

# Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-Positive Men

Miguel Ángel Hernán,<sup>1</sup> Babette Brumback,<sup>2</sup> and James M. Robins<sup>1,2</sup>

Standard methods for survival analysis, such as the time-dependent Cox model, may produce biased effect estimates when there exist time-dependent confounders that are themselves affected by previous treatment or exposure. Marginal structural models are a new class of causal models the parameters of which are estimated through inverse-probability-of-treatment weighting; these models allow for appropriate adjustment for confounding. We describe the marginal structural Cox proportional hazards model and use it to estimate the causal effect of zidovudine on the survival of human immunodeficiency virus-positive men participating in the Multicenter AIDS Cohort Study. In this study, CD4 lymphocyte count is both a time-dependent confounder of the causal effect of

zidovudine on survival and is affected by past zidovudine treatment. The crude mortality rate ratio (95% confidence interval) for zidovudine was 3.6 (3.0-4.3), which reflects the presence of confounding. After controlling for baseline CD4 count and other baseline covariates using standard methods, the mortality rate ratio decreased to 2.3 (1.9-2.8). Using a marginal structural Cox model to control further for time-dependent confounding due to CD4 count and other time-dependent covariates, the mortality rate ratio was 0.7 (95% conservative confidence interval = 0.6-1.0). We compare marginal structural models with previously proposed causal methods. (Epidemiology 2000;11:561-570)

**Keywords:** counterfactuals, causality, epidemiologic methods, longitudinal data, survival analysis, structural models, confounding, intermediate variables, AIDS.

Marginal structural models (MSMs) can be used to estimate the causal effect of a time-dependent exposure in the presence of time-dependent confounders that are themselves affected by previous treatment.<sup>1,2</sup> The use of MSMs can be an alternative to *g*-estimation of structural nested models (SNMs).<sup>3</sup>

In our companion paper we describe inverse-probability-of-treatment weighted (IPTW) estimation of a marginal structural logistic model.<sup>4</sup> In this paper, we introduce the marginal structural Cox proportional hazards model, show how to estimate its parameters by inverse-probability-of-treatment weighting, provide practical advice on how to use standard statistical software to obtain the IPTW estimates, and include, as an appendix, the SAS code necessary for the analysis. We use this Cox proportional hazards MSM to estimate the effect of zidovudine on the survival of human immunodeficiency

virus (HIV)-positive men enrolled in an observational cohort study, the Multicenter AIDS Cohort Study (MACS). We conclude by comparing methods based on MSMs with previously proposed methods based on *g*-estimation of SNMs and on the direct estimation of the *g*-computation algorithm formula.

We now begin by describing the MACS and then summarize why standard methods for survival analysis are not appropriate for estimating the effect of zidovudine on mortality in this cohort.

## The Multicenter AIDS Cohort Study and Bias of Standard Methods

Between 1984 and 1991, the MACS enrolled 5,622 homosexual and bisexual men, with no prior acquired immunodeficiency syndrome (AIDS)-defining illness, from the metropolitan areas of Los Angeles, Baltimore-Washington, Pittsburgh, and Chicago. Study participants were asked to return every 6 months to complete a questionnaire, undergo physical examination, and provide blood samples. The design and methods of the MACS have been described in detail elsewhere.<sup>5,6</sup>

We restricted our cohort to HIV-positive men alive in the period during which zidovudine was available for use (that is, after study visit 5; March 1986 through March 1987). Follow-up ended at study visit 21, October 1994, death, or 24 months after the last visit, whichever came

From the Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA.

Address correspondence to Miguel Hernán, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115.

This research was supported by NIH grant R01-AI32475.

Submitted March 13, 1999; final version accepted February 28, 2000.

Copyright © 2000 by Lippincott Williams & Wilkins, Inc.

first. Our analysis included the 2,178 men who attended at least one visit between visits 5 and 21 while HIV positive, and who did not have an AIDS-defining illness and were not on antiretroviral therapy at the first eligible visit. By the end of the follow-up (median duration-69 months), 1,296 men had initiated zidovudine treatment and 750 had died.

The usual approach to the estimation of the effect of a time-varying exposure, such as zidovudine, on survival is to model the hazard of failure at a given time as a function of past exposure history using a time-dependent Cox proportional hazards model. Robins<sup>7</sup> has shown this approach may be biased, whether or not one further adjusts for past covariate history, whenever (1) there exists a time-dependent covariate that is both a risk factor for mortality and also predicts subsequent exposure and (2) past exposure history predicts the risk factor. Covariates satisfying condition 1 are called time-dependent confounders. Past CD4 count is a time-dependent confounder for the effect of zidovudine on survival, because it is a risk factor for mortality and a predictor of subsequent initiation of zidovudine therapy,<sup>6</sup> and past zidovudine history is an independent predictor of subsequent CD4 count.<sup>8</sup> In fact, all standard methods (for example, Cox or Poisson regression) that predict the mortality rate at each time using a summary of zidovudine history up to that time may produce biased estimates of the causal effect of zidovudine whether or not one adjusts for past CD4 count in the analysis.

#### Marginal Structural Cox Proportional Hazards Model

In the absence of time-dependent confounding, a time-dependent Cox proportional hazards model is typically used. We treat visit 5, or the earliest subsequent visit at which a man was HIV positive, as start of follow-up time for our analysis. We define  $T$  to be a subject's time of death with time measured in months since start of follow-up, and  $A(t)$  to be 1 if a subject was on zidovudine at time  $t$ . We use overbars to represent a covariate history so, for example,  $\bar{A}(t) = \{A(u); 0 \leq u < t\}$  is a subject's treatment history up to  $t$ . Finally, let  $V$  be a vector of time-independent, baseline covariates measured before start of follow-up. Then the conditional hazard of death (that is, mortality rate)  $\lambda_T(t|\bar{A}(t), V)$  given treatment history  $\bar{A}(t)$  and baseline covariates  $V$  is modeled as

$$\lambda_T(t|\bar{A}(t), V) = \lambda_0(t) \exp(\gamma_1 A(t) + \gamma_2 V).$$

The subscript  $T$  in  $\lambda_T(t|\bar{A}(t), V)$  merely identifies this hazard function as being that corresponding to the variable  $T$ . In our analysis, the covariates in  $V$  are age, calendar year, CD4 count, CD8 count, white blood cell count (WBC), red blood cell count (RBC), platelets, and presence of symptoms. Symptomatic status was defined by previous presence of one or more of the following clinical symptoms or signs: fever (temperature  $> 37.9^\circ\text{C}$ ) for  $\geq 2$  weeks, oral candidiasis, diarrhea for  $\geq 2$  weeks, weight loss of  $\geq 4.5$  kg, oral hairy leukoplakia,

or herpes zoster. We assume, for simplicity, that patients remain on therapy once they start it and that the hazard of death at time  $t$  depends on a subject's zidovudine history only through its current value, but alternative specifications are possible. Suppose, for the moment, no censoring occurs, that is, death times  $T$  are observed for all subjects.

In the presence of time-dependent covariates  $L(t)$  satisfying the conditions 1 and 2, the estimate  $\hat{\gamma}_1$  obtained by maximizing the Cox partial likelihood is an (asymptotically) unbiased estimate of the association parameter  $\gamma_1$ . However, it is a biased estimate of the causal effect of zidovudine on mortality, even if we had included the time-dependent covariates  $L(t)$  as regressors in the model.

Arguing as in our companion paper,<sup>4</sup> we can eliminate or reduce this bias by fitting the above time-dependent Cox model with the contribution of a subject  $i$  to a risk-set calculation performed at time  $t$  weighted by the "stabilized" weights

$$sw_i(t) = \frac{\prod_{k=0}^{\text{int}(t)} \text{pr}(A(k) = a_i(k) | \bar{A}(k-1) = \bar{a}_i(k-1), V = v_i)}{\prod_{k=0}^{\text{int}(t)} \text{pr}(A(k) = a_i(k) | \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) = \bar{L}_i(k))},$$

to obtain an IPTW partial likelihood estimate. In the above  $\bar{A}(-1)$  is defined to be 0. Here,  $\text{int}(t)$  is the largest integer less than or equal to  $t$  and  $k$  is an integer-valued variable denoting whole months since start of follow-up. Because a subject's recorded treatment changes at most one per month, each factor in the denominator of  $sw_i(t)$  is, informally, the probability that the subject received his own observed treatment at month  $k$ , given his past treatment and prognostic factor history [ $V$  is included in  $L(0)$ ]. Each factor in the numerator is, informally, the probability that the subject received his observed treatment conditional on his past treatment history and baseline covariates, but not further adjusting for his past time-dependent prognostic factor history. "Nonstabilized" weights  $w_i(t)$ , in which the numerator of  $sw_i(t)$  is replaced by 1, can be used in lieu of  $sw_i(t)$ . Although this choice will not influence the consistency of our causal estimates, the stabilized weights  $sw_i(t)$  are preferred because they generally yield 95% confidence intervals that not only are narrower (that is, more efficient) but also have actual coverages rates that are closer to 95%. In a latter section, we describe how these stabilized inverse-probability-of-treatment weights  $sw_i(t)$  can be estimated from the data.

Suppose all relevant time-dependent confounders are measured and included in  $L(t)$ . Then, weighting by  $sw_i(t)$  effectively creates, for a risk set at time  $t$ , a pseudopopulation in which (1)  $\bar{L}(t)$  no longer predicts initiation of zidovudine at  $t$  (that is,  $\bar{L}(t)$  is not a confounder), and (2) the causal association between zidovudine and mortality is the same as in the original study population.<sup>4</sup> As argued in Ref 4, this implies that an IPTW estimator, say  $\hat{\beta}_1$ , of the parameter  $\gamma_1$  of our time-

dependent Cox model will converge to a quantity  $\beta_1$  that can be appropriately interpreted as the causal effect, on the log rate ratio scale, of zidovudine on mortality.

To formalize the above, we introduce counterfactual outcomes.<sup>4</sup> For each possible treatment history  $\bar{a} = \{a(t); 0 \leq t < \infty\}$ , let  $T_{\bar{a}}$  be the random variable representing the subject's time to death had he followed, possibly contrary to fact, the zidovudine history  $\bar{a}$  from the start of follow-up, rather than his observed history. For example,  $T_{\bar{a}}$  with  $\bar{a}$  such that  $a(t) = 0$  for  $t < 2.5$  and  $a(t) = 1$  for  $t \geq 2.5$  is the subject's survival time when he started zidovudine therapy 2.5 months after the start of follow-up. We only observe  $T_{\bar{a}}$  for those treatment histories  $\bar{a}$  that agree with the subject's observed treatment history until the subject's observed death time  $T$ . For these histories  $T_{\bar{a}}$  equals  $T$ . For each  $\bar{a}$ , we specify the marginal structural Cox proportional hazards model

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

where  $\lambda_{T_{\bar{a}}}(t|V)$  is the hazard of death at  $t$  among subjects with baseline covariates  $V$  in the source population had, contrary to fact, all subjects followed zidovudine history  $\bar{a}$  through time  $t$ , the scalar  $\beta_1$  and the row vector  $\beta_2$  are unknown parameters, and  $\lambda_0(t)$  is an unspecified baseline hazard. We refer to this model as an MSM because, within levels of  $V$ , it is a structural (that is, causal) model for the marginal distribution of the counterfactual variables  $T_{\bar{a}}$ .

The parameter  $\beta_1$  of our MSM is the causal log rate ratio for zidovudine. Hence,  $\exp(\beta_1)$  has a causal interpretation as the ratio of the mortality (hazard) rate at any time  $t$  had all subjects been continuously exposed to zidovudine compared with the hazard rate at time  $t$  had all subjects remained unexposed.  $\beta_1$  is consistently estimated by our IPTW estimator  $\hat{\beta}_1$ , under the untestable assumption of no unmeasured confounders given the measured risk factors in  $L(t)$ .<sup>1</sup> We shall make this assumption with  $L(t)$  being the covariate vector with the following elements: the most recent recorded CD4, CDS, WBC, RBC, platelets, presence of an AIDS-defining illness, and symptomatic status before  $t$ .

It is difficult to get standard Cox model software to compute our IPTW estimator  $\hat{\beta}_1$  because our subject-specific weights  $sw_i(t)$  vary over time, and most standard Cox model software programs, even those that allow for subject-specific weights, do not allow for subject-specific time-varying weights. The approach we shall adopt to overcome this software problem is to fit a weighted pooled logistic regression treating each person-month as an observation. (In the MACS, our 2,178 men contribute 143,194 person-months of observation.) That is, we will fit, by weighted logistic regression using weights  $sw_i(t)$ , the model

$$\text{logit } pr[D(t) = 1 | D(t-1) = 0, \bar{A}(t-1), V] = \beta_2(t) + \beta_1 A(t-1) + \beta_2 V$$

where, henceforth  $t$ , like  $k$ , is integer valued denoting whole months since start of follow-up,  $D(t) = 0$  if a

subject was alive in month  $t$  and 1 if the subject died in month  $t$ , and  $\beta_2(t)$  is a time-specific (that is, month-specific) intercept. This method has the advantage of being easily programmed in many standard statistical packages. In the unweighted case, it is essentially equivalent to fitting an unweighted time-dependent Cox model, because the hazard in any single month is small.<sup>9</sup> However, the use of weights induces within-subject correlation, which invalidates the standard error estimates outputted by a standard logistic program (they can be either too large or too small). To overcome this difficulty, the above weighted logistic model should be fit using a generalized estimating equations<sup>10</sup> program (for example, option "repeated" in SAS Proc Genmod) that outputs "robust" variance estimators that allow for correlated observations. The robust variance estimator provides a conservative confidence interval for the  $\beta$ . That is, under our assumptions, the 95% confidence interval calculated as  $\hat{\beta} \pm 1.96 \times \text{robust standard error}$  is guaranteed to cover the true  $\beta$  at least 95% of the time in large samples.

### Censoring

The analysis just described assumes that there is no dropout or censoring by end of follow-up. We define the censoring indicator  $C(t)$  to be 1 if a subject is right-censored by time  $t$  and  $C(t) = 0$  otherwise, where a subject is right-censored if he either dropped out of the study or reached the administrative end of follow-up alive. To estimate  $\beta_1$  in the presence of censoring, we fit a weighted Cox model in which, for a subject at risk at month  $t$ , we use the weight  $sw_i(t) \times sw_i^*(t)$ , where

$$sw_i^*(t) =$$

$$\prod_{k=0}^t \frac{pr[C(k) = 0 | \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1), V = v_i]}{pr[C(k) = 0 | \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k-1) = \bar{l}_i(k-1)]}$$

where  $\bar{C}(k-1)$  and  $\bar{A}(k-1)$  are defined to be 0.  $sw_i^*(t)$  is, informally, the ratio of a subject's probability of remaining uncensored up to month  $t$ , calculated as if there had been no time-dependent determinants of censoring except past zidovudine history, divided by the subject's conditional probability of remaining uncensored up to month  $t$ . The denominator of the product  $sw_i(t) \times sw_i^*(t)$  is, informally, the probability that a subject had had his observed zidovudine and censoring history through month  $t$ . Because  $sw_i(t)$  and  $sw_i^*(t)$  are unknown, they must be estimated from the data as described below. Weighting by  $sw_i(t) \times sw_i^*(t)$  produces a consistent estimate of the causal parameter  $\beta_1$  under the assumption that the measured covariates are sufficient to adjust for both confounding and selection bias due to loss to follow-up.<sup>4</sup>

### Estimation of the Weights

The practical problem faced by the investigator is how to obtain the quantities  $sw_i(t) \times sw_i^*(t)$  necessary to run

the pooled weighted logistic regression model. Consider first the estimation of  $sw_i(t)$ . We need to estimate consistently the denominator and numerator of  $sw_i(t)$  for each subject and time point. Because any subject starting zidovudine was assumed to remain on it thereafter, we can regard the time to starting zidovudine as a failure time variable and model the probability of starting zidovudine through a pooled logistic model that treats each person-month as an observation and allows for a time-dependent intercept. Specifically, we can, for example, fit the model  $\logit pr[A(k) = 0 | \bar{A}(k-1) = 0, L(k)] = \alpha_0(k) + \alpha_1 L(k) + \alpha_2 V$  and obtain estimates  $\hat{\alpha} = (\hat{\alpha}_0(k), \hat{\alpha}_1, \hat{\alpha}_2)$  for the unknown parameters. It is only necessary to fit the model for subjects alive and uncensored in month  $k$  who had yet to begin zidovudine (that is, the 85,115 person-months in the MACS with  $\bar{A}(k-1) = 0$ ).

The estimated predicted values  $\hat{p}_i(k) = \text{expit}(\hat{\alpha}_0(k) + \hat{\alpha}_1 L_i(k) + \hat{\alpha}_2 V_i)$  from this model are the estimated probabilities of subject  $i$  not starting zidovudine in month  $k$  given that zidovudine had not been started by month  $k-1$ , where  $\text{expit}(x) = e^x/(1 + e^x)$ . Our estimate of the denominator of  $sw_i(k)$  for person

$i$  in month  $k$  is the product  $\bar{p}_i(k) = \prod_{u=0}^k \hat{p}_i(u)$  if subject  $i$  did not start zidovudine up to month  $k$  and is

$\bar{p}_i(k) = [1 - \hat{p}_i(t)] \prod_{u=0}^{t-1} \hat{p}_i(u)$  if subject  $i$  started zidovudine at month  $t$  for  $t \leq k$ . [Note that, in calculating  $\bar{p}_i(k)$ , we have used our assumption that no subject stops zidovudine once begun.] Similarly, we estimate the numerator of  $sw_i(k)$  by fitting the above logistic model except with the covariates  $L(k)$  removed from the model.

There is a small but important technical detail we have yet to discuss. For our IPTW estimates of  $\beta$  to be

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   |
| Stabilized weights       |  | 0.74 | 0.57-0.96                   |
| Nonstabilized weights    |  | 0.75 | 0.54-1.05                   |
|                          |  | RR   | Invalid Model-Based‡ 95% CI |
| Stabilized weights       |  | 0.74 | 0.62-0.87                   |
| Nonstabilized weights    |  | 0.75 | 0.71-0.80                   |

RR = mortality rate ratio (in 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  $\geq 500/\mu\text{l}$ ), CD8+ ( $<500$ , 500-999, or  $\geq 1,000$  per  $\mu\text{l}$ ), WBC ( $<3,000$ , 3,000-4,999, or  $\geq 5,000$  per  $\mu\text{l}$ ), RBC ( $<35$ , 35-44, or  $\geq 45 \times 10^3$  per  $10^3$ ), platelets ( $<150$ , 150-249, or  $\geq 250 \times 10^3$  per  $\mu\text{l}$ ), presence of symptoms (fever, oral candidiasis, diarrhea, weight loss, oral hairy leukoplakia, or neither/never, or not if otherwise).

† Weights calculated as described in the text using data on baseline covariates plus most recent CD4, CD8, WBC, RBC, or platelet count, presence of symptoms, and previous zidovudine use.

‡ The model-based intervals are shown for weighted models because they fail to account for the within-subject differences induced by weighting.

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% CI |
|---|--------------------|-----------------------|---------------------|
| Zidovudine  | -0.301             | 0.132                 | -0.559, -0.043      |
| Age at baseline                                       | 0.038              | 0.008                 | 0.224, 0.053        |
| Baseline CD4+, per $\mu\text{l}$                      |                    |                       |                     |
| <200  | 1.913              | 0.136                 | 1.645, 2.180        |
| 200-499   | 0.634              | 0.088                 | 0.461, 0.806        |
| $\geq 500$  | 0.000              |                       |                     |
| Baseline CD8+, per $\mu\text{l}$                      |                    |                       |                     |
| <500  | -0.648             | 0.122                 | -0.885, -0.409      |
| 500-99  | -0.477             | 0.087                 | -0.645, -0.306      |
| $\geq 1000$   | 0.000              |                       |                     |
| Baseline WBC, per $\mu\text{l}$                       |                    |                       |                     |
| <3000   | 0.790              | 0.085                 | 0.575, 1.205        |
| 3000-4999   | 0.273              | 0.420                 | 0.106, 0.440        |
| $\geq 5000$   | 0.000              |                       |                     |
| Baseline RBC, $\times 10^3$ , per $\mu\text{l}$       |                    |                       |                     |
| <35   | 0.231              | 0.420                 | -0.593, 1.055       |
| 35-45   | 0.254              | 0.105                 | 0.045, 0.460        |
| $\geq 45$   | 0.000              |                       |                     |
| Baseline platelets, $\times 10^3$ , per $\mu\text{l}$ |                    |                       |                     |
| <150  | 0.616              | 0.129                 | 0.364, 0.869        |
| 150-249   | 0.189              | 0.087                 | 0.019, 0.359        |
| $\geq 250$  | 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                 | -0.426, 1.287       |
| 1987-89   | 0.250              | 0.442                 | -0.617, 1.120       |
| 1990-93   | 0.000              |                       |                     |

CI = confidence limits.

\* Weighted logistic model including the covariates listed in the table plus a time-varying intercept (not shown). Weights were estimated by  $sw_i(t) \times sw_i'(t)$  as defined in the text.

consistent, it is necessary that the denominator of  $sw_i(t)$  be consistently estimated. To do so, we cannot estimate a separate intercept  $\alpha_0(k)$  for each month  $k$ . Rather, we need to "borrow strength" from subjects starting zidovudine in months other than  $k$  to estimate  $\alpha_0(k)$ . This can be accomplished by assuming that  $\alpha_0(k)$  is constant in windows of, say, 3 months. An alternative approach is to assume the  $\alpha_0(k)$  are a smooth function of  $k$  and thus can be estimated by smoothing techniques (such as regression splines, smoothing splines, or kernel regression).<sup>11</sup>

To correct for censoring, we estimate  $sw_i'(k)$  in a manner analogous to the estimation of  $sw_i(k)$  except with  $A(k)$  replaced by  $C(k)$  as the outcome variable, with  $A(k-1)$  added as an additional regressor, and not conditioning on  $\bar{A}(k-1) = 0$  but rather on  $\bar{C}(k-1) = 0$ .

### Causal Effect of Zidovudine in the Multicenter AIDS Cohort Study

Using a standard Cox proportional hazards model—or the equivalent pooled logistic regression model—with no covariates, the crude mortality rate ratio for zidovudine was 3.6 (95% confidence interval 3.0-4.3). The

TABLE 3. Estimated Probability of Having One's Own Observed Treatment History [Estimated Denominator of  $sw_i(t)$ ] and Censoring History [Estimated Denominator of  $sw_i^*(t)$ ] 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 (1955, 1956, 1957-1989, or 1990-1993), CD4 (<200, 200-499, or  $\geq 500$  per  $\mu\text{l}$ ), CD8 (<500, 500-999,  $\geq 1,000$  per  $\mu\text{l}$ ), WBC (<3,000, 3,000-4,999, or  $\geq 5,000$  per  $\mu\text{l}$ ), RBC (<35, 35-44, or  $\geq 45 \times 10^3$  per  $\mu\text{l}$ ), platelets (<150, 150-249, or  $\geq 250 \times 10^3$  per  $\mu\text{l}$ ), 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,  $\geq 45 \times 10^3$  per  $\mu\text{l}$ ), platelets, presence of symptoms, or presence of an AIDS-defining illness.

addition of the baseline covariates  $V$  to the model decreased this rate ratio to 2.3 (1.9-2.8).

To adjust further for confounding due to the time-dependent factors  $L(t)$ , we estimated the parameters of our marginal structural Cox model by calculating a stabilized weight  $sw_i(t) \times sw_i^*(t)$  for each person-month and fitting a weighted pooled logistic model. The estimated causal mortality rate ratio  $\exp(\beta_1)$  was 0.7 (95% conservative confidence interval-0.6-1.0), indicating that, under our assumptions, zidovudine therapy appears to decrease the risk of death. When nonstabilized weights  $\hat{w}_i(t) \times \hat{w}_i^*(t)$  were used, the rate ratio was virtually identical but the 95% conservative confidence interval was 30% wider, compared with the stabilized results (Table 1). We also report the invalid model-based intervals obtained using an ordinary weighted logistic regression program that does not account for within-subject correlations. The point estimates and 95% conservative confidence intervals for each of the parameters of our marginal structural Cox model are displayed in Table 2.

The stabilized weights were calculated by means of four pooled logistic regression models, as described in the previous section. In two of the models the outcome was "initiation of zidovudine." Using the estimated predicted values from each of these models, we calculated two quantities for each observation: the probability of each person having his own observed zidovudine history up to month  $t$  given baseline covariates  $V$ , and, then, given also time-varying covariates  $L(t)$ . Similar models were fit for the outcome "censoring," after adding zidovudine history as a time-varying dichotomous variable indicating whether the subject had started zidovudine by month  $t - 1$ .

Table 3 shows the center and dispersion parameters of the distribution of the four estimated probabilities at two arbitrary time points: 24 and 84 months of follow-up.

The estimated probabilities of having one's own observed zidovudine history at 24 months of follow-up, given time-varying covariates, range from 0.939 to 0.002. This would be translated into (nonstabilized) inverse-probability-of-treatment weights  $\hat{w}_i(t)$  ranging from 1.06 (1/0.939) to 500 (1/0.002). Thus, in the pseudopopulation, some observations would be represented by 1.06 copies of themselves, whereas others would be represented by 500 copies. The use of stabilized inverse-probability-of-treatment weights  $sw_i(t)$  "normalizes" or stabilizes the range of these inverse probabilities and increases the efficiency of the analysis by preventing just a few people from contributing most of the observations in the pseudopopulation. Thus, the values  $sw_i(t)$  for  $t = 24$  are centered around 1.01 and show a narrower range (0.14-6.67).

The estimated probabilities of being uncensored at 24 months follow a more peaked distribution, tightly centered around values close to 1 (0.96). This is expected, as 95% of the men were uncensored at 24 months of follow-up. Inverse probabilities  $\hat{w}_i^*(t)$  range from 1.00 (1/0.997) to 5.03 (1/0.199). The stabilized weights  $sw_i^*(t)$ , for  $t = 24$ , are centered around 1 and range from 0.93 to 1.23. The estimated probabilities of being uncensored at 84 months are lower, as expected.

The distribution of the final weights, which combine information on zidovudine and censoring history, is presented in Figures 1 and 2 for several follow-up times (a logarithmic transformation was applied for display purposes only). Two sets of weights were estimated: the stabilized weights  $sw_i(t) \times sw_i^*(t)$  and the nonstabilized weights  $\hat{w}_i(t) \times \hat{w}_i^*(t)$ . The distribution of stabilized weights is symmetric and centered around 1 at all times, whereas its variance increases over time. The distribution of the nonstabilized weights is skewed, and its variance greatly exceeds that of the stabilized weights.

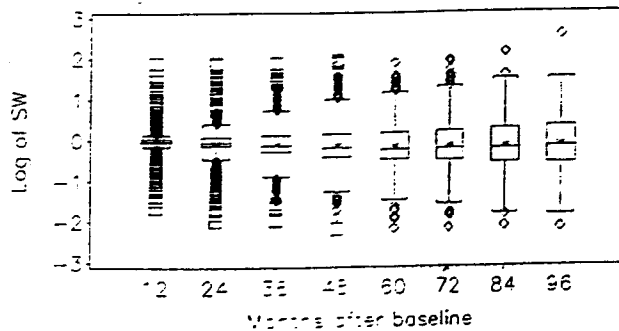


FIGURE 1. Distribution of stabilized weights SW. The box for each group shows the location of the mean (\*), median (middle horizontal bar) and quartiles (border horizontal bars). Vertical lines extend to the most extreme observations which are no more than  $1.5 \times \text{IQR}$  beyond the quartiles.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

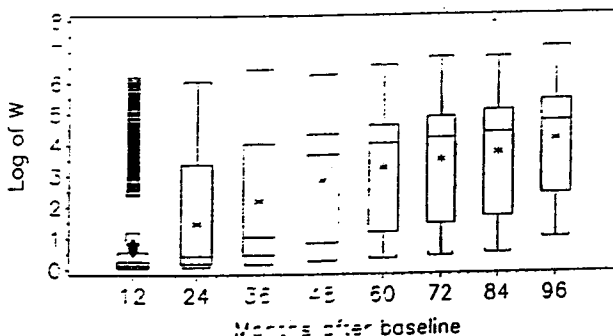


FIGURE 2. Distribution of nonstabilized weights SW. The box for each group shows the location of the mean (\*), median (middle horizontal bar) and quartiles (border horizontal bars). Vertical lines extend to the most extreme observations which are no more than  $1.5 \times \text{IQR}$  beyond the quartiles.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

The weight estimates were robust with respect to the method used to estimate the time-dependent baseline logit  $\alpha_3(k)$  in the logistic models for zidovudine and censoring, provided that sufficient flexibility was allowed. The weights in Figures 1 and 2 were obtained by modeling the time-dependent intercept  $\alpha_3(k)$  with natural cubic splines with five knots (on months 23, 44, 71, 94, and 100, which correspond to the percentiles 5, 27.5, 50, 72.5, and 95, respectively).<sup>11</sup>

#### Adjusting for Time-Dependent Confounders in a Cox Model

We also fit a standard (unweighted) time-dependent Cox model in which we included, at each month  $t$ , the current value  $L(t)$  of the time-dependent covariates, the baseline covariates  $V$ , and the treatment  $A(t)$ . We obtained a point estimate of 0.4 (95% confidence interval: 0.3–0.5) for the zidovudine coefficient, which was considerably less than our stabilized IPTW estimate of 0.7. Nevertheless, as discussed in section 7.1 of our compar-

ion paper,<sup>4</sup> because the covariates in  $L(t)$  are affected by earlier treatment, the zidovudine coefficient cannot be causally interpreted as either the overall zidovudine effect or the direct effect of zidovudine mediated by pathways not through the covariates  $L(t)$ . In contrast, under our assumption of no unmeasured confounders, the coefficient of zidovudine in our marginal structural Cox model represents the overall effect of zidovudine.

More specifically, if we had included in the above covariate-adjusted time-dependent Cox model both a term for current zidovudine exposure [that is,  $A(t)$ ] and several terms for past zidovudine exposure [for example, cumulative months on treatment before  $t$  [ $\text{cum}(t)$ ], and the indicator  $A(t-6)$  of whether a subject was on treatment 6 months previously], then, in the absence of unmeasured confounders and model misspecification, the coefficient of  $A(t)$  would have a causal interpretation but the coefficients  $\text{cum}(t)$  and  $A(t-6)$  would not, because only current zidovudine does not affect  $L(t)$ . The coefficient of  $A(t)$  would represent the effect on a log rate ratio scale of recent zidovudine on mortality in month  $t$  within strata defined by zidovudine and covariate history up to  $t$  and would generally differ from the coefficient  $\beta_1$  of  $A(t)$  in our MSM, as the coefficients in the two models represent different causal contrasts. [Indeed, it can be shown that if our MSM is correct and  $\beta_1$  is nonzero, the causal rate ratio for current zidovudine in the covariate and past treatment-adjusted time-dependent Cox model will not be constant over strata defined by past treatment and covariate history. Hence, we would need to include interaction terms between  $A(t)$  and the variables  $L(t)$ ,  $A(t-6)$ , and  $\text{cum}(t)$  in our covariate and past treatment-adjusted time-dependent Cox model to avoid model misspecification.] However, as mentioned above, the coefficient (estimated to be 0.4) of  $A(t)$  in the covariate-adjusted time-dependent Cox model that does not include terms for past zidovudine exposure does not have a causal interpretation, because past treatment is a confounder for current treatment and thus must be adjusted for. This is true even under the null hypothesis of no direct, indirect, or overall effect of zidovudine on mortality whenever, as will be essentially always the case, a component of  $L(t)$ , say RBC, and mortality in month  $t$  have an unmeasured common cause (for example, the baseline number of bone marrow stem cells); adjusting for a variable  $L(t)$  affected by past zidovudine makes past zidovudine a noncausal independent (protective) risk factor for mortality within strata of  $L(t)$  and  $A(t)$ , and thus, to estimate the effect of recent zidovudine exposure, past zidovudine must be controlled as a confounder in the analysis.

#### Comparison of Marginal Structural Models with Previously Proposed Methods

Before introducing MSMs, Robins and co-workers introduced three methods for estimation of the causal effect of a time-varying treatment in the presence of time-varying confounders: the parametric g-computation algorithm formula estimator,<sup>12,13</sup> g-estimation of structural

nested models,<sup>12,14,15</sup> and the iterative conditional expectations (ICE) estimator.<sup>3,12</sup> IPTW estimation of MSMs constitutes a fourth method. 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. They differ when, owing to sparse multivariate data, one must introduce modeling assumptions.

ICE estimators can only rarely be used, because they often lead to logically incompatible models and will not be discussed further.<sup>12</sup> Of the remaining three methods, inference based on SNMs and MSMs is preferable to that based on the parametric *g*-computation algorithm. The reason is that MSM and SNM models, in contrast to models based directly on the conditional probabilities in the *g*-computation algorithm formula, include parameters that represent the null hypothesis of no treatment effect.<sup>12,15</sup> As a consequence, when using the parametric *g*-computation algorithm estimator, it is quite difficult to determine whether one's confidence interval for the treatment effect includes the null hypothesis of no effect.

MSMs have two major advantages over SNMs. Although useful for survival time outcomes, continuous measured outcomes (for example, blood pressure), and Poisson count outcomes, logistic SNMs cannot be conveniently used to estimate the effect of treatment on dichotomous (0, 1) outcomes unless the outcome is rare.<sup>1,2,12</sup> This is because logistic SNMs cannot be fit by *g*-estimation. In contrast, as we have seen,<sup>4</sup> IPTW estimation of logistic MSMs can be used to estimate the effect of a time-dependent treatment on a binary outcome.

The second major advantage of MSMs is that they resemble standard models, whereas SNMs do not. For example, the logistic MSM described in our companion paper<sup>4</sup> and the Cox proportional hazards MSM described here are the natural way to extend the ordinary logistic and time-dependent Cox models to allow for estimation of causal effects. The close resemblance of MSMs to standard statistical models makes their application more intuitive for researchers and easier for programmers.

Nevertheless, SNMs have a number of advantages over MSMs. For example, as discussed in Ref 4, MSMs should not be used to estimate effects in studies (such as occupational cohort or cancer screening 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 exposure *a*(*k*).<sup>1,2</sup>

A second major drawback of MSMs is that one must be able to specify a correct model for the conditional probability of exposure,

$$\text{pr}(A(k) = a(k) | \bar{L}(k) = \bar{l}(k), \bar{A}(k-1) = \bar{a}(k-1)),$$

for each time *k* up to end of follow-up. This is unfortunate because, if the *L*(*k*) and *A*(*k*) are discrete, we could use nonparametric saturated models for small values of *k*,

(say *k* = 0, 1, 2), but for large *k* we require strong modeling assumptions, as there are many variables in  $\bar{l}(k) = (l(0), l(1), \dots, l(k))$  and in  $\bar{a}(k-1)$ . It is unlikely that these modeling assumptions would be precisely correct. Furthermore, even when these modeling assumptions are correct, if the distribution of the stabilized weights is highly variable and skewed as a result of very strong covariate-treatment associations, 95% confidence intervals based on IPTW estimation of an MSM will be very wide and may fail to cover the true parameter at least 95% of the time, unless the sample size is very large.

The use of *g*-estimation of SNMs overcomes the above difficulties. For example, one can use SNMs to estimate the effect of an exposure on mortality in occupational cohort studies.<sup>14,16</sup> Similarly, one can unbiasedly estimate the causal parameter of a SNM without having to model the probability of treatment given the past through end of follow-up. Instead, in the setting of a discrete *A*(*k*) and *L*(*k*) described above, one can unbiasedly estimate the parameters of SNMs by using a saturated model for the probability of exposure *A*(*k*) given the past for *k* = 0, 1, 2 periods and ignoring exposure at later periods, thus preventing bias due to misspecification of the model for exposure. Of course, as always in statistical analysis, there will be a loss of efficiency of estimation associated with this protection against bias. Finally, in the presence of strong covariate-treatment associations, theoretical arguments imply that it should be possible to construct confidence intervals based on *g*-estimation of SNMs that are both narrower and have better coverage properties than those based on IPTW estimation of MSMs. However, further research is required to see whether this theoretical prediction is borne out in practice.

Another advantage of SNMs over MSMs is that, although MSMs are useful for estimating the causal effect of the prespecified treatment regime  $\bar{a}$  (for example, always treat, treat on alternate months, etc), they are much less useful than SNMs for modeling the interaction of treatment with a time-varying covariate and for estimating the effect of dynamic treatment plans in which treatment on a given month is decided in part on the basis of a subject's evolving covariate history.<sup>1,2</sup> It is important to recognize that actual medical treatment regimes are usually dynamic, because if a patient develops a toxic reaction to a drug, the drug must be stopped. Nonetheless, causal questions concerning prespecified treatment plans, such as estimating the effect of a continuous exposure at a certain level vs no exposure, are of great interest in many areas of epidemiology, including nutritional and environmental.

## Discussion

We have used a marginal structural Cox proportional hazards model to estimate the causal effect of zidovudine on mortality of HIV-positive patients in the MACS. This method was used because standard statistical methods are not appropriate when there exists time-depen-

dent confounding by variables, such as CD4 count, that are affected by previous exposure.

Because of the presence of confounding, the crude mortality rate ratio for zidovudine was 3.6 (95% confidence interval-3.0-4.3), erroneously suggesting that zidovudine increased risk of death. The rate ratio estimated by the (unweighted) standard model that included only baseline covariates, and that therefore does not adjust for time-dependent confounding, was 2.3 (95% confidence interval-1.9-2.8), which still suggests a detrimental effect of zidovudine.

In fact, the mortality rate ratio for zidovudine was 0.7 (95% conservative confidence interval-0.6-1.0) in the weighted analysis that provides, under our assumptions, an unbiased estimate of the causal rate ratio,  $\exp(\beta_1)$ , of the marginal structural Cox model, because the weighted analysis appropriately adjusts for time-dependent confounders affected by earlier treatment.

The difference between the unweighted and weighted estimates is an indication of the amount of confounding due to the time-dependent prognostic factors. The weights can be interpreted as the number of copies of each observation that are necessary to form a pseudopopulation in which censoring does not exist and in which the time-dependent prognostic factors do not predict initiation of zidovudine history (that is, treatment is unconfounded).

Like all causal inferences, the validity of our analyses depends on a number of assumptions. First, we assume that the information on month of zidovudine initiation and month of death is accurate. Second, we assume that the measured covariates in  $L(t)$  are sufficient to adjust for both confounding and selection bias due to loss to follow-up. This assumption implies that we have available, for each month  $t$ , accurate data recorded in  $L(t)$  on all time-dependent covariates that (1) are independent predictors of death and (2) independently predict the probability of starting zidovudine and/or of being censored in that month. Unfortunately, as in all observational studies, these two assumptions cannot be tested from the data. In our analysis, we assume that this goal has been realized, while recognizing that, in practice, this assumption would never be precisely or sometimes even approximately true. Recently, Robins et al.<sup>16</sup> have developed extensions of IPTW estimation of MSMs that allow one to evaluate the sensitivity of one's estimates to increasing violation of these fundamental assumptions.

Third, we assume that the models for initiation of zidovudine and censoring, given the past, are correctly specified. Fourth, we assume that our MSM for the effect of zidovudine on mortality, within levels of baseline covariates  $V$ , is correctly specified.

Although the stated assumptions may seem heroic, note that in point-exposure studies the same assumptions (accurate information, no unmeasured confounders, noninformative censoring, and no misspecification of the model) are required to give a causal interpretation to the parameters of standard statistical models. Furthermore, when studying the effect of a time-dependent treatment such as zidovudine, the assumptions of MSMs

are less restrictive than those of standard methods; MSMs do not require the absence of time-dependent confounding by variables affected by previous exposure.

## Acknowledgments

We thank Sander Greenland and several referees for helpful comments.

Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS) with centers (Principal Investigators) at The Johns Hopkins School of Public Health (Joseph Margolick, Alvaro Muñoz); Howard Brown Health Center and Northwestern University Medical School (John Phair); University of California, Los Angeles (Roger Detels, Janis V. Giorgi, Beth Jamieson); and University of Pittsburgh (Charles Rinaldo). The MACS is funded by the National Institute of Allergy and Infectious Diseases, with additional supplemental funding from the National Cancer Institute: UO1-AI-35042, 5-M01-RR-0052 (GCRC), UO1-AI-35043, UO1-AI-37984, UO1-AI-35039, UO1-AI-35040, UO1-AI-37613, UO1-AI-35041.

## References

1. Robins JM. Marginal Structural Models. 1997 Proceedings of the Section on Bayesian Statistical Science. Alexandria, Virginia: American Statistical Association 1998;1-10.
2. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran E, Berry D, eds. *Statistical Models in Epidemiology: The Environment and Clinical Trials*. New York: Springer-Verlag, 1999;95-134.
3. Robins JM. Structural nested failure time models. In: Andersen PK, Keiding N, section eds. *Survival Analysis*. In: Armitage P, Colton C, eds. *The Encyclopedia of Biostatistics*. Chichester, UK: John Wiley and Sons, 1998; 4372-4389.
4. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550-560.
5. Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. *Am J Epidemiol* 1987;126:310-315.
6. Graham NM, Zeger SL, Park LP, Vermund SH, Derels R, Rinaldo CR, Phair JP. The effects on survival of early treatment of human immunodeficiency virus infection. *N Engl J Med* 1992;326:1037-1042.
7. Robins JM. Causal inference from complex longitudinal data. In: Berkane M, ed. *Latent Variable Modeling and Applications to Causality: Lecture Notes in Statistics* 120. New York: Springer-Verlag, 1997;69-117.
8. Kinloch-De-Loes S, Hirschel BJ, Hoen B, Cooper DA, Tindall B, Carr A, Saurnt JH, Clumeck N, Lazzarin A, Mathiesen L, et al. A controlled trial of zidovudine in primary human immunodeficiency virus infection. *N Engl J Med* 1995;333:409-413.
9. D'Agostino RB, Lee M-L, Belanger AJ. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med* 1990;9:1501-1515.
10. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13-22.
11. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. New York: Chapman and Hall, 1990;335.
12. Robins JM. Correction for non-compliance in equivalence trials. *Stat Med* 1998;17:369-382.
13. Robins JM. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *J Chron Dis* 1987;40(suppl. 2):139S-161S.
14. Robins JM, Blevins D, Ritter G, Wolfson M. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients (abstract in *Epidemiology* 1993;4:139). *Epidemiology* 1992; 3:319-336.
15. Robins JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest L, Freeman H, Mulley A, eds. *Health Service Research Methodology: A Focus on AIDS*. Rockville, MD: National Center for Health Services Research, U.S. Public Health Service, 1989;113-159.
16. Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran E, Berry D, eds. *Statistical Methods in Epidemiology: The Environment and Clinical Trials*. New York: Springer-Verlag, 1999;1-92.

## APPENDIX

SAS CODE FOR THE MARGINAL STRUCTURAL COX  
PROPORTIONAL HAZARDS MODEL

In this appendix, we provide SAS code to fit the Cox proportional hazards MSM described in the text. The original MACS data file contains one record per man, but here we use a transformed, or pooled, file (MAIN) with each person-month as a separate record. This file format is necessary to fit pooled logistic models. The code used to generate the pooled dataset from the original one is available from the first author upon request. The records in the file MAIN must be sorted by patient identification number (variable ID) and, within each ID level, by month of follow-up (MONTH).

The SAS code shown below is organized as follows. First, we use Proc Logistic to fit four pooled logistic models (two for the probability of remaining off zidovudine and two for the probability of remaining uncensored) and obtain their predicted values. Second, we use a SAS data step to calculate the weights for each person-month from the predicted values of the previous four models. Last, we use Proc Genmod to fit the final weighted pooled logistic model that estimates the causal parameter of interest and its robust standard error.

The outcome variable in models 1 and 2 is a dichotomous variable A indicating whether the patient had started ( $A = 1$ ) or remained off ( $A = 0$ ) zidovudine on that month. When the option "descending" is not specified, Proc Logistic models the probability that the outcome variable is 0. Hence, models 1 and 2 model the probability of remaining off zidovudine. The "where" statement restricts the analysis to patients not previously on zidovudine by specifying that either month of follow-up (MONTH) is less than or equal to month of onset of zidovudine (ZDV\_M) or zidovudine was never initiated during the follow-up period (ZDV\_M is coded as missing, if this is the case). Model 1 includes as regressors a time-dependent intercept and the baseline covariates V: baseline age, calendar year, CD4, CD8, WBC, RBC, platelets, and presence of symptoms. Model 2 includes, in addition, the time-dependent covariates  $L(t)$ : most recently available CD4, CD8, WBC, RBC, platelets, symptoms, and AIDS-defining illness. We estimate the time-dependent intercept by a smooth function of the time since beginning of follow-up (MONTH) using natural cubic splines with five knots. To do so, we need to include, as regressors, the variables MONTH1, MONTH2, and MONTH3, that are specific polynomial functions of MONTH (calculated with the cubic splines SAS macro RCSPLINE in survrisk.pak, by Frank Harrel, which is publicly available on <http://jse.stat.ncsu.edu:7015/software/sas>).

The outcome variable in models 3 and 4 is a dichotomous variable C indicating whether the patient was censored ( $C = 1$ ) or uncensored ( $C = 0$ ) in that month. Thus, models 3 and 4 model the probability of remaining uncensored for each person-month. All available person-months are used. Model 3 includes the baseline covariates and the time-dependent intercept, whereas

model 4 includes the time-dependent covariates (to which we add A) as well.

For each model, we output a new data file (option "out =" in Proc Logistic) that contains, for each person-month, the original variables plus the predicted values from the model (option "p="). As an example, the first Proc Logistic creates the data set MODEL1 with its predicted values as the variable PZDV.0.

In the following data step, we merge the four output files in the file MAIN\_W that contains the predicted values from the four logistic models. We then compute the numerator K2.0 and the denominator K2\_W of the  $sw_i^*(t)$  from models 3 and 4.

Similarly, we calculate the numerator K1.0 and the denominator K1\_W of the weights  $sw_i(t)$  for months in which the subject has not yet started zidovudine from models 1 and 2. For a month in which a subject did begin zidovudine, we multiply by 1 minus the predicted value. For months subsequent to starting zidovudine, we no longer update K1.0 and K1\_W. Then we use the numerators and denominators to calculate the "stabilized" weights  $sw_i(t) \times sw_i^*(t)$  (STABW), and use the denominators alone to calculate the "nonstabilized" weights  $w_i(t) \times w_i^*(t)$  (NSTABW).

Finally, we call Proc Genmod to fit a weighted pooled logistic model for survival to obtain consistent estimates of the parameters of our Cox MSM. The outcome variable of this model, D, is a dichotomous variable indicating whether the patient died ( $D = 1$ ) or remained alive ( $D = 0$ ) in that month. The program will provide robust standard errors for the model parameters when the option "repeated" is included. The patient identification variable and the independent working correlation matrix ("subject = ID/type = ind") must be specified. We fit the model using the stabilized weights by specifying the variable STABW in the "scwgt" statement. Specifying the variable NSTABW fits the model with non-stabilized weights.

The SAS code given below can also be used to fit the logistic MSM of our companion paper.<sup>4</sup> The only difference is that the final weighted logistic model in Proc Genmod includes a single observation per person using as outcome variable the logistic variable Y of our companion paper, rather than the survival variable D considered in this paper.

```
/* MODEL 1 */
proc logistic data=MAIN;
  where MONTH <= ZDV_M or ZDV_M=.;
  model A=AGE.0 YEAR.01 YEAR.02 YEAR.03
        CD4.01 CD4.02 CD8.01 CD8.02
        WBC.01 WBC.02 RBC.01 RBC.02
        PLAT.01 PLAT.02 SYMPT.0
        MONTH MONTH1-MONTH3;
  output out=model1 p=pzdv.0;
run;

/* MODEL 2 */
proc logistic data=MAIN;
  where MONTH <= ZDV_m or ZDV_M=.;
```

```

model A=AGE_0 YEAR_01 YEAR_02 YEAR_03
      CD4_01 CD4_02 CD8_01 CD8_02
      WBC_01 WBC_02 RBC_01 RBC_02
      PLATE_01 PLATE_02 SYMPT_0
      CD4_1 CD4_2 CD8_1 CD8_2
      WBC_1 WBC_2 RBC_1 RBC_2
      PLAT_1 PLAT_2 SYMPT AIDS
      MONTH MONTH1-MONTH3;
output out=model2 p=pzdv_w;
run;

/* MODEL 3 */
proc logistic data=MAIN;
  model C=A AGE_0 YEAR_01 YEAR_02 YEAR_03
        CD4_01 CD4_02 CD8_01 CD8_02
        WBC_01 WBC_02 RBC_01 RBC_02
        PLATE_01 PLATE_02 SYMPT_0
        MONTH MONTH1-MONTH3;
output out=model3 p=punc_0;
run;

/* MODEL 4 */
proc logistic data=MAIN;
  model C=A AGE_0 YEAR_01 YEAR_02 YEAR_03
        CD4_01 CD4_02 CD8_01 CD8_02
        WBC_01 WBC_02 RBC_01 RBC_02
        PLATE_01 PLATE_02 SYMPT_0
        CD4_1 CD4_2 CD8_1 CD8_2
        WBC_1 WBC_2 RBC_1 RBC_2
        PLAT_1 PLAT_2 SYMPT AIDS
        MONTH MONTH1-MONTH3;
output out=model4 p=punc_w;
run;

data main_w;
  merge model1 model2 model3 model4;
  by ID MONTH;
/* variables ending with _C refer to the numerator of the
   weights
   variables ending with _w refer to the denominator of
   the weights */

/* reset the variables for a new patient */
if first.id then do;
  k1_0=1; k2_0=1; k1_w=1; k2_w=1;
end;
retain k1_0 k2_0 k1_w k2_w;
/* Inverse probability of censoring weights */
k2_0=k2_0*punc_0;
k2_w=k2_w*punc_w;
/* Inverse probability of treatment weights */
/* patients not on zidovudine */
if zdv_m>day or zdv_m =. then do;
  k1_0=k1_0*pzdv_0;
  k1_w=k1_w*pzdv_w;
end;
/* patients that start zidovudine this month */
else if zdv_m=day then do;
  k1_0=k1_0*(1-pzdv_0);
  k1_w=k1_w*(1-pzdv_w);
end;
/* patients that have already started zidovudine */
else do;
  k1_0=k1_0;
  k1_w=k1_w;
end;
/* Stabilized and non stabilized weights */
stabw=(k1_0*k2_0)/(k1_w*k2_w);
nstabw=1/(k1_w*k2_w);
run;

proc genmod data=main_w;
  class id;
  model D=A AGE_0 YEAR_01 YEAR_02 YEAR_03
        CD4_01 CD4_02 CD8_01 CD8_02
        WBC_01 WBC_02 RBC_01 RBC_02
        PLAT_01 PLAT_02 SYMPT_0
        MONTH MONTH1-MONTH3/
  link=logit dist=bin;
  scswgt stabw;
  repeated subject=ID/ type=ind;
run;

```