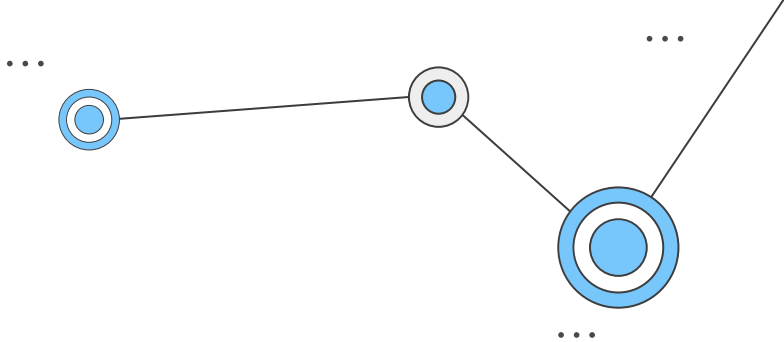
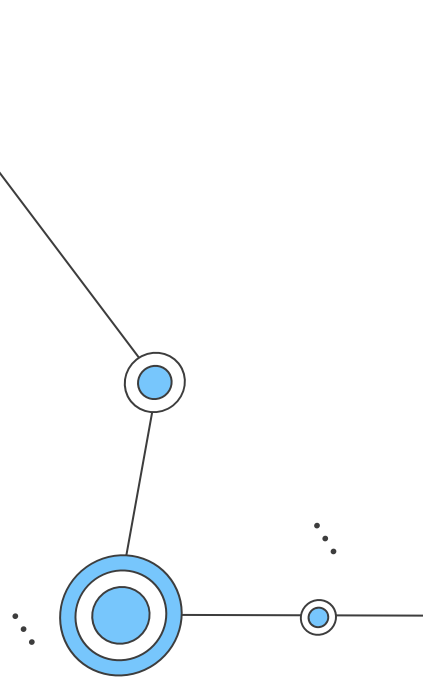


Marginal Structural Model

HYEMIN CHO
MINKYEONG KIM
HYESUNG CHOI
SUNGKYUNKWAN UNIVERSITY
MAY 28 2023

Table of Contents



01
Introduction

Marginal Structural Model (MSM)

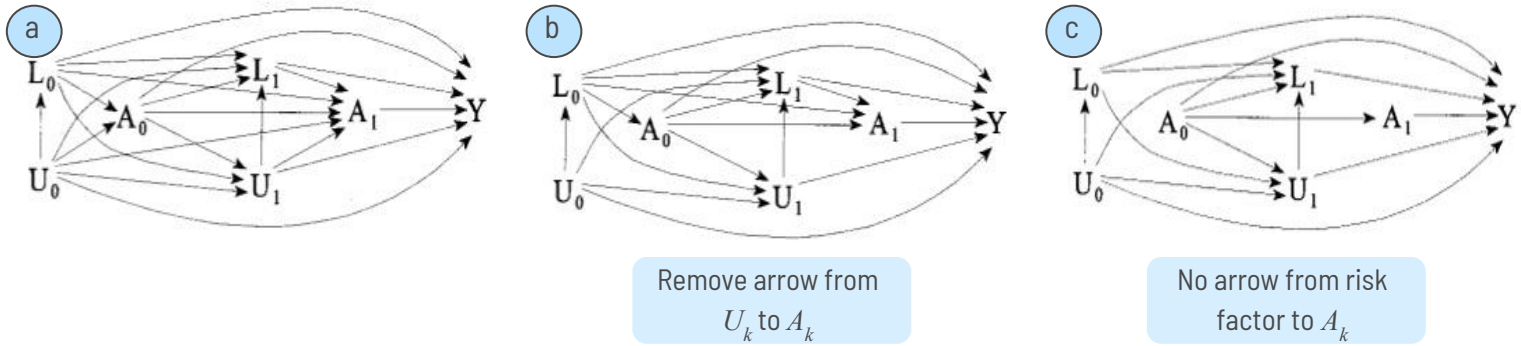
- Models for estimation of the causal effect of a time-dependent treatment in the presence of time-dependent covariates, from observational data.
- The parameters of a MSM are consistently estimated using IPTW.
- The usual approach to estimate the effects is to model the probability of disease as a function of past treatment and past confounder "history".
- The standard approaches may be biased, whether or not one further adjusts for past confounder history, when:
    - There exists a time-dependent covariate that is a risk factor for, or predictor of, the event of interest and also predicts subsequent treatment.
    - Past treatment history predicts subsequent level of the covariate.
- These conditions will always hold when there are time-dependent covariates that are simultaneously confounders and intermediate variables.

02
Theoretical Background

Time-Dependent Confounding

- Goal : Estimate the causal effect of the time-dependent treatment A_k on the outcome Y.
- Variables to consider:
  - A_k: Treatment of interest on the k th day since the start of follow-up.
  - Y: Binary outcome of interest measured at the end of follow-up on day $K+1$.
  - L_k : Vector of all measured risk factors for the outcome on kth day.
  - U_k : All unmeasured causal risk factors for Y on kth day.

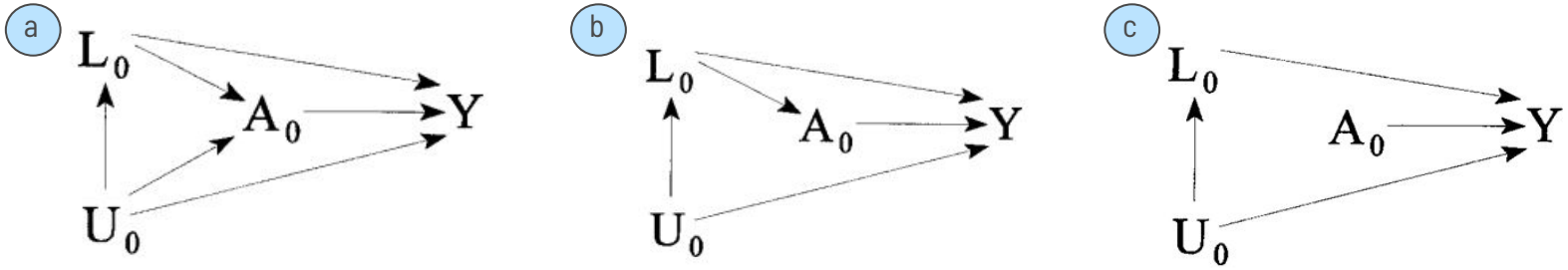
Figure 1. (K=1)



Time-Dependent Confounding

- Figure 2 contains the analogs of Figure 1, for a point-treatment study.
- Unable to determine from the observed data $L_{k'}$, $A_{k'}$ and Y whether there is confounding by unmeasured risk factors.
- Under the untestable assumption that there is no unmeasured confounding given the $L_{k'}$ we can, however, empirically test from the data whether treatment is unconfounded.

Figure 2. ($K = 0$)



Counterfactuals in Point-Treatment Studies

- Suppose A_0 is binary and Figure 2 (c) is the true causal graph.
- There are neither measured nor unmeasured covariates that confound the relation between the treatment and the outcome.
- The crude Risk Difference (RD), Risk Ratio (RR), and Odds Ratio (OR) measure the causal effect.
  - $crude\, RD = pr[Y = 1 \mid A_0 = 1] - pr[Y = 1 \mid A_0 = 0]$
  - $crude\, RR = pr[Y = 1 \mid A_0 = 1] / pr[Y = 1 \mid A_0 = 0]$
  - $crude\, OR = pr[Y = 1 \mid A_0 = 1] pr[Y = 0 \mid A_0 = 0] / pr[Y = 1 \mid A_0 = 0] pr[Y = 0 \mid A_0 = 1]$

Counterfactuals in Point-Treatment Studies

- The causal contrasts that correspond to these associational parameters involve counterfactual variables.
  - $Y_{a0} = 1$: subject's outcome if treated.
  - $Y_{a0} = 0$: subject's outcome if left untreated.
  - No subjects are observed for both $Y_{a0} = 1$ and $Y_{a0} = 0$. (i.e. If a subject is treated ($A_0 = 1$), the subject's observed outcome Y (or Y_{obs}) is equal to $Y_{a0} = 1$, and $Y_{a0} = 0$ is unobserved.
- If A_0 is unconfounded, the crude RD, RR, OR equals to the causal ones respectively.
  - $causal\ RD = pr[Y_{a_0=1} = 1] - pr[Y_{a_0=0} = 1]$
  - $causal\ RR = pr[Y_{a_0=1} = 1]/pr[Y_{a_0=0} = 1]$
  - $causal\ OR = pr[Y_{a_0=1} = 1]pr[Y_{a_0=1} = 0]/pr[Y_{a_0=0} = 1]pr[Y_{a_0=1} = 0]$

Models for Point-Treatment Studies

- The causal RD, RR, OR can be expressed in terms of the parameters of the following linear, log linear, and linear logistic models for the two counterfactual probabilities $\text{pr}[Y_{a0=1} = 1]$ and $\text{pr}[Y_{a0=0} = 1]$:

$$pr[Y_{a_0} = 1] = \psi_0 + \psi_1 a_0 \quad (1)$$

$$\log pr[Y_{a_0} = 1] = \theta_0 + \theta_1 a_0 \quad (2)$$

$$\text{logit } pr[Y_{a_0} = 1] = \beta_0 + \beta_1 a_0 \quad (3)$$

$$where\, causal\, RD = \psi_1,\ causal\, RR = \exp(\theta_1),\ and\, causal\, OR = \exp(\beta_1)$$

- These models are (1) Saturated, (2) Marginal, (3) Structural, and (4) does not include covariates. Therefore, they called saturated MSMs.

Models for Point-Treatment Studies

- The crude RD, RR, and OR can also be expressed in terms of the parameters of the following saturated linear, log linear, and linear logistic models for the observed outcome Y.

$$pr[Y = 1 \mid A_0 = a_0] = \psi_0' + \psi_1' a_0 \tag{4}$$

$$\log pr[Y = 1 \mid A_0 = a_0] = \theta_0' + \theta_1' a_0 \tag{5}$$

$$\text{logit } pr[Y = 1 \mid A_0 = a_0] = \beta_0' + \beta_1' a_0 \tag{6}$$

$$where\, crude\, RD = \psi_1',\ crude\, RR = \exp(\theta_1'),\ and\, crude\, OR = \exp(\beta_1')$$

- Models for associations observed when comparing subpopulations (defined by levels of treatment) of the source population.
- The parameters of the associational models 4 – 6 will differ from the parameters of the MSMs 1 – 3, except when treatment is unconfounded.
- When treatment is unconfounded, these estimates will also be unbiased for the corresponding causal parameters of models 1 – 3.

No Unmeasured Confounders

- If the treatment is confounded (Figure 2 (a) & (b)), the crude association parameter will not be equal to the corresponding causal parameter.

- However, assuming we have no unmeasured confounders given data on $L_{0'}$ unbiased estimates of the causal parameters can be obtained by performing a weighted analysis. Each subject i is assigned a weight w_i :

$$w_i = \frac{1}{pr[A_0 = a_0 \mid L_0 = l_{0i}]}$$

No Unmeasured Confounders

- True weight is unknown but it can be estimated with the logistic regression model of A_0 on L_0.

$$\text{logit}\, pr[A_0 = 1 \mid L_0 = l_0] = \alpha_0 + \alpha_1 l_0$$
$$\rightarrow\ pr[A_0 = 1 \mid L_0 = l_{0i}] = \frac{\exp(\hat{\alpha_0} + \hat{\alpha_1} l_{0i})}{1 + \exp(\hat{\alpha_0} + \hat{\alpha_1} l_{0i})} \qquad (7)$$

- Weighting creates pseudo population consisting of w_i copies of each subject
- Two properties of pseudo population:
  - A_0 is unconfounded by L_0.
  - $pr[Y_{a0=1} = 1]$ and $pr[Y_{a0=0} = 1]$ are the same as in the true study population.

Unmeasured Confounding

- In the presence of U_0, unbiased estimates of the causal RD, RR, and OR can be estimated by using the weight:

$$w_i = \frac{1}{pr[A_0 = a_0 \mid L_0 = l_{0i},\, U_0 = u_{0i}]}$$

- However, U_0 are not observed and thus, it is not possible to estimate these weights unbiasedly.
- Unbiased estimation by any method is impossible in the presence of unmeasured confounding factors without strong additional assumptions.

Multilevel Treatment and Unsaturated MSMs

- Suppose that treatment is unconfounded(Figure 2 (c)) and is an ordinal variable.
- A subject has a separate counterfactual variable for each of the possible treatments a_0.
- Due to the large number of potential outcome variables $Y_{a0'}$ it is difficult to perform a saturated analysis.
- By specifying a linear logistic MSM such as

$$\text{logit } pr[Y_{a_0} = 1] = \beta_0 + \beta_1 a_0, \tag{8}$$

parsimonious dose-response relationship can be assumed.

- MSM model 8 is contrasted with the following linear logistic association model for the observed data:

$$\text{logit } pr[Y = 1 \mid A_0 = a_0] = \beta_0' + \beta_1' a_0 \tag{9}$$

Multilevel Treatment and Unsaturated MSMs

- For example, similar to the saturated model, one might specify polytomous logistic model. If possible values of A_0 in Figure 2 (a) are 0, 1, ..., p, then the denominator of the weight is :

$$pr[A_0 = a_0 \mid L_0 = l_0] = \begin{cases} \frac{\exp(\alpha_{0a_0} + \alpha_1 l_0)}{1 + \sum_{j=1}^{p} \exp(\exp(\alpha_{0j} + \alpha_1 l_0))} & , a_0 = 1, 2, \ldots, p \\ \frac{1}{1 + \sum_{j=1}^{p} \exp(\exp(\alpha_{0j} + \alpha_1 l_0))} & , a_0 = 0 \end{cases} \quad (10)$$

Stabilized Weights

- $pr[A_0 = a_0 | L_0 = l_{0i}]$ may vary greatly between subjects when components of L_0 are strongly associated with A_0.
- For saturated MSM, this problem is unavoidable but for unsaturated MSMs, variability can be mitigated by replacing w_i by the stabilized weight sw_i:

$$sw_i = \frac{pr[A_0 = a_{0i}]}{pr[A_0 = a_{0i} \mid L_0 = l_{0i}]} \qquad (*)$$

- Estimates of the parameters β of MSM remain unbiased when sw_i is used.
- In the case of an unsaturated MSM, estimates of β will generally be less variable. For saturated MSMs, the variability will be the same whether we use the stabilized or unstabilized weights.

Stabilized Weights

- The numerator of the stabilized weight is:

$$\Pr[A_0 = a_0] = \begin{cases} \dfrac{e^{\left(\alpha_{0a_0}*\right)}}{\left\{1+\sum_{j=1}^{p} e^{\left(\alpha_{0j}*\right)}\right\}} & , a_0 = 1, \ldots, p \\ \dfrac{1}{\left\{1+\sum_{j=1}^{p} e^{\left(\alpha_{0j}*\right)}\right\}} & , \ a_0 = 0 \end{cases} .$$

- If treatment is unconfounded, this parameter will be the same with α_0 of model (10).

Continuous Treatment

- Suppose that the treatment is continuous and greater than or equal to 0. Assume that no subject has treatment near 0, and that we can model the distribution of A_0 as normal.
- Now each subject has an extremely large number of counterfactual outcomes Y_{a0}.
- One can still obtain unbiased estimates of the causal parameter β_1 of model 8 by fitting the logistic model 9 if one uses the $sw_i = f(a_{0i}) / f(a_{0i} \mid l_{0i})$.
    - To estimate $f(a_{0i}|l_{0i})$, given L_0, A_0 is normal with mean $\alpha_0 + \alpha_1 L_0$ and variance σ^2. The unbiased estimates can be obtained by OLS of A_0 on L_0.
    - To estimate $f(a_{0i})$, A_0 is normal with mean α_0* and variance σ^{*2}. The unbiased estimates $\widehat{\alpha_0^*}$ and $\widehat{\sigma_0^{*2}}$ are respectively the average of the observed A_0s and their empirical variance.
- If we use unstabilized weights, the estimates have infinite variance and thus cannot be used.

Time-Dependent Treatments

- Let the variables be:
  - $\bar{A}_k = (A_0, A_1, \ldots, A_k)$: Treatment or exposure history through day k and let $\bar{A} = \bar{A}_k$.
  - $\bar{L} = \bar{L}_k$: History vectors of measured risk factors.
  - $Y_{\bar{a}}$: Value of Y that would have been observed had all subjects received dose history $\bar{a} = (a_0, a_1, \ldots, a_k)$ rather than their observed dose history $\bar{A}$.
- It may not be possible to estimate a saturated MSM. therefore assume some sort of parsimonious dose-response relationship by specifying a linear logistic MSM such as

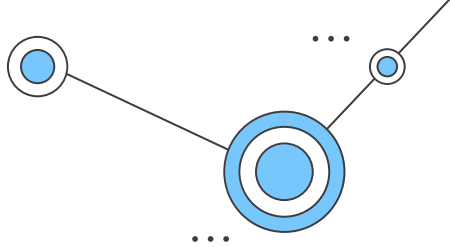
$$\text{logit } pr[Y_{\bar{a}} = 1] = \beta_0 + \beta_1 \text{cum}(\bar{a}) \qquad (12)$$

- Contrasting MSM 12 with the following linear logistic model for the observed data :

$$\text{logit } pr\left[Y = 1 \mid \bar{A} = \bar{a}\right] = \beta_0' + \beta_1' \text{cum}(\bar{a}) \qquad (13)$$

- Assuming no loss to follow-up selection bias or measurement error, the parameters β'_1 can be unbiasedly estimated by fitting the linear logistic model 13.
- If the treatment is unconfounded, the parameters of models 12 and 13 are equal.

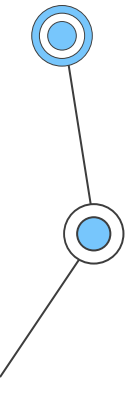
Time-Dependent Treatments



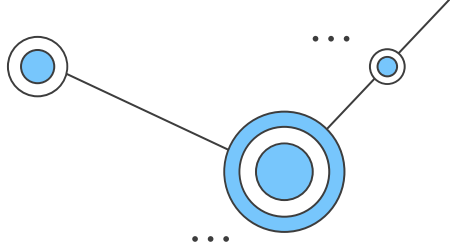
- When treatment is confounded and if there is no unmeasured confounder given the $L_{k'}$ then one can still obtain unbiased estimates of the causal parameter β_1 of model 12 by fitting the logistic model 13 with the stabilized weights :

$$sw_i = \frac{\left\{\prod_{k=0}^{K} pr\left(A_k = a_{ki} | \overline{A}_{k-1} = \overline{a}_{(k-1)i}\right)\right\}}{\left\{\prod_{k=0}^{K} pr\left(A_k = a_{ki} | \overline{A}_{k-1} = \overline{a}_{(k-1)i}, \overline{L}_k = \overline{l}_{ki}\right)\right\}}. \quad (14)$$

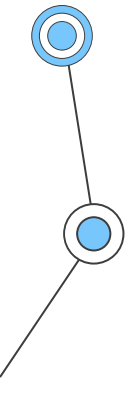
- When $K = 0$, models 12 and 13 reduce to our previous models 8 and 9 and sw_i 14 reduces to our previous sw_i.
- With time-dependent treatments the variation in the unstabilized weights will often be enormous, resulting with the estimator of β being highly variable with a markedly non-normal sampling distribution.



Bias Induced by Controlling for a Variable Affected by Treatment



- Standard regression methods adjust for measured covariates by including them in the model as regressors.
- Standard methods may fail to adjust appropriately for confounding due to measured confounders L_k when treatment is time varying.
    - L_k may be a confounder for later treatment and thus must be adjusted.
    - It may also be affected by earlier treatment and thus should not be adjusted for by standard methods.
- A solution to this conundrum is to adjust for the time-dependent covariates L_k by using them to calculate the weights sw_i rather than by adding the L_k to the regression model as regressors.



Estimation of weights

- Assume that A_k is binary. Considering model (14), pooled logistic model is used to estimate the unknown probability. For example :
  - Denominator :

$$\begin{aligned}&\text{logit } pr[A_k = 1 \mid A_{k-1} = a_{k-1},\ L_k = l_k] \\ &= \alpha_0 + \alpha_1 k + \alpha_2 a_{k-1} + \alpha_3 a_{k-2} + \alpha_4 l_k + \alpha_5 l_{k-1} + \alpha_6 a_{k-1} l_k + \alpha_7 l_0 \end{aligned} \tag{15}$$

  - Numerator :

$$\text{logit } pr[A_k = 1 \mid A_{k-1} = a_{k-1}] = \alpha_0^* + \alpha_1^* k + \alpha_2^* a_{k-1} + \alpha_3^* a_{k-2} \tag{16}$$

- With the estimated predicted values $\hat{p}_{0i}, \ldots, \hat{p}_{Ki}$ from model (15) and $\hat{p}^*_{0i}, \ldots, \hat{p}^*_{Ki}$ from model (16), sw_i is estimated by

$$sw_i = \frac{\prod_{k=0}^{K} \left(\hat{p}^*_{ki}\right)^{a_{ki}} \left(1 - \hat{p}^*_{ki}\right)^{1-a_{ki}}}{\prod_{k=0}^{K} \left(\hat{p}_{ki}\right)^{a_{ki}} \left(1 - \hat{p}_{ki}\right)^{1-a_{ki}}} \tag{17}$$

Effect Modification by Pretreatment Covariates

- MSMs can be generalized to allow one to include pretreatment covariates.

$$\text{logit } pr[Y_{\bar{a}} = 1 \mid V = v] = \beta_0 + \beta_1 \text{cum}(\bar{a}) + \beta_2 v + \beta_3 \text{cum}(\bar{a}) v \qquad (18)$$

  $\beta_1 + \beta_3 v$ represents the effect of cumulative treatment on a linear logistic scale within level v of the baseline variable V.

- Unbiased estimates of the parameters of model 18 is obtained under the assumption of no unmeasured confounders by fitting an association model such as

$$\text{logit } pr\left[Y = 1 \,\middle|\, \bar{A} = \bar{a},\, V = v\right] = \beta_0 + \beta_1 \text{cum}(\bar{a}) + \beta_2 v + \beta_3 \text{cum}(\bar{a}) v \qquad (19)$$

  with the estimated weights sw_i of equation 17, modified only in that p_k* is now the estimated predicted value from the fit :

$$\text{logit } pr\left[A_k = 1 \,\middle|\, \bar{A}_{k-1} = \bar{a}_{k-1},\, V = v\right] = \alpha_0^* + \alpha_1^* k + \alpha_2^* a_{k-1} + \alpha_3^* a_{k-2} + \alpha_4^* v \qquad (20)$$

- MSMs cannot be used to model the interaction of treatment with a time-varying covariate.

Censoring by Loss to Follow-up

- Assume that there is censoring by loss to follow-up and that once a subject is lost to follow-up, he/she is not able to re-enter. Let $C_k = 1$ if a subject was lost to follow-up by day k, and $C_k = 0$ otherwise.
- Censoring is conceptualized as just another time-varying treatment.
- Variable C_k is added in between L_k and A_{k-1} at each timepoint k in figure 1. In order to do this, assumption of no unmeasured confounders for treatment and censoring is used.
- The required subject-specific weight is $sw_i \times sw_i^{\dagger}$, where

$$sw_i^{\dagger} = \frac{\prod_{k=0}^{K+1} pr\left[C_k = 0 \,\middle|\, \bar{C}_k = 0,\ \bar{A}_{k-1} = \bar{a}_{(k-1)i}\right]}{\prod_{k=0}^{K+1} pr\left[C_k = 0 \,\middle|\, \bar{C} = 0,\ \bar{A}_{k-1} = \bar{a}_{(k-1)i},\ \bar{L}_k = \bar{l}_{ki}\right]}$$

Limitations of MSM

- IPTW estimators will be biased and thus MSMs should not be used in studies in which at each time k there is a covariate level l_k such that all subjects with that level of the covariate are certain to receive the identical treatment a_k
- Nevertheless, g-estimation of structural nested models can always be used to estimate the exposure effects, even in studies in which MSMs cannot be used.
- In many studies, based on substantive considerations, the above difficulty does not occur and MSMs are practical methods.

03
Application to Real Data

Multicenter AIDS Cohort Study (MACS)

- Goal: Estimate the effect of zidovudine on the survival of human immunodeficiency virus (HIV) - positive men.
- Homosexual and bisexual men from Los Angeles, Washington, Pittsburgh, and Chicago enrolled in this study.
- Two requirements needed to be enrolled:
    - No prior acquired immunodeficiency syndrome (AIDS) defining illness.
    - Not on antiretroviral therapy at the first eligible visit .
- Return every 6 months to complete a questionnaire, undergo physical examination, and provide blood samples.
- Follow-up ended when : study visit 21 / October 1994 / deaths / 24 months after the last visit.

Multicenter AIDS Cohort Study (MACS)

- Analysis focuses on 2178 men who attended at least on visit between visits 5 and 21.
- By the end of the follow up, 1296 men had initiated zidovudine treatment and 750 died.
- Robinson shown that the usual time-dependent Cox proportional hazards model approach may be biased due to the time-dependent confounders and past exposure history predicting the risk factors.
- Therefore, Marginal Structural Cox proportional hazards model introduced.

Marginal Structural Cox Proportional Hazards Model

- Notations:
  - T : Subject's time of death with time measured in months since the start of follow up.
  - $A(t)$: 1 if subject was on zidovudine at time t and 0 otherwise.
  - $\bar{A}(t)$: Covariate history (eg. $\{A(u);\ 0 \leq u < t\}$ is a subject's treatment history up to t).
  - $L(t)$: Time-dependent covariates $\rightarrow$ "before t".
  - V : Vector of time-dependent baseline covariates measured before the start of follow-up.
- Assumptions made for simplicity:
  - Patients remain on therapy once they start it and that the hazards of death at time t depend on a subject's zidovudine history only through its current value.

Marginal Structural Cox Proportional Hazards Model

- Conditional Hazard (mortality rate) given the treatment history and baseline covariates:

$$\lambda_T\left(t \mid \bar{A}(t), V\right) = \lambda_0 \exp\left(\gamma_1 A(t) + \gamma_2 V\right)$$

- After implementing weights, MSM Conditional Hazard given the treatment history and baseline covariates is created:

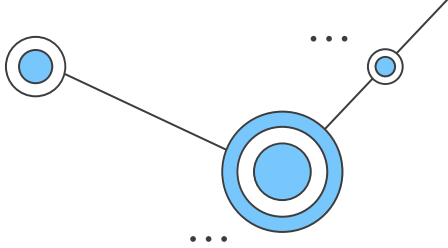
$$\lambda_{T_{\bar{a}}}(t \mid V) = \lambda_0 \exp\left(\beta_1 a(t) + \beta_2 V\right)$$

$\lambda_{T_{\bar{a}}}(t \mid V)$: hazard of death at t among subjects with baseline V in the source population.

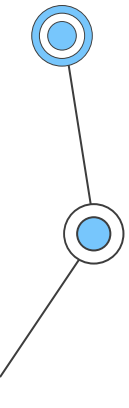
$\lambda_0(t)$: unspecified baseline hazard.

β_1 is the log of RR and β_2 is a row vector. Both of them are unknown parameters.

Marginal Structural Cox Proportional Hazards Model



- Causal interpretation : $exp(\beta_1)$ is the ratio of the mortality rate at any time t had all subjects been continuously exposed to zidovudine, compared with the mortality rate at time t had all subjects remain unexposed.
- β_1 is consistently estimated by IPTW, under the untestable assumption.
- If censoring exist, weighted Cox model can be fitted using the weights $sw_i(t) \times sw_i^{\dagger}(t)$.



Estimation of the Weights

- Denominator:
  - $$\bar{\hat{p}}_i(k) = \begin{cases} \prod_{u=0}^{k} \hat{p}_i(u) & \text{, if subject } i \text{ didn't start zidovudine up to month } k \\ [1 - \hat{p}_i(t)] \prod_{u=0}^{t-1} \hat{p}_i(u) & \text{, if subject } i \text{ started zidovudine at month } t \text{ for } t \leq k \end{cases}$$
  - Where $\hat{p}_i = \dfrac{\exp\left(\hat{\alpha_0}(k) + \hat{\alpha_1} L_i(k) + \hat{\alpha_2}(V_i)\right)}{1 + \exp\left(\hat{\alpha_0}(k) + \hat{\alpha_1} L_i(k) + \hat{\alpha_2}(V_i)\right)}.$
- Numerator: $\bar{\hat{p}}_i(k)$ without the covariate $L(k)$.
- For censored data, estimate the $sw_i^{\dagger}(k)$ same method as above using:

$$p_i(k) = \text{logit}\, pr\left[C(k) = 0 \,\middle|\, \bar{C}(k-1) = 0, \bar{A}(k-1) = 0, \bar{L}(k)\right] = \alpha_0'(k) + \alpha_1' L(k) + \alpha_2' V$$

Causal Effect of Zidovudine in MACS

Adjusted vs Unadjusted & Stabilized vs Nonstabilized

TABLE 1. Inverse-Probability-of-Treatment Weighted Estimates of the Causal Effect of Zidovudine Therapy on Mortality in the Multicenter AIDS Cohort Study

Unweighted estimates*	RR	95% CI
Unadjusted	3.55	2.95–4.27
Only baseline covariates	2.32	1.92–2.81

Weighted estimates†	RR	Valid 95% Conservative CI	Invalid Model-Based‡ 95% CI
Stabilized weights	0.74	0.57–0.96	0.62–0.87
Nonstabilized weights	0.76	0.54–1.05	0.71–0.80

RR = mortality rate ratio (zidovudine users vs nonusers); CI = confidence interval.

* Noncausal models, shown for comparison purposes only. The unadjusted model includes only the time-varying intercept and zidovudine use (yes or no). The model with baseline covariates includes also: age, calendar year (1985, 1986, 1987–89, or 1990–1993), CD4 (<200, 200–499, or ≥500/μl), CD8 (<500; 500–999; or ≥1,000 per μl), WBC (<3,000; 3,000–4,999; or ≥5,000 per μl), RBC (<35, 35–44, or ≥45 × 10^5 per μl), platelets (<150, 150–249, or ≥250 × 10^3 per μl), presence of symptoms (yes if fever, oral candidiasis, diarrhea, weight loss, oral hairy leukoplakia, or herpes zoster, or no if otherwise).

† Weights calculated as described in the text using data on baseline covariates plus most recent CD4, CD8, WBC, RBC (<30, 30–39, or ≥40 × 10^5 per μl), platelets, presence of symptoms, presence of AIDS-defining illness, and previous zidovudine use.

‡ The model-based intervals are not valid for weighted models because they fail to account for the within-subject covariances induced by weighting.

Estimated parameters and their robust errors

TABLE 2. Inverse-Probability-of-Treatment Weighted Estimates of the Parameters of a Marginal Structural Model for the Causal Effect of Zidovudine on Mortality in the Multicenter AIDS Cohort Study

Covariates*	Parameter Estimate	Robust Standard Error	Conservative 95% CL
Zidovudine	−0.301	0.132	−0.559, −0.043
Age at baseline	0.038	0.008	0.224, 0.053
Baseline CD4, per μl			
<200	1.913	0.136	1.645, 2.180
200–499	0.634	0.088	0.461, 0.806
≥500	0.000		
Baseline CD8, per μl			
<500	−0.648	0.122	−0.888, −0.409
500–99	−0.477	0.087	−0.648, −0.306
≥1000	0.000		
Baseline WBC, per μl			
<3000	0.790	0.085	0.375, 1.205
3000–4999	0.273	0.420	0.106, 0.440
≥5000	0.000		
Baseline RBC, ×10^5, per μl			
<35	0.231	0.420	−0.593, 1.055
35–45	0.254	0.105	0.048, 0.460
≥45	0.000		
Baseline platelets, ×10^3, per μl			
<150	0.616	0.129	0.364, 0.869
150–249	0.189	0.087	0.019, 0.359
≥250	0.000		
Presence of symptoms	0.618	0.086	0.448, 0.787
Calendar year at baseline			
1985	0.405	0.426	−0.431, 1.240
1986	0.440	0.432	−0.408, 1.287
1987–89	0.250	0.442	−0.617, 1.120
1990–93	0.000		

Causal Effect of Zidovudine in MACS

Center and dispersion parameters of the distribution of the four estimated probabilities at two arbitrary time points 24 & 84 months of follow-up

TABLE 3. Estimated Probability of Having One's Own Observed Treatment History [Estimated Denominator of $sw_i(t)$] and Censoring History [Estimated Denominator of $sw_i^{\dagger}(t)$] at 24 and 84 Months of Follow-Up, Multicenter AIDS Cohort Study

	Mean	SD	Median	IQR	Minimum	Maximum
24 Months (N = 2,063)						
Probability of having observed zidovudine history						
Given baseline covariates*	0.52	0.34	0.69	0.77	0.005	0.917
Given time-varying covariates†	0.54	0.35	0.68	0.82	0.002	0.939
Probability of being uncensored						
Given baseline covariates*	0.96	0.08	0.99	0.04	0.228	0.997
Given time-varying covariates†	0.96	0.08	0.99	0.04	0.199	0.997
84 Months (N = 836)						
Probability of having observed zidovudine history						
Given baseline covariates*	0.11	0.17	0.01	0.22	0.001	0.556
Given time-varying covariates†	0.15	0.24	0.02	0.26	0.001	0.728
Probability of being uncensored						
Given baseline covariates*	0.78	0.17	0.86	0.08	0.244	0.932
Given time-varying covariates†	0.78	0.17	0.86	0.08	0.256	0.940

SD = standard deviation, IQR = interquartile range.

* Age (years), calendar year (1985, 1986, 1987–1989, or 1990–1993), CD4 (<200, 200–499, or ≥500 per μl), CD8 (<500; 500–999; ≥1,000 per μl), WBC (<3,000; 3,000–4,999; or ≥5,000 per μl), RBC (<35, 35–44, or ≥45 × 10^5 per μl), platelets (<150, 150–249, or ≥250 × 10^3/μl), presence of symptoms (yes if fever, oral candidiasis, diarrhea, weight loss, oral hairy leukoplakia, or herpes zoster, or no otherwise), and previous zidovudine use.

† Baseline covariates plus most recent CD4, CD8, WBC, RBC (<35, 35–44, ≥45 × 10^5 per μl), platelets, presence of symptoms, or presence of an AIDS-defining illness.

Causal Effect of Zidovudine in MACS

Range of the weights estimated

24 months	Stabilized Weight	Non-stabilized Weight
Pr of Observed Zidovudine history	**(0.14 , 6.67)**	**(1.06 , 500)** = **(1/0.939 , 1/0.002)**
Pr of Uncensored (†)	(0.93 , 1.23)	(1.00 , 5.03) = (1/0.997 , 1/0.199)

Interpretation : In the pseudo population, some observations would be represented by 1.06 copies of themselves, whereas others would be represented by 500 copies.

$sw_i^{\dagger}(t)$ normalizes/stabilizes the range and increases the efficiency of the analysis by reducing the number of people in the pseudo population

→ narrower range

Causal Effect of Zidovudine in MACS

Stabilized vs Nonstabilized : $s\hat{w}_i(t) \times \hat{sw}_i^{\dagger}(t)$ vs $\hat{w}_i(t) \times \hat{w}_i^{\dagger}(t)$

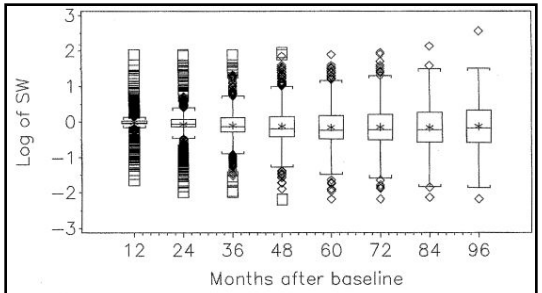


FIGURE 1. Distribution of stabilized weights SW. The box for each group shows the location of the mean (*), median (middle horizontal bar) and quartiles (border horizontal bars). Vertical lines extend to the most extreme observations which are no more than 1.5 × IQR beyond the quartiles. Observations beyond the vertical lines are plotted individually, if they lie within the limits of the frame.

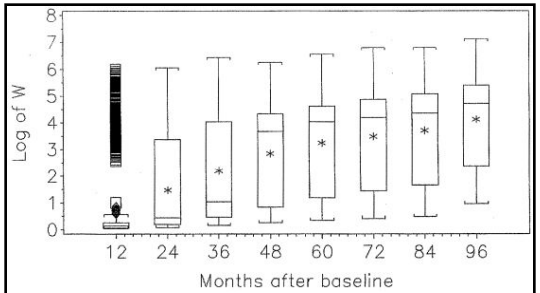


FIGURE 2. Distribution of nonstabilized weights SW. The box for each group shows the location of the mean (*), median (middle horizontal bar) and quartiles (border horizontal bars). Vertical lines extend to the most extreme observations which are no more than 1.5 × IQR beyond the quartiles. Observations beyond the vertical lines are plotted individually, if they lie within the limits of the frame.

- Distribution of weights:
  - Stabilized : weight symmetric and centered around 1 at all times, variance increases over time
  - Unstabilized : weights skewed, variance greatly exceeds that of the stabilized weights
- Weights were obtained by modeling the time-dependent intercept with natural cubic splines with 5 knots (months 23, 44, 71, 94, and 100 ⟶ percentiles 5, 27.5, 50, 72.5, and 95, respectively)

04
Discussion

Comparison of MSM with Previously Proposed Methods

- 4 different methods:
  - Iterative conditional expectations (ICE) estimator
  - Parametric g-computation algorithm formula
  - G-estimation of Structural Nested Models (SNM)
  - Marginal Structural Models (MSM)
- When (1) both treatment and the confounders are discrete variables, (2) they are measured at only a few time points, and (3) the study size is large, then estimation can be carried out using fully saturated models (that is, nonparametrically), and all four methods are precisely equivalent.
- ICE ⟶ rarely used as often lead to logically incompatible models
- G-computation ⟶ quite difficult to determine whether the CI for the treatment effect includes the null hypothesis of no effect whereas SNM and MSM do.

Comparison of MSM with Previously Proposed Methods

- SNM cannot be conveniently used to estimate the effect of treatment on dichotomous outcomes because it cannot be fitted by g-estimation whereas IPTW estimation of logistic MSM can.
- SNM do not resemble the standard models whereas MSM does. Close resemblance makes their application more intuitive for researchers and easier for programmers.
- Although MSM looks superior here, there are some drawbacks where SNM overrules MSM. Therefore, MSM prominently presents better result than ICE and parametric g-computation algorithm but it is still in question for SNM.

Conclusion

- From the data analysis, the more we adjust the time-dependent confounders, the more closer we get to the true effect of the treatment on the potential outcomes.

Adjustment Status	Mortality Rate
No Adjustment	**3.6** (erroneous effect of treatment)
Unweighted Standard Model, No Adjustment to Time-Dependent Confounders	**2.3** (detrimental effect of treatment)
Weighted and Time-Dependent Confounders Adjusted	**0.7** (increments the survival rate)

- Difference between the weighted and unweighted estimates is an indication of the amount of confounding due to the time-dependent prognostic factors.
- MSM is superior to standard statistical models when there exists time-dependent confounding by variables that are affected by previous exposure.

References

- Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 550-560.

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

THANK YOU