

# *The International Journal of Biostatistics*

---

*Volume 6, Issue 2*

2010

*Article 18*

CAUSAL INFERENCE

---

## When to Start Treatment? A Systematic Approach to the Comparison of Dynamic Regimes Using Observational Data

**Lauren E. Cain**, *Harvard School of Public Health*

**James M. Robins**, *Harvard School of Public Health*

**Emilie Lanoy**, *INSERM U943*

**Roger Logan**, *Harvard School of Public Health*

**Dominique Costagliola**, *INSERM U943 and Université*

*Pierre et Marie Curie*

**Miguel A. Hernán**, *Harvard School of Public Health and  
Harvard-MIT Division of Health Sciences and Technology*

### **Recommended Citation:**

Cain, Lauren E.; Robins, James M.; Lanoy, Emilie; Logan, Roger; Costagliola, Dominique; and Hernán, Miguel A. (2010) "When to Start Treatment? A Systematic Approach to the Comparison of Dynamic Regimes Using Observational Data," *The International Journal of Biostatistics*: Vol. 6: Iss. 2, Article 18.

**DOI:** 10.2202/1557-4679.1212

# When to Start Treatment? A Systematic Approach to the Comparison of Dynamic Regimes Using Observational Data

Lauren E. Cain, James M. Robins, Emilie Lanoy, Roger Logan, Dominique Costagliola, and Miguel A. Hernán

## Abstract

Dynamic treatment regimes are the type of regime most commonly used in clinical practice. For example, physicians may initiate combined antiretroviral therapy the first time an individual's recorded CD4 cell count drops below either 500 cells/mm<sup>3</sup> or 350 cells/mm<sup>3</sup>. This paper describes an approach for using observational data to emulate randomized clinical trials that compare dynamic regimes of the form "initiate treatment within a certain time period of some time-varying covariate first crossing a particular threshold." We applied this method to data from the French Hospital database on HIV (FHDH-ANRS CO4), an observational study of HIV-infected patients, in order to compare dynamic regimes of the form "initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>" where  $x$  takes values from 200 to 500 in increments of 10 and  $m$  takes values 0 or 3. We describe the method in the context of this example and discuss some complications that arise in emulating a randomized experiment using observational data.

**KEYWORDS:** dynamic treatment regimes, marginal structural models, HIV infection, antiretroviral therapy

**Author Notes:** This research was supported by NIH grant R01-AI073127. We thank Dr. Andrea Rotnitzky for her helpful comments.

## 1. INTRODUCTION

The goal of many observational studies is to compare the effects of two or more treatment regimes on a clinical outcome. Treatment regimes are dynamic when they depend on time-dependent covariates and static otherwise. For example, some guidelines for the use of combined antiretroviral therapy (cART) as a treatment for HIV recommend initiating treatment the first time the CD4 cell count drops below 350 cells/mm<sup>3</sup> (Panel on Antiretroviral Guidelines, 2008; Recommandations du groupe d'experts, 2008). This recommendation is an example of a dynamic regime (Robins and Hernán, 2008) because the initiation of treatment depends on the evolution of a time-varying covariate, CD4 cell count. In contrast, most randomized trials have compared static regimes like “initiate treatment at the beginning of the study” and “do not initiate treatment during the study”.

Static regimes are the type of regime most commonly compared in randomized clinical trials, but they are rarely used in clinical practice. Dynamic regimes, on the other hand, are rarely compared in clinical trials, but they are the type of regime most commonly used in clinical practice. For example, randomized trials have demonstrated that cART is an effective therapy to prevent AIDS and death in HIV-infected individuals (Cameron et al., 1998; Hammer et al., 1997), but the comparison of static regimes does not provide information on the optimal CD4 cell count at which to initiate treatment. Until randomized trials comparing dynamic regimes are conducted, we need to rely on observational studies of HIV-infected individuals to identify the optimal CD4 cell count at which to initiate cART.

Hernán et al. (2006) described an approach to emulate randomized clinical trials with two dynamic regimes using observational data. They applied the method to data from the French Hospital database on HIV (FHDH-ANRS CO4), an observational study of HIV-infected individuals, to compare dynamic regimes of the form “initiate treatment when the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  takes two values, e.g., 200 and 500. Later van der Laan and Petersen (2007), Orellana et al. (2010a, b), and Robins et al. (2008) proposed generalizations of the method to simultaneously compare many dynamic regimes. In this paper, we extend the analysis of the FHDH data from 2 to 31 dynamic regimes of the form “initiate treatment when the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  takes values from 200 to 500 in increments of 10.

The analysis by Hernán et al. compared regimes in which individuals initiate treatment immediately (during the same month) after their CD4 cell count crosses a particular threshold. Since the expectation of immediate action is often unrealistic due to administrative delays and other factors, regimes that allow delayed action may be more clinically relevant than regimes that require immediate action. In this paper, we extend the method to allow for delayed action by consi-

dering regimes of the form “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  takes values from 200 to 500 in increments of 10 and  $m > 0$ . The analysis by Hernán et al. was restricted to regimes with  $m = 0$ .

Below we describe how to emulate a randomized experiment involving multiple dynamic regimes using observational data, and how to address some complications that arise. For pedagogic reasons, we initially restrict our attention to dynamic regimes with  $m = 0$ , and then extend the method to dynamic regimes with  $m > 0$ . First, we provide a brief description of the FHDH data used in our analyses and introduce the notation used throughout the paper.

## **2. DATA AND NOTATION**

The French Hospital database on HIV (FHDH ANRS CO4) (Piketty et al., 2008) includes HIV-infected individuals seen at 62 French teaching hospitals belonging to 29 HIV Treatment and Information Centres (COREVIH) in mainland France and French overseas territories. Data have been collected since 1992 through medical records review by trained research assistants. Quality control is performed via monitored on-site source documentation (AUDIT on randomized individuals). Individuals are followed-up at the time of their clinic appointments, usually every three to four months, and the study attempts to collect information at least every six months. Death is ascertained through medical records review.

Our analysis was restricted to the 4,237 HIV-infected individuals who met the following eligibility criteria: age 18 years or older, antiretroviral therapy-naïve, no history of AIDS-defining illness (Ancelle-Park et al., 1993; CDC, 1992), no pregnancy, HIV-RNA  $> 500$  copies/mL, CD4 cell count and HIV-RNA measurements within six months of each other, and CD4 cell count between 200 and 500 cells/mm<sup>3</sup> with no history of CD4 cell count less than 500 cells/mm<sup>3</sup>. All analyses were conducted with SAS 9.2 (Cary, North Carolina).

An individual’s time zero was defined as the first time all of the above criteria were met. Time  $t$  is measured in months and ranges from 0 to 142. For each individual, follow-up ended at the time the outcome occurred, 12 months after the most recent laboratory measurement, pregnancy, or the administrative end of follow-up (December 2008), whichever occurred earlier. Initiation of cART was defined as the date at which an individual initiated use of either three or more antiretroviral drugs, or two ritonavir-boosted protease inhibitors, or one non-nucleoside reverse transcriptase inhibitor plus one boosted protease inhibitor.  $A_{it} = 1$  indicates that individual  $i$  has initiated treatment by time  $t$ , 0 otherwise. The outcome of interest was clinical AIDS or death. The date of death was identified as described elsewhere (HIV-CAUSAL Collaboration, 2010) and AIDS was ascertained by the treating physicians.  $D_{it} = 1$  indicates that individual  $i$  developed the outcome dur-

ing time  $t$ , 0 otherwise.  $L_{it}$  is a vector of individual  $i$ 's covariates measured at time  $t$ .  $V_i$  is a vector of the time-fixed covariates measured at time zero (a subset of  $L_{i0}$ ) that includes sex, age ( $<35$ ,  $35-49$ ,  $\geq 50$  years), geographic origin (Europe/North America, Sub-Saharan Africa, Latin America/Caribbean, other), mode of transmission (heterosexual, homosexual/bisexual, injection drug use, other or unknown), CD4 cell count (restricted cubic spline with three knots), HIV-1 RNA ( $<10,000$ ,  $10,000-100,000$ ,  $>100,000$  copies/mL), calendar year (1997-1998, 1999-2000, 2001-2003, 2004-2008), and years since HIV diagnosis ( $<1$ ,  $1-4$ ,  $\geq 5$  years, unknown). For simplicity, the analyses below assume that censoring due to infrequent laboratory measurement was ignorable given the measured baseline covariates  $V$ . This assumption can be relaxed easily by using time-varying covariates to estimate inverse probability weights as described by Hernán, Brumback and Robins (2001).

We use overbars to denote the history of a time-dependent variable. For example,  $\bar{A}_{it}$  represents individual  $i$ 's treatment history through time  $t$ , or  $\bar{A}_{it} = [A_{i0}, A_{i1}, A_{i2}, \dots, A_{it}]$ . Likewise,  $\bar{L}_{it}$  is individual  $i$ 's covariate history through time  $t$ . We often suppress the  $i$  subscript because we assume that the random vector for each individual is drawn independently from a distribution common to all individuals.

A static regime is defined as  $\bar{a} = [a_0, a_1, a_2, \dots, a_t]$ . For example,  $\bar{a} = [1, 1, 1, \dots, 1]$  and  $\bar{a} = [0, 0, 0, \dots, 0]$  are the static regimes “initiate treatment at the beginning of the study and continue throughout” and “do not initiate treatment during the study”, respectively. In contrast, we cannot define a dynamic regime as a sequence of 0s and 1s because the actual treatment at each time is not known until the time-varying covariates are measured. In this paper, we first consider the 31 dynamic regimes “initiate treatment in the same month that the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  takes values from 200 to 500 in increments of 10 and treatment initiation occurs during the same month that CD4 cell count crosses the threshold  $x$  (i.e.,  $m = 0$ ). In a slight abuse of notation, we index the dynamic regimes by  $x$ . Therefore,  $x = 350$  corresponds to the regime “initiate treatment in the same month that the recorded CD4 cell count first drops below 350 cells/mm<sup>3</sup>”.

### 3. EMULATION OF A RANDOMIZED EXPERIMENT

The preferred method for comparing the 31 dynamic regimes indexed by  $x$  is to conduct a randomized clinical trial with 31 arms. The first step of this hypothetical trial would be to identify individuals who meet the eligibility criteria given above. Second, we would randomly assign eligible individuals to one of the 31 regimes and follow them until AIDS, death, or the administrative end of follow-up. Third, we would compare the regime-specific AIDS-free survival. One simple

approach would be to compare the AIDS-free survival at a predefined time point (e.g., five years from randomization). Then, among these 31 regimes, the optimal regime would be the one that results in the greatest proportion surviving without AIDS after five years.

This experiment would require extremely large sample sizes and is unlikely to be conducted. In the absence of a randomized clinical trial, we can try to emulate one using observational data (Hernán et al., 2006). The first step in emulating the trial is to identify eligible individuals and observations using the same criteria as the randomized clinical trial. Second, we review each individual's CD4 cell count and treatment initiation histories to determine with which of the 31 regimes their data are consistent. If an individual's data are consistent with regime  $x$  through time  $t$ , we consider him to be "following" regime  $x$  through time  $t$ . If and when the individual's data are no longer consistent with following a particular regime, we artificially censor him at that time. For example, if an individual enters the study with a CD4 cell count of 205 cells/mm<sup>3</sup> and does not initiate treatment at that time, his data are consistent with the regime  $x = 200$  (i.e., initiate treatment in the same month that the recorded CD4 cell count first drops below 200 cells/mm<sup>3</sup>), so we say he is following the regime  $x = 200$ . He does not follow the other 30 regimes given he did not initiate treatment the first time his CD4 cell count dropped below 500, 490, 480, ..., nor 210 cells/mm<sup>3</sup>. Because his CD4 cell count is above 200 cells/mm<sup>3</sup>, he could still initiate treatment the first time his CD4 cell count drops below 200 cells/mm<sup>3</sup>. However, if he does not initiate treatment when his CD4 cell count first drops below 200 cells/mm<sup>3</sup>, he is artificially censored at that time. As in the randomized clinical trial, we follow each individual until AIDS, death, artificial censoring, or the administrative end of follow-up.

Third, we compare the (appropriately adjusted) AIDS-free survival across regimes at a predefined time point (e.g., five years from time zero). Among these 31 regimes, the optimal regime is the one that results in the greatest proportion surviving without AIDS after five years.

In order to emulate randomized experiments using observational data, we consider "having data consistent with a regime in the observational study" analogous to "following a regime in a randomized experiment with perfect adherence". There are, however, some key complications in the implementation of this approach. We now review them and propose some solutions.

### **Complication #1: Multiple regimes**

The individual described above (who enters the study with a CD4 cell count of 205 cells/mm<sup>3</sup> and does not initiate treatment at that time) is unusual in that he follows only one regime  $x$ . Most individuals have data that are consistent with

more than one regime  $x$ . For example, Table 1 shows data for six hypothetical individuals and the regimes they follow when  $m = 0$ . All six individuals enter the study with CD4 cell counts of 352 cells/mm<sup>3</sup> and are followed for four months. In the first month, their CD4 cell count increases to 380 cells/mm<sup>3</sup>. In the second month, their CD4 cell count drops to 273 cells/mm<sup>3</sup>. In the third month, there is no CD4 cell count measurement and thus the most recent value is carried forward. In the fourth month, their CD4 cell count drops to 198 cells/mm<sup>3</sup>. The only difference between the six individuals is the time at which they initiate treatment. Individual 1 initiates treatment the first time his CD4 cell count drops below 500 cells/mm<sup>3</sup> when his CD4 cell count was 352 cells/mm<sup>3</sup>. Therefore, individual 1's data are consistent with 15 regimes where  $x$  takes values from 360 to 500 in increments of 10 cells/mm<sup>3</sup> given he initiated treatment the first time his CD4 cell count dropped below 500, 490, 480, ..., and 360 cells/mm<sup>3</sup>. The other five individuals do not initiate treatment the first time their CD4 cell counts drop below 500 cells/mm<sup>3</sup>. Therefore, at time zero, their data are consistent with 16 regimes where  $x$  takes values from 200 to 350 in increments of 10 cells/mm<sup>3</sup> given they could still initiate treatment the first time their CD4 cell counts drop below 350, 340, 330, ..., and 200 cells/mm<sup>3</sup>.

Thus it is necessary to either randomly allocate an individual to one of the multiple regimes he is following or use a more statistically efficient approach where individuals are allowed to follow more than one regime simultaneously. To allow individuals to follow more than one regime, we make a replicate of each individual for each regime he follows at some time during the follow-up. We add a replicate-specific variable  $X$  to the dataset. The variable  $X$  for each individual's replicate is assigned a different value  $x$ . If and when a replicate with  $X = x$  deviates from the regime  $x$ , we artificially censor that replicate. Table S1 in the Supplemental Materials shows the expanded data for the same six hypothetical individuals from Table 1 for  $m = 0$ . Recall that at time zero individual 1 is following 15 regimes and the other five individuals are following 16 regimes. Therefore, 15 replicates are made of individual 1's follow-up and 16 replicates are made of each of the other five individuals' follow-ups. Individual 1 continues to follow the same 15 regimes for his entire follow-up and is never artificially censored. Individual 2 initiates treatment in the first month when his CD4 cell count is above his minimum CD4 cell count to date and, as a result, is censored from all 16 regimes. Individual 3 initiates treatment in the second month when his CD4 cell count is at a new minimum, 273 cells/mm<sup>3</sup>. The eight regimes where  $x$  takes values from 200 to 270 cells/mm<sup>3</sup> are censored in the second month, but individual 3 continues to follow the other eight regimes for the remainder of his follow-up.

Table 1: Six hypothetical individuals who follow multiple regimes in the class “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  takes the values 200 to 500 in increments of 10 and  $m$  takes values 0 and 3.

Individual	Time (months)	CD4 cell count	Treatment (1: yes, 0: no)	No. of regimes followed (range of $x$ )	
				$m = 0$	$m = 3$
1	0	352	1	15 (360-500)	15 (360-500)
1	1	380	1	15 (360-500)	15 (360-500)
1	2	273	1	15 (360-500)	15 (360-500)
1	3	273	1	15 (360-500)	15 (360-500)
1	4	198	1	15 (360-500)	15 (360-500)
2	0	352	0	16 (200-350)	31 (200-500)
2	1	380	1	0	15 (360-500)
2	2	273	1	0	15 (360-500)
2	3	273	1	0	15 (360-500)
2	4	198	1	0	15 (360-500)
3	0	352	0	16 (200-350)	31 (200-500)
3	1	380	0	16 (200-350)	31 (200-500)
3	2	273	1	8 (280-350)	23 (280-500)
3	3	273	1	8 (280-350)	23 (280-500)
3	4	198	1	8 (280-350)	23 (280-500)
4	0	352	0	16 (200-350)	31 (200-500)
4	1	380	0	16 (200-350)	31 (200-500)
4	2	273	0	8 (200-270)	31 (200-500)
4	3	273	1	0	23 (280-500)
4	4	198	1	0	23 (280-500)
5	0	352	0	16 (200-350)	31 (200-500)
5	1	380	0	16 (200-350)	31 (200-500)
5	2	273	0	8 (200-270)	31 (200-500)
5	3	273	0	8 (200-270)	16 (200-350)
5	4	198	1	8 (200-270)	16 (200-350)
6	0	352	0	16 (200-350)	31 (200-500)
6	1	380	0	16 (200-350)	31 (200-500)
6	2	273	0	8 (200-270)	31 (200-500)
6	3	273	0	8 (200-270)	16 (200-350)
6	4	198	0	0	16 (200-350)



Individual 4 is censored from the eight regimes where  $x$  takes values from 280 to 350 cells/mm<sup>3</sup> when his CD4 cell count drops in the second month. When he initiates treatment in the third month, he is censored from the remaining eight regimes because he initiates at either a repeated or carried forward CD4 cell count. Like individual 4, individual 5 is censored from eight regimes when his CD4 cell count drops in the second month. In the fourth month, he initiates treatment with a CD4 cell count that is a new minimum, 198 cells/mm<sup>3</sup>. He continues to follow the remaining eight regimes for the rest of his follow-up. Finally, individual 6 is identical to individual 5 except that he does not initiate treatment. Therefore, he is censored from eight regimes in the second month and from the remaining eight regimes in the fourth month.

The artificial censoring procedure to emulate this randomized experiment can be summarized as follows. Replicates can be censored for two different reasons. First, replicates with  $X = x$  are censored if and when the individual's CD4 cell count first drops below  $x$  and the individual does not initiate treatment. Second, replicates with  $X = x$  are censored if and when the individual initiates treatment before his CD4 cell count drops below  $x$ . Once an individual initiates treatment at an appropriate CD4 cell count for a particular replicate, that replicate becomes ineligible to be censored at a later time point.

The left portion of Table 2 shows the number of individuals that were eligible for 7 of the 31 dynamic regimes with  $m = 0$  where  $x$  takes values from 200 to 500 by 50 cells/mm<sup>3</sup> as well as the observed number of outcomes under each regime. When  $m = 0$ , individuals in the FHDH followed, on average, 22 regimes. Therefore, our expanded dataset had 92,460 individuals. Each replicate was “assigned” to a different regime and had a potentially different length of follow-up because of the differential artificial censoring by regime. Thus, before estimating the five-year AIDS-free survival that would have been observed if all individuals had followed every regime, we need to adjust for the bias this artificial censoring may create.

### Complication #2: Selection bias due to censoring

In the absence of confounding by unmeasured factors, as formalized in the strengthened identifiability conditions of exchangeability, positivity, and consistency described by Robins and Hernán (2008) we can eliminate the bias introduced by the artificial censoring if we weight each replicate of an individual by the individual's time-varying, unstabilized inverse probability weight

$W_t = \prod_{k=0}^t \frac{1}{f(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k)}$  where  $f(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k)$  is by definition the conditional probability mass function  $f_{A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k}(a_k | \bar{a}_{k-1}, d_k = 0, \bar{l}_k)$  with  $(a_k | \bar{a}_{k-1}, d_k = 0, \bar{l}_k)$  evaluated at the random argument  $(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k)$ ,

and  $A_{-1} = 0$ . Informally, the denominator of an individual's inverse probability weight at time  $t$  is the probability of having his own observed treatment history conditional on not developing the outcome by time  $t$  and his observed values of covariate history  $\bar{L}_t$  and treatment history  $\bar{A}_{t-1}$ . In our models for  $(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k)$ , we summarized the dependence on  $\bar{L}_t$  by  $V$  and  $L_t$ , the most recently available values of CD4 cell count (restricted cubic spline with five knots) and HIV-1 RNA ( $<10,000$ ,  $10,000-100,000$ ,  $\geq 100,000$  copies/mL) at time  $t$ , and months between time  $t$  and the most recent laboratory measurement (0, 1-2, 3-4, 5-6,  $\geq 7$ ).

Like in previous analyses of observational HIV data (Cole et al., 2003; Hernán et al., 2000; Hernán et al., 2002; Sterne et al., 2005) we assumed that treatment was never stopped once initiated. Therefore, for each individual, the factors in the denominator of the weights  $W_t$  were set to 1 for times  $t$  subsequent to treatment initiation, and estimated from the data for all other times, i.e., times when  $\bar{A}_{t-1} = 0$ . The conditional probability of treatment initiation was estimated by fitting the pooled logistic regression model  $\text{logit Pr}(A_t = 1 | \bar{A}_{t-1} = 0, D_t = 0, V, L_t) = \beta_{0t} + \beta_1' V + \beta_2' L_t$  where  $\beta_{0t}$  is a month-specific intercept (restricted cubic splines with four knots),  $\beta_1'$  and  $\beta_2'$  are the transposes of the column vectors of log hazard ratios for the components of the baseline covariates  $V$  and the time-varying covariates  $L_t$ , respectively. The logistic model was fit to and the weights  $W_t$  were estimated in the original, unexpanded study population. The time-varying weights estimated for an individual were then applied to all of his replicates when employing the approach described in the previous section. In the Appendix (Section 2) we show that, even when  $\beta_1' = \beta_2' = 0$  (i.e., the probability of initiating treatment did not depend on past covariate history), it is still necessary to use these inverse probability weights to prevent bias when estimating the effect of dynamic regimes on survival.

### **Complication #3: Unstable estimates**

After inverse probability weighting and artificial censoring of the replicates, one still needs to compare the AIDS-free survival across the regimes. A simple solution would be to create one inverse-probability-weighted Kaplan-Meier curve per regime of interest, and identify the regime with the greatest five-year AIDS-free survival. However, this method would result in unstable survival estimates when, as in most real applications, few individuals in the study population follow any given regime for a long time. In practice, to estimate the five-year AIDS-free survival for any given regime, we need to use a model that uses a smooth function  $h(X)$  to combine information from many different regimes.

We fit (to the expanded data including all replicates) the inverse probability weighted pooled logistic model  $\text{logit Pr}(D_{t+1} = 0 | D_t = 0, C_t = 0, X, V) =$

$\theta_{0t} + \theta_1' h(X) + \theta_2' V + \theta_3' h(X)t_m$  where  $h(X)$  is a restricted cubic spline for  $\frac{500-X}{100}$  with four knots at 0.50, 1.17, 1.83, and 2.50,  $h(X)t$  is the product (“interaction”) of  $h(X)$  with follow-up time  $t_m$  ( $\leq 6$ , 7-12, 13-24,  $>24$  months). The inclusion of the product terms allows the log hazard ratio for  $x$  to vary over time which is crucial since the assumption of a constant hazard ratio is substantively untenable for dynamic regimes.

Under the assumptions of strengthened exchangeability, positivity, and consistency (Robins and Hernán, 2008), the parameters of this inverse probability weighted model  $\gamma$  consistently estimate the parameters  $\psi$  of the dynamic marginal structural pooled logistic model  $\text{logit } \Pr(D_{t+1}^x = 0 | D_t^x = 0, V) = \theta_{0t} + \theta_1' h(x) + \theta_2' V + \theta_3' h(x)t_m$ . Here  $D_t^x$  is the counterfactual indicator that an individual would have developed the outcome during time  $t$  under regime  $X = x$ .

We then used the predicted values of the inverse probability weighted model to estimate the survival probability at each time  $t$  under each of the 31 regimes  $x$ .

#### Complication #4: Stabilization of the weights

The use of unstabilized inverse probability weights may result in highly unstable estimates, which makes this approach problematic. In practical implementations, stabilized weights are preferred. Unfortunately, as described below and in the Appendix (Section 3), the stabilization procedures commonly used for static regimes (Cole and Hernán, 2008) are not valid for dynamic regimes. We now describe one approach to stabilize the inverse probability weights  $W_t$ .

First note that, for a replicate with  $X = x$  who is uncensored through time  $t$ , the contribution  $f(A_t | \bar{A}_{t-1}, D_t = 0, \bar{L}_t)$  to the denominator of  $W_t$  is equal to the probability  $\Pr(C_t = 0 | C_{t-1} = 0, D_t = 0, X, \bar{L}_t, \bar{A}_{t-1})$  that the replicate remains uncensored through time  $t$  conditional on not developing the outcome by time  $t$ , covariate history, and treatment history. Table S1 in the Supplemental Materials shows the relation between the probability of treatment in the original dataset and the probability of remaining uncensored in the expanded dataset.

We argue in the Appendix (Section 3) that the numerator of any stabilized weight can depend on  $(X, V, D_t = 0)$  but cannot depend on  $\bar{A}_{t-1}$  or  $\bar{L}_t$ , which makes  $\Pr(C_t = 0 | C_{t-1} = 0, D_t = 0, X, V)$  a natural choice for the numerator of the stabilized weights. Thus, we define a replicate’s time-varying stabilized inverse probability weight to be  $SW_{t,x} = \prod_{k=0}^t \frac{\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, V)}{\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, \bar{L}_k, \bar{A}_{k-1})}$ . Note the denominator of an individual’s stabilized weight  $SW_{t,x}$  is equal to the denominator of his unstabilized weight  $W_t$ . Unfortunately, the stabilized weights  $SW_{t,x}$  are not guaranteed to produce estimates that are less variable than those obtained

using the unstabilized weights  $W_t$ . See the Appendix for more detail on the stabilized weights  $SW_{t,x}$ . Optimal, locally semiparametric efficient weights have been derived (Orellana et al., 2010a, b) for certain dynamic marginal structural models, but their implementation is less straightforward.

We now consider estimation of the numerator of the weights, the stabilizing factor. The probabilities in the numerator of  $SW_{t,x}$  were estimated by fitting the pooled logistic regression model  $\text{logit } \Pr(C_t = 0 | C_{t-1} = 0, D_t = 0, X, V) = \alpha_{0t} + \alpha'_1 h(X) + \alpha'_2 V$ . Note that we do not restrict this model to times when  $\bar{A}_{t-1} = 0$  since the numerator cannot be a function of  $A_t$  for any time  $t$  as mentioned above. Under the assumption that the denominator model is correct, our estimates of our marginal structural model will be consistent even if the model for the numerator of the weights is misspecified.

We truncated the stabilized weights to protect against misspecification of the model for the denominator of the weights, and against the near violations of positivity expected in small samples. When we truncated at 10.00, the mean estimated stabilized weight was 0.99 (range 0.01 to 10.00). The right portion of Table 2 shows the five-year survival for 7 of the 31 dynamic regimes with  $m = 0$  where  $x$  takes values from 200 to 500 by 50 cells/mm<sup>3</sup>. For example, the five-year survival was 0.95, 0.94, and 0.91 for the regimes  $x = 500$ , 350, and 200, respectively. We used a nonparametric bootstrap to calculate 95% confidence intervals for these survival estimates.

Table 2: Numbers of individuals, numbers of outcomes, and estimated five-year AIDS-free survivals under the regimes “initiate treatment in the same month that the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  ranges from 200 to 500 in increments of 50 cells/mm<sup>3</sup>.

Regime	No. of individuals*	No. of outcomes*	5-year AIDS-free survival	95% CI
500	288	10	0.95	0.91, 0.98
450	1,835	34	0.94	0.91, 0.97
400	2,891	54	0.93	0.90, 0.97
350	3,507	72	0.94	0.92, 0.96
300	3,764	89	0.94	0.92, 0.96
250	3,885	100	0.93	0.91, 0.95
200	3,949	114	0.91	0.89, 0.94

\* Each individual’s data may be consistent with his following several regimes.

#### 4. DYNAMIC REGIMES WITH A GRACE PERIOD

Thus far, we have considered regimes of the form “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $m = 0$  and  $x$  takes values from 200 to 500 in increments of 10. Under regimes with  $m = 0$  individuals are forced to initiate treatment immediately after the threshold  $x$  is crossed. We now describe how to extend our approach to the comparison of dynamic regimes where  $m > 0$ . Under regimes with  $m > 0$  individuals are given a grace period of  $m$  months after the threshold  $x$  is crossed before they are forced to initiate treatment. To illustrate the idea of dynamic regimes with a grace period  $m$ , we discuss regimes of the form “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $m = 3$ .

Recall that for  $m = 0$ , it was necessary to make a replicate of each individual for each regime he follows at some time during the follow-up and then artificially censor the replicate if and when the replicate deviates from a particular regime. The same applies when  $m > 0$ . However, the number of regimes followed and the timing of artificially censoring are affected by  $m$ . For example, consider the individuals in Table 1.

Individuals who initiate treatment the first time their CD4 cell count drops below 500 cells/mm<sup>3</sup> will follow the same regimes for their entire follow-up regardless of the value of  $m$ . For instance, individual 1 initiates treatment the first time his CD4 cell count drops below 500 cells/mm<sup>3</sup> when his CD4 cell count was 352 cells/mm<sup>3</sup>. Therefore, individual 1’s data are consistent with 15 regimes where  $x$  takes values from 360 to 500 in increments of 10 cells/mm<sup>3</sup> given he initiated the first time his CD4 cell count dropped below 500, 490, 480, ..., and 360 cells/mm<sup>3</sup>. Individual 1 will follow these same 15 regimes for his entire follow-up when  $m = 0$  and when  $m > 0$ .

Individuals who do not initiate treatment the first time their CD4 cell count drops below 500 cells/mm<sup>3</sup> follow different regimes over the course of their follow-up depending on the value of  $m$ . Individuals 2-6 do not initiate treatment the first time their CD4 cell counts drop below 500 cells/mm<sup>3</sup>. When  $m = 0$ , their data were consistent with the 16 regimes where  $x$  takes values from 200 to 350 in increments of 10 cells/mm<sup>3</sup> at time zero. However, when  $m > 0$ , their data are also consistent with the 15 regimes where  $x$  takes values from 360 to 500 in increments of 10 cells/mm<sup>3</sup> at time zero given that they could still initiate treatment within  $m$  months of their CD4 cell count first dropping below 500, 490, 480, ..., and 360 cells/mm<sup>3</sup>. The last column of Table 1 shows which regimes are followed by the six hypothetical individuals when  $m = 3$ .

As before, we make a replicate of each individual for each regime he follows at some time during the follow-up. When  $m = 3$ , 15 replicates are made of individual 1’s follow-up and 31 replicates are made of each of the other five indi-

viduals' follow-ups. Individual 1 continues to follow the same 15 regimes for his entire follow-up and is never artificially censored when  $m = 0$  nor when  $m = 3$ . Individual 2 initiates treatment in the first month when his minimum CD4 cell count to date was 352 cells/mm<sup>3</sup> and, as a result, is censored from the 16 regimes where  $x$  takes values from 200 to 350 cells/mm<sup>3</sup>. When  $m = 3$ , he continues to follow the other 15 regimes for the remainder of his follow-up. Individual 3 initiates treatment in the second month when his CD4 cell count is at a new minimum, 273 cells/mm<sup>3</sup>. The 8 regimes where  $x$  takes values from 200 to 270 cells/mm<sup>3</sup> are censored in the second month, but individual 3 continues to follow the other 23 regimes for the remainder of his follow-up when  $m = 3$ . Individual 4 initiates treatment in the third month when his CD4 cell count remains 273 cells/mm<sup>3</sup> having been carried forward from the second month. At that time, he is censored from the 8 regimes where  $x$  takes values from 200 to 270 cells/mm<sup>3</sup>, but continues to follow the other 23 regimes for the remainder of his follow-up when  $m = 3$ . Individual 5 does not initiate treatment until the fourth month. Therefore, when  $m = 3$ , he is censored from the 15 regimes where  $x$  takes values from 360 to 500 cells/mm<sup>3</sup> in the third month, but continues to follow the other 16 regimes for the remainder of his follow-up. Like individual 5, when  $m = 3$ , individual 6 is censored from the 15 regimes where  $x$  takes values from 360 to 500 cells/mm<sup>3</sup> in the third month, but continues to follow the other 16 regimes for the remainder of his follow-up. Table S2 in the Supplemental Materials shows the expanded data for these six hypothetical individuals from Table 1 for  $m = 3$ .

For  $m = 0$ , the factors in the denominator of the weights were estimated from the data for all times when  $\bar{A}_{t-1} = 0$ , and set to 1 for all times when  $\bar{A}_{t-1} = 1$ . For  $m > 0$ , there are additional times when the factor in the denominator of the weights  $SW_{t,x}$  must be set to 1. Specifically, the month-specific factor in the denominator of the weight must be set to 1 for any month in which a replicate was not eligible to be censored. Note that when the value of  $x$  is greater than the value of the CD4 cell count at time zero, a replicate is ineligible for censoring prior to month  $m$ . For example, in the  $m = 3$  column of Table 1, individuals 5 and 6 are censored from the 15 regimes where  $x$  takes values from 360 to 500 cells/mm<sup>3</sup> in the third month. Additionally, when the CD4 cell count first drops below  $x$ , regimes where  $X = x$  cannot be censored for  $m$  months. For example, individual 6 cannot be censored from regimes where  $x$  takes values from 280 to 350 cells/mm<sup>3</sup> in the second, third, or fourth month because his CD4 cell count first drops below 350, 340, 330, ..., 280 cells/mm<sup>3</sup> in the second month. These regimes become eligible for censoring again in the fifth month.

For  $m = 0$ , the factors in the numerator of the weights were estimated from the data for all times including when  $\bar{A}_{t-1} = 1$  since the numerator cannot be a function of  $A_t$  for any time  $t$  as mentioned above. For  $m > 0$ , there are additional times when the factor in the numerator of the weights  $SW_{t,x}$  must be set to 1.

When the value of  $x$  is greater than the value of the CD4 cell count at time zero, a replicate is ineligible for censoring prior to month  $m$  as is the case in the denominator. However, when the CD4 cell count first drops below  $x$ , regimes where  $X = x$  are immediately eligible for censoring in the numerator since the numerator cannot be a function of time-varying covariates. Individuals follow more regimes when  $m > 0$  than when  $m = 0$ . As  $m$  increases individuals follow more regimes for longer periods of time, which results in more precise survival estimates because more events are included in the analysis.

When  $m = 3$ , individuals in the FHDH followed, on average, 30 regimes. Therefore, our expanded dataset had 125,010 individuals. Table 3 shows the number of individuals that were eligible for the seven regimes where  $x$  takes values from 200 to 500 by 50 cells/mm<sup>3</sup>, as well as the observed number of outcomes and the estimated five-year AIDS-free survival under each regime when the stabilized weights were truncated at 10.00 (mean: 1.03, range 0.00 to 10.00). For example, the five-year AIDS-free survival was 0.92, 0.93, and 0.91 for the regimes  $x = 500$ , 350, and 200, respectively. In the Appendix (Section 4) we provide additional discussion of the clinical meaning of the regime “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” with grace period  $m > 0$ .

Table 3: Numbers of individuals, numbers of outcomes, and estimated five-year AIDS-free survivals under the regimes “initiate treatment within 3 months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>” where  $x$  ranges from 200 to 500 in increments of 50 cells/mm<sup>3</sup>.

Regime	No. of individuals*	No. of outcomes*	5-year AIDS-free survival	95% CI
500	4,237	58	0.92	0.86, 0.98
450	4,146	82	0.92	0.88, 0.96
400	4,072	94	0.93	0.89, 0.96
350	4,027	105	0.93	0.91, 0.95
300	3,986	119	0.93	0.91, 0.95
250	3,970	119	0.92	0.90, 0.94
200	3,949	125	0.91	0.88, 0.93

\* Each individual’s data may be consistent with his following several regimes.

## 5. DISCUSSION

We have described how to use dynamic marginal structural models to identify the optimal treatment regime in a set of regimes. Our results, although imprecise, suggest that the optimal CD4 threshold to initiate treatment was above 350 cells/mm<sup>3</sup> whether we considered immediate or delayed (up to 3 months) treatment initiation after the threshold is crossed. We defined “optimal regime” in terms of time to AIDS or death, whichever happened first. Had we considered an outcome other than AIDS or death (e.g., death alone or serious non-AIDS defining illnesses), we might have found that a different regime is the optimal one.

A strength of our approach is that it eliminates the possibility of “lead time” bias (Cole et al., 2004) by design. However, the validity of our effect estimates relies on a number of assumptions.

First, we made the untestable assumption of no unmeasured confounding (or conditional exchangeability) given the measured covariates, i.e., the assumption that  $\bar{L}_t$  includes all joint predictors of treatment initiation and the outcome. This assumption may approximately hold here because  $\bar{L}_t$  included time-varying CD4 cell count and HIV-RNA, the most important clinical measures used by physicians as indications for treatment initiation. To further protect our estimates from unmeasured confounding, we analyzed our data under the intent-to-treat principle used in the analysis of randomized clinical trials. We defined the dynamic treatment regimes in terms of treatment initiation under whatever degree of subsequent adherence to treatment existed in our study population. This strategy makes it unnecessary to adjust for joint determinants of treatment discontinuation and the outcome, which are less well-measured in most observational studies, at the expense of potential bias due to misclassification of treatment status. Note, however, that adjusting for treatment discontinuation would not be appropriate, nor clinically interesting, if most individuals who discontinue treatment do so for toxicity-related reasons.

Second, we assumed a correct specification of the model for treatment initiation as a function of the measured confounders. If this model is misspecified, the weights could be extreme and lead to bias. To prevent this bias and to mitigate the effect of any near violations of positivity, we truncated the estimated weights at the value 10.00, which was at or above the 99<sup>th</sup> percentile of the distribution of the estimated weight. Other levels of truncation (e.g., 50.00) yielded virtually the same estimates. Finally, the validity of our method also depends on modeling assumptions involving  $h(x)$ , the function of the time-fixed covariate regime. Our choice to consider  $x$  in increments of 10 cells/mm<sup>3</sup> was not completely arbitrary. Increments of less than 10 cells/mm<sup>3</sup> are not clinically relevant and may also be too dependent on the accuracy of measurement. Increments of greater than 10



cells/mm<sup>3</sup> will limit the flexibility of  $h(x)$ . Alternative parameterizations of  $h(x)$  (quadratic, cubic, and quartic polynomials) resulted in similar estimates.

The methods presented in this paper can be extended to more complex dynamic regimes. For instance, if we were interested in death alone as the outcome, we might consider regimes of the form “initiate cART within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup> or an AIDS diagnosis, whichever occurs earlier”. In these dynamic regimes, the initiation of treatment depends on the evolution of two time-varying covariates, CD4 cell count and AIDS. We might also consider another randomized experiment where individuals enter the study with CD4 cell counts above 500 cells/mm<sup>3</sup> and receive a fixed intervention until their CD4 cell count drops below 500 cells/mm<sup>3</sup>. The regimes in this experiment would be of the form “do not initiate treatment when CD4 cell count is above 500 cells/mm<sup>3</sup> and initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$  cells/mm<sup>3</sup>”.

In summary, this paper described a method to emulate randomized experiments with dynamic regimes using observational data. To obtain more stable estimates, these analyses will need to be conducted using a larger dataset. Also, as more longitudinal observational data become available, more questions involving dynamic regimes can be answered. Though estimates obtained from observational data will always be suspect because of the possibility of residual confounding, the application of this method may help identify promising regimes to be compared in randomized clinical trials.

## APPENDIX

### A1. General theory

The main text considers regimes  $x$  with grace period  $m$  of the form “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ ”. Formally this means that, under regime  $x$  with grace period  $m$ , individuals are prevented from starting treatment before their CD4 cell count drops below  $x$ , and are then forced to initiate treatment exactly  $m$  months after the time that their CD4 cell count first drops below  $x$ , if they are still alive and have not initiated treatment during the  $m$ -month grace period on their own.

Let  $T^x$  denote an individual’s counterfactual failure time under regime  $x$ . Let counterfactual death indicator  $D_t^x = 1$  if  $T^x \leq t$  and  $D_t^x = 0$  otherwise for  $t = 0, 1, 2, \dots, K$ , where  $K$  is the known maximum possible follow-up time

of any individual. Consider the dynamic marginal structural discrete hazard model  $\Pr [D_{t+1}^x = 1 | D_t^x = 0, V = v] = \lambda(t, x, v; \beta^*)$  where  $V$  is a subset of the baseline covariates  $L(0)$ ,  $\lambda(t, x, v; \beta)$  is a known function taking values between 0 and 1, and  $\beta^*$  is the true value of the parameter  $\beta$ .

Let  $\bar{L}_t, \bar{A}_t$  be an individual's covariate history and treatment history through  $t$ , respectively. As in the text, we assume  $A_t = 1$  implies  $A_{t+1} = 1$  since treatment once begun is never stopped. Let  $Q_x$  denote the time at which an individual's recorded CD4 cell count drops below  $x$ . Define  $C_{t,x} = 0$  if an individual's observed data is consistent with having followed regime  $x$  through time  $t$  and  $C_{t,x} = 1$  otherwise. Then, by definition,

$$\begin{aligned} C_{t,x} &= 0 \text{ if and only if both} \\ A_j &= 0 \text{ for } j < \min(Q_x, t+1, T) \\ \text{and} \\ A_{Q_x+m} &= 1 \text{ whenever } Q_x + m \leq \min(t, T) \end{aligned}$$

Thus  $C_{t,x}$  is a deterministic function of  $\bar{L}_t, \bar{A}_t, \bar{D}_t$ , and  $x$ . Note, in particular, if  $C_{t,x} = 0$  and either  $D_{t+1} = 1$  or  $A_t = 1$ , then  $C_{t',x} = 0$  for  $t' > t$  since, by definition, the individual continues to follow regime  $x$  after failure or if he started treatment while following the regime. We make the consistency assumption that

$$D_{t+1}^x = D_{t+1} \text{ if } C_{t,x} = 0,$$

the exchangeability assumption

$$\{\bar{D}_K^x\} \perp\!\!\!\perp A_t | \bar{L}_t, \bar{A}_{t-1} = 0, D_t = 0,$$

where  $\bar{D}_K^x = \{D_0^x, \dots, D_K^x\}$ , and the positivity assumption that

$$1 > \Pr[A_t = 1 | \bar{L}_t, \bar{A}_{t-1} = 0, D_t = 0] > 0 \text{ with probability 1.}$$

The above exchangeability and positivity assumptions for treatment imply the following exchangeability and positivity assumptions for censoring:

$$\{\bar{D}_K^x\} \perp\!\!\!\perp C_{t,x} | C_{t-1,x} = 0, \bar{L}_t, \bar{A}_{t-1}, D_t = 0, \text{ and}$$

$$\Pr [C_{t,x} = 0 | C_{t-1,x} = 0, \bar{L}_t, \bar{A}_{t-1}, D_t = 0] > 0 \text{ with probability 1}$$

because  $C_{t-1,x}$  is a deterministic function of  $\bar{L}_t, \bar{A}_{t-1}, D_t = 0$  and, conditional on  $D_t = 0$  and a given realization of  $\bar{L}_t, \bar{A}_{t-1}$ , either (i)  $C_{t,x} = A_t$ , (ii)  $C_{t,x} = 1 - A_t$ , or (iii)  $C_{t,x}$  has a degenerate distribution at 0.

Define

$$W_{t,x} = 1 / \prod_{j=0}^t \Pr [C_{j,x} = 0 | C_{j-1,x} = 0, \bar{L}_j, \bar{A}_{j-1}, D_j = 0]$$

Given the consistency, exchangeability, and positivity assumptions for censoring, we have  $E [I \{C_{t,x} = 0\} W_{t,x} | V, \bar{D}_K^x] = 1$ . Arguing as in Orellana et al. (2010a, b), the set of all unbiased estimating functions (which is also the linear space spanned by all influence functions) for  $\beta^*$  when  $f(A_t | \bar{A}_{t-1}, \bar{L}_t, D_t = 0)$  is known is given by  $P_n\{U(\beta, \mathbf{q}) + M(\mathbf{h})\}$  as  $\mathbf{q}$  and  $\mathbf{h}$  are varied arbitrarily, where  $P_n$  denotes a sample average over the  $n$  study population members,

$$U(\beta, \mathbf{q}) = \sum_x \sum_{t=0}^K \{D_{t+1} - \lambda(t, x, V; \beta)\} (1 - D_t) q_t(x, V) I\{C_{t,x} = 0\} W_{t,x},$$

$$M(\mathbf{h}) = \sum_{t=0}^K \{A_t - \Pr(A_t = 1 | \bar{A}_{t-1}, \bar{L}_t, D_t = 0)\} \{1 - D_t\} I\{A_{t-1} = 0\} h_t(\bar{L}_t),$$

and  $\mathbf{q} = \{q_1, \dots, q_K\}$ ,  $\mathbf{h} = \{h_1, \dots, h_K\}$  are sets of arbitrary functions of  $(x, V)$  and  $\bar{L}_t$  respectively (Robins and Rotnitzky, 1992).

The estimators of  $\beta^*$  in the main text solve  $P_n\{U(\beta, \mathbf{q}) + M(\mathbf{h})\} = 0$  with each  $h_t(\bar{L}_t) \equiv 0$  with the estimated weight function  $\widehat{W}_{t,x}$  replacing  $W_{t,x}$  for two different choices of  $\mathbf{q}$  depending on whether unstabilized or stabilized weights (see below) are used. As noted in the text, when  $D_t = 0$ ,  $I\{C_{t,x} = 0\} W_{t,x}$  is equal to

$$I\{C_{t,x} = 0\} \left\{ \begin{aligned} & f[A_{Q_x+m} | \bar{L}_{Q_x+m}, \bar{A}_{Q_x+m-1}, D_{Q_x+m} = 0]^{I(Q_x+m \leq t, A_{Q_x+m-1}=0)} \\ & \times \prod_{j=0}^{\min(t, Q_x-1)} f[A_j | \bar{L}_j, \bar{A}_{j-1}, D_j = 0] \end{aligned} \right\}^{-1}$$

which can be written as  $I\{C_{t,x} = 0\} / \prod_{j=0}^t f[A_j | \bar{L}_j, \bar{A}_{j-1}, D_j = 0]$  when  $m = 0$ .

When  $m > 0$ , the time-specific contribution to the weight  $W_{t,x}$  is 1 at times  $Q_x, Q_x + 1, \dots, Q_x + m - 1$ .

## A2. Need for inverse probability weighting even when treatment probabilities are constant

Suppose that the grace period  $m$  equals 0 and that  $\Pr[A_t = 1 | \bar{L}_t, \bar{A}_{t-1} = 0, D_t = 0]$  is a constant  $p_t$  that does not depend on  $\bar{L}_t$  as might be the case in

an experiment that randomly assigned a month of treatment initiation to each individual. Then one might hope that one could ignore the weight  $W_{t,x}$  and use  $P_n \left\{ \sum_x \sum_{t=0}^K \{D_{t+1} - \lambda(t, x, V; \beta)\} (1 - D_t) q_t(x, V) I\{C_{t,x} = 0\} \right\}$  as an unbiased estimating equation. Since this latter estimating equation can be written as

$$P_n \left\{ \sum_x \sum_{t=0}^K \{D_{t+1} - \lambda(t, x, V; \beta)\} (1 - D_t) q_t(x, V) I\{C_{t,x} = 0\} (W_{t,x}/W_{t,x}) \right\},$$

it is clear that this approach gives an unbiased estimating equation only if  $1/W_{t,x}$  is a function just of  $(x, V)$ . However, we now show that  $W_{t,x}$  is in fact a function of  $\bar{L}_K$  through the CD4 cell count history. To see this note that the weight  $W_{t,x}$  for an individual with  $C_{t,x} = 0$  is given by  $W_{t,x} = \{p_{Q_x}\}^{I\{Q_x \leq t\}}$

$\prod_{j=0}^{\min(t, Q_x-1)} (1 - p_j)$ , which is a function of an individual's time-dependent CD4 cell count through  $Q_x$ . Note that  $W_{t,x}$  is a function of  $Q_x$ , even if  $p_j = 1/2$  for every  $j$  because then  $W_{t,x} = (1/2)^{\min(t, Q_x)}$ .

### A3. Restrictions on the stabilized weights

If we wish to substitute a stabilized weight  $SW_{t,x} = Num_{t,x} \times W_{t,x}$  for the unstabilized weight  $W_{t,x}$ ,  $Num_{t,x}$  must be a function of  $(x, V)$  only; otherwise  $\sum_x \sum_{t=0}^K \{D_{t+1} - \lambda(t, x, V; \beta)\} (1 - D_t) q_t(x, V) I\{C_{t,x} = 0\} SW_{t,x}$  would not be in the aforementioned set of unbiased estimating functions. In particular,

we cannot take  $Num_{t,x} = \prod_{j=0}^t f(A_t | \bar{A}_{t-1}, D_t = 0)$ , as we typically do for a nondynamic (static) marginal structural model.

### A4. Alternative, possibly more clinically relevant regimes

Following Robins and Rotnitzky (unpublished technical report) we now argue that one may wish to consider regimes consistent with “initiating treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ ” other than the regimes  $x$  with grace period  $m$  in the main text. To see why, suppose that  $m = 6$  and the observed probability of starting treatment in each of the months  $0, 1, \dots, m$  after the CD4 cell count first drops below  $x$  is 1%. Then,

under regime  $x$  with grace period  $m$ , roughly 94% of individuals would start treatment in the 6th month after their CD4 cell count fell below  $x$ . However, if one recommended to physicians that patients should “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ ”, the distribution of start times might actually be closer to uniform over the months  $0, 1, \dots, m$ . If that were so, we would like to find the  $x$  that optimizes survival under the set of regimes “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ , such that there is a uniform probability of starting in each of months  $0, 1, \dots, m$ .” Below we describe how to estimate the optimal  $x$  for a wide variety of regime sets, each consistent with the requirement that an individual “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ .”

The larger point of this subsection is that counterfactual survival corresponding to the regime “initiate treatment within  $m$  months after the recorded CD4 cell count first falls below  $x$ ” is vague and ill-defined because there are many different regimes (i.e., versions of treatment) consistent with this regime. The problem of vague counterfactuals due to many versions of treatment has been frequently discussed in the literature (Robins and Greenland, 2000; Hernán and Taubman, 2008; VanderWeele, 2009); what is interesting in our setting is that, as we next show, one can actually identify the effect on survival of many different versions of this regime, once the versions are formally defined. See Taubman et al. (2008) for related results and discussion on different versions of treatment regimes.

Consider the set of regimes defined as follows. For fixed  $m$  and  $x$ , and a set of conditional probabilities  $\Pr_{j,t,x} [A_{t+j} = 1 | \bar{L}_{t+j}, A_{t+j-1} = 0, D_{t+j} = 0]$  of starting treatment conditional on  $\bar{L}_{t+j}$ , indexed by  $x, t, j$ , with  $j = 0, \dots, m$ , satisfying  $\Pr_{m,t,x} [A_{t+m} = 1 | \bar{L}_{t+m}, A_{t+m-1} = 0, D_{t+m} = 0] = 1$ , consider the (random) dynamic regime  $x$  in which (i) individuals are prevented from starting treatment before their CD4 cell count drops below  $x$ , and (ii) individuals initiate treatment at time  $t+j$  with probability  $\Pr_{j,t,x} [A_{t+j} = 1 | L_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$ , provided they are alive and at risk to initiate treatment at  $t+j$ , if the time  $Q_x$  their recorded CD4 cell count first drops below  $x$  is equal to  $t$ . Under this regime, all individuals alive at  $t+m$  who are yet to initiate treatment will initiate at  $t+m$ .

Note an individual’s counterfactual failure time  $T^x$  under this regime  $x$  is well defined (as a stochastic counterfactual) but different for each choice of  $m$  and the initiation probabilities  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, A_{t+j-1} = 0, D_{t+j} = 0]$ . To estimate the optimal  $x$  for a given  $m$  and a given choice of the probabilities  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, A_{t+j-1} = 0, D_{t+j} = 0]$ , we can estimate the parameters

of a dynamic marginal structural discrete hazard model

$$\Pr [D_{t+1}^x = 1 | D_t^x = 0, V = v] = \lambda(t, x, v; \beta^*)$$

for  $T^x$  with the weights  $W_{t,x}$  redefined to be

$$W_{t,x} = \left\{ \prod_{j=0}^{\min(t, Q_x-1)} \Pr [A_j = 0 | \bar{L}_j, \bar{A}_{j-1} = 0, D_j = 0] \right\}^{-1} \times \prod_{j=0}^m \left\{ \frac{f_{j,Q_x,x}^* [A_{Q_x+j} | \bar{L}_{Q_x+j}, \bar{A}_{Q_x+j-1} = 0, D_{Q_x+j} = 0]}{f[A_{Q_x+j} | \bar{L}_{Q_x+j}, \bar{A}_{Q_x+j-1} = 0, D_{Q_x+j} = 0]} \right\}^{I(Q_x+j \leq t, D_{Q_x+j}=0, \bar{A}_{Q_x+j-1}=0)}$$

where

$$f_{j,t,x}^* [a | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0] = \Pr_{j,t,x}^* [A_{t+j} = a | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$$

for  $a = 0, 1$ . If we choose  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$  to be the observed probability  $\Pr [A_{t+j} = 1 | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$ , then the distribution of  $T^x$  is the same as it is under the regime  $x$  with grace period  $m$  considered in the main text and the redefined  $W_{t,x}$  equals the earlier  $W_{t,x}$ . On the other hand, if we choose  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0] = 1/(m+1-j)$ , then our regime has a roughly uniform distribution of starting times with the probability of starting in each of months  $0, 1, \dots, m$  after the CD4 cell count first drops below  $x$  approximately  $1/(m+1)$ . The distribution would be precisely uniform in the limit as the number of deaths in any  $m+1$  month interval goes to zero. Note that although the choice  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0] = 1/(m+1-j)$  produces a regime that satisfies the condition “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ , such that there is a uniform probability of starting in each of months  $0, 1, \dots, m$ ”, it is not the only one. Specifically, there will exist choices of  $\Pr_{j,t,x}^* [A_{t+j} = 1 | \bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$  that allow dependence on  $\bar{L}_{t+j}$  but still satisfy the requirement that the marginal probabilities of starting in each of months  $0, 1, \dots, m$  are uniform. That is, the regime “initiate treatment within  $m$  months after the recorded CD4 cell count first drops below  $x$ , such that there is a uniform probability of starting in each of months  $0, 1, \dots, m$ ” still contains many versions of the treatment regime.

We now sketch the proof for the consistent estimation of the parameters of the dynamic marginal structural model for  $T^x$  using the redefined  $W_{t,x}$ .

*Proof sketch:* Under the above exchangeability, positivity, and consistency assumptions, it follows from Robins (1987) that the joint distribution of  $D_t^x$ ,  $t = 0, \dots, k$  given  $V$  under the regime defined by  $x, m, f_{j,t,x}^*, [a|\bar{L}_{t+j}, \bar{A}_{t+j-1} = 0, D_{t+j} = 0]$ ,  $j = 0, \dots, m$ , is given by the marginal distribution of  $D_t$ ,  $t = 0, \dots, k$  and  $V$  under the (so-called g-formula) joint distribution

$$\prod_{j=0}^{\min(k, \text{int}(T))} f(D_{j+1}, L_{j+1} | \bar{D}_j = 0, \bar{L}_j, \bar{A}_j) \tilde{f}[A_j | \bar{L}_j, \bar{A}_j - 1, D_j = 0],$$

where  $f(D_{j+1}, L_{j+1} | \bar{D}_j = 0, \bar{L}_j, \bar{A}_j)$  is the conditional density based on the distribution of the observed data, and  $\tilde{f}[A_t | \bar{L}_t, \bar{A}_{t-1}, D_t = 0]$  is (i) 1 if  $\bar{A}_{t-1} = 1$ , (ii) 0 if  $t < Q_x$ , (iii)  $f_{j,t-j,x}^*[A_t | \bar{L}_t, \bar{A}_{t-1} = 0, D_t = 0]$  if  $t = Q_x + j$ ,  $j = 0, \dots, m-1$ ,  $A_{t-1} = 0$ ,  $D_t = 0$ , (iv) 1 if  $t = Q_x + m$ ,  $A_{t-1} = 0$ . The likelihood under the observed data generating mechanism for data up to  $k$  is

$$\prod_{j=0}^{\min(k, \text{int}(T))} f(D_{j+1}, L_{j+1} | \bar{D}_j = 0, \bar{L}_j, \bar{A}_j) f[A_j | \bar{L}_j, \bar{A}_j - 1, D_j = 0].$$

The likelihood ratio is

$$\prod_{j=0}^{\min(k, \text{int}(T))} \frac{\tilde{f}[A_j | \bar{L}_j, \bar{A}_j - 1, D_j = 0]}{f[A_j | \bar{L}_j, \bar{A}_j - 1, D_j = 0]}$$

which is precisely the redefined  $W_{k,x}$  for an individual with  $D_k = 0$ . Because

$$\sum_x \sum_{k=0}^K \{D_{k+1} - \lambda(k, x, V; \beta)\} (1 - D_k) q_k(x, V)$$

is an unbiased estimating function for  $\beta^*$  under the law

$$\prod_{j=0}^{\min(k, \text{int}(T))} f(D_{j+1}, L_{j+1} | \bar{D}_j = 0, \bar{L}_j, \bar{A}_j) \tilde{f}[A_j | \bar{L}_j, \bar{A}_j - 1, D_j = 0],$$

the result now follows from the Radon-Nikodyn theorem.

## REFERENCES

- Ancelle-Park, R., Klein, J. P., Stroobant, A., *et al.* (1993). Expanded European AIDS case definition. *The Lancet* **341**, 441.
- Cameron, D. W., Heath-Chiozzi, M., Danner, S., *et al.* (1998). Randomised placebo-controlled trial of zidovudine in advanced HIV-1 disease. The Advanced HIV Disease Zidovudine Study Group. *The Lancet* **351**, 543-549.
- CDC (1992). 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *Morbidity and Mortality Weekly Report* **41**, 1-19.
- Cole, S. R., and Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology* **168**, 656-664.
- Cole, S. R., Hernán, M. A., Robins, J. M., *et al.* (2003). Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *American Journal of Epidemiology* **158**, 687-694.
- Cole, S. R., Li, R., Anastos, K., *et al.* (2004). Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies. *Statistics in Medicine* **23**, 3351-3363.
- Hammer, S. M., Squires, K. E., Hughes, M. D., *et al.* (1997). A controlled trial of two nucleoside analogues plus zidovudine in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *The New England Journal of Medicine* **337**, 725-733.
- 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.
- Hernán, M. A., Brumback, B., and Robins, J. M. (2001). Marginal structural models to estimate the joint causal effect of non-randomized treatments. *Journal of the American Statistical Association* **96**, 440-448.



- Hernán, M. A., Brumback, B. A., and Robins, J. M. (2002). Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Statistics in Medicine* **21**, 1689-1709.
- Hernán, M. A., Lanoy, E., Costagliola, D., and Robins, J. M. (2006). Comparison of dynamic treatment regimes via inverse probability weighting. *Basic & Clinical Pharmacology & Toxicology* **98**, 237-242.
- Hernán M.A., and Taubman S.L. (2008). Does obesity shorten life? The importance of well defined interventions to answer causal questions. *International Journal of Obesity* **32**, S8-S14.
- HIV-CAUSAL Collaboration (2010). The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. *AIDS* **24**, 123-137.
- Orellana, L., Rotnitzky, A., and Robins, J. M. (2010a). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main Content. *The International Journal of Biostatistics* **6**, Article 7.
- Orellana, L., Rotnitzky, A., and Robins, J. M. (2010b). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part II: Proofs and Additional Results. *The International Journal of Biostatistics* **6**, Article 8.
- Panel on Antiretroviral Guidelines for Adults and Adolescents (2008). *Guidelines for the use of antiretroviral agents in HIV-1 infected adults and adolescents*: Department of Health and Human Services.
- Piketty, C., Selinger-Leneman, H., Grabar, S., *et al.* (2008). Marked increase in the incidence of invasive anal cancer among HIV-infected patients despite treatment with combination antiretroviral therapy. *AIDS* **22**, 1203-1211.
- Recommandations du groupe d'experts sous la direction du Pr P Yeni (2008). *Prise en charge des personnes infectées par le VIH*: Flammarion Médecines-Sciences.

- Robins, J.M. (1987). Addendum to "A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect." *Computers & Mathematics with Applications* **14**, 923-945 (errata in *Computers & Mathematics with Applications* **18**, 477).
- Robins, J. M., and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology: Methodological Issues*, N. P. Jewell, K. Dietz, and V. T. Farewell (eds), 297-331 (includes errata sheet). Boston, MA: Birkhäuser.
- Robins, J.M., and Greenland, S. (2000). Comment on "Causal Inference Without Counterfactuals" by A.P. Dawid. *Journal of the American Statistical Association -- Theory and Methods* **95**, 477-482.
- Robins, J., and Hernán, M. (2008). Estimation of the causal effects of time-varying exposures. In *Longitudinal Data Analysis*, G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs (eds), 553-599. New York: Chapman and Hall/CRC Press.
- Robins, J., Orellana, L., and Rotnitzky, A. (2008). Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine* **27**, 4678-4721.
- Sterne, J. A., Hernán, M. A., Ledergerber, B., *et al.* (2005). Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: a prospective cohort study. *The Lancet* **366**, 378-384.
- Taubman, S.L., Robins, J.M., Mittleman, M.A., and Hernán, M.A. (2008). Alternative approaches to estimating the effects of hypothetical interventions. In: *JSM Proceedings, Health Policy Statistics Section*. Alexandria, VA: American Statistical Association.
- van der Laan, M. J., and Petersen, M. L. (2007). Causal effect models for realistic individualized treatment and intention to treat rules. *The International Journal of Biostatistics* **3**, Article 3.
- VanderWeele, T.J. (2009). Concerning the consistency assumption in causal inference. *Epidemiology* **20**, 880-883.