional delta method requires the derivation of the components,

$$\frac{dU(\mathbf{S}, \boldsymbol{\beta})}{d\boldsymbol{\beta}} = -\sum_{l,t} X_{l,t} X_{l,t}^T \frac{\exp(X_{l,t}^T \boldsymbol{\beta})}{\{1 + \exp(X_{l,t}^T \boldsymbol{\beta})\}^2},$$

a matrix with dimension $R \times R$, and

$$\frac{dU(\mathbf{S}, \boldsymbol{\beta})}{dS_{\bar{a}^l}(t)} = X_{l,t},$$

a column vector of length R. The two above components can be numerically evaluated and combined to form a column vector of length R using Equation (2).

The efficient influence function can be derived by combining Equation (2) with Equation (1) and simplifying slightly:

$$D_{\boldsymbol{\beta}} = \left[\sum_{l,t} \frac{\exp(X_{l,t}^T \boldsymbol{\beta})}{\{1 + \exp(X_{l,t}^T \boldsymbol{\beta})\}^2} X_{l,t} X_{l,t}^T \right]^{-1} \sum_{l,t} X_{l,t} D_{\bar{a}^l,t}.$$

Since the influence curve $D_{\bar{a}^l,t}$ can be numerically evaluated for each of the n subjects, we obtain a matrix of dimension $n \times R$, representing the joint influence components for $\boldsymbol{\beta}$.

Treating each of the estimated values of $S_{\bar{a}^l}(t)$ as an outcome vector (so that there are M "observations", one for each unique exposure pattern and time), fit a logistic regression with

a chosen linear specification and a logit link. This will produce the point estimate of $\boldsymbol{\beta}$. To obtain variance estimates, fit the efficient influence curve for $\boldsymbol{\beta}$ for each subject by estimating each of the components as described in Section 4 and combining them as indicated. Then, for each of the R columns of the resulting matrix the empirical variance is the estimated variance for the corresponding MSM coefficient estimate of $\boldsymbol{\beta}$.

5. Simulation study

5.1 *Data generation*

We generated data of the form $(W, A_1, L_1, S_1, ..., A_5, L_5, S_5)$ using known data generating functions. W is a continuous baseline confounder, $A_t, t = 1, ..., 5$ are the binary exposure variables and $S_t, t = 1, ..., 5$ are the survival indicators at each time point. The exposure generated was monotone (once exposed, always exposed). L_t is a binary variable that acts as a time-varying confounder. Each variable (unless determined by the monotonicity of exposure and survival) was generated dependent on the baseline and the covariate values at the previous time point according to the general rule that exposure reduces the probability of survival at the next time point as do higher values of L_t. Censoring was not included in the simulation study.

We used the following function (written in R Statistical Software version 2.13.2, R Development Core Team 2011) to generate the data:

```
data_surv_new<-function(i,ssize){
set.seed(i*5436)

W<-rnorm(n=ssize)/4+1

#TP1

p1<-expit(-4.2+2.5*W)
A1<-rbinom(n=ssize,size=1,prob=p1)
```

```
mu1<-expit(1+W+0.5*A1)
L1<-rbinom(n=ssize,size=1,prob=mu1)

s1<-expit(2+W-0.7*L1-0.5*A1)
S1<-rbinom(n=ssize,size=1,prob=s1)

#TP2

A2<-rep(0,length=ssize)
p2<-expit(-3.2+1*W+1.2*L1)
A2[A1==0&S1==1]<-rbinom(n=sum(A1==0&S1==1),size=1,prob=p2[A1==0&S1==1])
A2[A1==1&S1==1]<-1

mu2<-expit(1+L1+0.5*A2)
L2<-rbinom(n=ssize,size=1,prob=mu2)

S2<-rep(1,ssize)
s2<-expit(1.6+W-0.7*L2-0.5*A2)
S2[S1==1]<-rbinom(n=sum(S1==1),size=1,prob=s2[S1==1])
S2[S1==0]<-0

#TP3

A3<-rep(0,length=ssize)
p3<-expit(-2.9+1*W+1.2*L2)
A3[A2==0&S2==1]<-rbinom(n=sum(A2==0&S2==1),size=1,prob=p3[A2==0&S2==1])
A3[A2==1&S2==1]<-1

mu3<-expit(1+L2+0.5*A3)
L3<-rbinom(size=1,prob=mu3,n=ssize)

S3<-rep(0,ssize)
s3<-expit(2.5+0.8*W-0.7*L3-0.5*A3)
S3[S2==1]<-rbinom(n=sum(S2==1),size=1,prob=s3[S2==1])
S3[S2==0]<-0

#TP4

A4<-rep(0,length=ssize)
p4<-expit(-2+0.5*W+1.2*L3)
A4[A3==0&S3==1]<-rbinom(n=sum(A3==0&S3==1),size=1,prob=p4[A3==0&S3==1])
```

```
A4[A3==1&S3==1]<-1

mu4<-expit(1+L3+0.5*A4)
L4<-rbinom(prob=mu4,size=1,n=ssize)

S4<-rep(0,ssize)
s4<-expit(1.2+W-0.7*L4-0.5*A4)
S4[S3==1]<-rbinom(n=sum(S3==1),size=1,prob=s4[S3==1])
S4[S3==0]<-0

#TP5

A5<-rep(0,length=ssize)
p5<-expit(-1+0.5*W+1.2*L4)
A5[A4==0&S4==1]<-rbinom(n=sum(A4==0&S4==1),size=1,prob=p5[A4==0&S4==1])
A5[A4==1&S4==1]<-1

mu5<-expit(1+L4+0.5*A5)
L5<-rbinom(prob=mu5,size=1,n=ssize)

S5<-rep(0,ssize)
s5<-expit(1+0.5*W-0.7*L5-0.5*A5)
S5[S4==1]<-rbinom(n=sum(S4==1),size=1,prob=s5[S4==1])
S5[S4==0]<-0

#If dead, make missing
A2[S1==0]<-NA
L2[S1==0]<-NA
A3[S2==0]<-NA
L3[S2==0]<-NA
A4[S3==0]<-NA
L4[S3==0]<-NA
A5[S4==0]<-NA
L5[S4==0]<-NA


return(as.data.frame(cbind(W,L1,L2,L3,L4,L5,A1,A2,A3,A4,A5,S1,S2,S3,S4,S5)))

}
```

5.2 *Methods*

The TMLE method to estimate the parameters of a marginal structural model (MSM) for the hazard as described in the main manuscript was evaluated in its ability to predict $S^{\bar{a}=1}(5)$, the probability of survival at the fifth time point under the counterfactual condition of having all subjects exposed by the first time point (and thereafter, since once exposed means always exposed in our setting). The TMLE was compared to the Adjusted Kaplan-Meier Estimator (AKME), the inverse probability of treatment weighting method for the Kaplan-Meier curve described in Xie and Liu (2005). Both methods were implemented using logistic regressions to estimate all probabilities. The standard error for TMLE was estimated using its efficient influence function, and for AKME using the non-parametric bootstrap. The non-parametric bootstrap was performed by taking 500 resampled data sets with replacement from the complete data set. Each resampled data set was the same size as the original. The standard error was found by taking the standard deviation of the estimates calculated from the resampled data sets. The 95% confidence intervals for TMLE were estimated using the Normal approximation and the standard error from the efficient influence function. The confidence intervals for AKME used the 2.5th and 97.5th quantiles of the estimates from 500 bootstrap resamples.

Because of the way the data were generated, the models for each of the $Q_t^{\bar{a}}(j)$'s in the TMLE procedure were always misspecified (even when they included the correct set of confounders). It was possible to correctly specify the exposure model, so the unbiasedness of the TMLE in this simulation study is a result of the method's double-robustness.

5.3 *Results*

In the simulation study, 1,000 data sets were drawn with sample sizes 2,500 and 5,000. AKME and TMLE were both implemented so that the exposure models were correctly specified. Table 1 (top) shows the simulation results for the estimation of the counterfactual

probability of survival at the final time point under a history of always being exposed. For both sample sizes, AKME and TMLE perform very similarly, both producing unbiased estimates, and identical mean-squared errors (MSE) and standard errors (SE). Coverage was also close to 95% for both methods.

[Table 1 about here.]

TMLE was then evaluated in its ability to estimate the parameters of a marginal structural model for the hazard (see the main manuscript for the procedure). The model evaluated was $\text{logit}\lambda_{\bar{a}^l}(t) = \beta_0 + \beta_1 \text{cum}(\bar{a}_t^l) + \beta_2 t$ where β_1, the coefficient for the cumulative number of past times exposed, was the parameter of interest. Since the data was not generated from this MSM, the β parameters represent a likelihood projection of the survival probabilities at each time point onto a linear model (Neugebauer and van der Laan, 2007). TMLE was compared to the IPTW method for fitting the hazard MSM described in Hernán et al. (2000), which estimates an identical parameter (see Section 2). The IPTW was fit with unstabilized weights and its standard error was estimated using the nonparametric bootstrap (with the same specifications as for AKME). The MSM results in Table 1 (bottom) indicate that while both methods produced unbiased inference, for the lower sample size, TMLE had a slightly higher estimated standard error, resulting in slightly inflated confidence intervals and coverage. The standard errors for $n = 5,000$ coincided for the two methods.

The slightly higher standard errors obtained for TMLE when compared to IPTW can be explained by the different methods used to estimate the variance. For IPTW, we used nonparametric bootstrap resampling. For TMLE, we used the influence curve-based sandwich estimator, which is known to be conservative for misspecified Q-models (van der Laan and Rose, 2011).

References

Hernán, M. A., Brumback, B., and Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* **11,** 561–570.

Kang, J. D. Y. and Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science* **22,** 523–539.

Neugebauer, R. and van der Laan, M. (2007). Nonparametric causal effects based on marginal structural models. *Journal of Statistical Planning and Inference* **137,** 419–434.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rosenblum, M. and van der Laan, M. J. (2010). Simple examples of estimating causal effects using targeted maximum likelihood estimation. *U.C. Berkeley Division of Biostatistics Working Paper Series* .

Tsiatis, A. A. (2006). *Semiparametric Theory and Missing Data*. Springer Series in Statistics. Springer.

van der Laan, M. J. and Robins, J. M. (2003). *Unified Methods for Censored Longitudinal Data and Causality*. Springer Series in Statistics. Springer Verlag: New York.

van der Laan, M. J. and Rose, S. (2011). *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer Series in Statistics. Springer.

van der Laan, M. J. and Rubin, D. (2006). Targeted maximum likelihood learning. *The International Journal of Biostatistics* **2,** Article 11.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer Series in Statistics. Springer.

Xie, J. and Liu, C. (2005). Adjusted Kaplan-Meier estimator and log-rank test with inverse

probability of treatment weighting for survival data. *Statistics in Medicine* **24,** 30893110.

Received October 2000. *Revised February* 2000.

Accepted March 2000.

Table 1

Simulation results for (top) the probability of survival at time five under always-exposed, and (bottom) the coefficient of cumulative exposure in the hazard model (β_1). Correct exposure model used. Estimates taken over 1,000 generated datasets. True value for survival = 0.274; true value for MSM = 0.099

Method	Bias	MSE	rSE	% Coverage
Survival				
		$n = 2500$		
TMLE	<0.001	0.001	0.154	94.3
AKME	<0.001	0.001	0.154	94.7
		$n = 5000$		
TMLE	<0.001	<0.001	0.130	96.0
AKME	<0.001	<0.001	0.129	95.9
MSM				
		$n = 2500$		
TMLE	<0.001	<0.001	0.166	96.3
IPTW-MSM	<0.001	<0.001	0.160	94.4
		$n = 5000$		
TMLE	<0.001	<0.001	0.140	96.2
IPTW-MSM	<0.001	<0.001	0.134	94.5