


RESEARCH ARTICLE

Emulating a trial of joint dynamic strategies: An application to monitoring and treatment of HIV-positive individuals

Ellen C. Caniglia^{1,2}  | James M. Robins^{1,3} | Lauren E. Cain¹ | Caroline Sabin⁴ | Roger Logan¹ | Sophie Abgrall⁵ | Michael J. Mugavero⁶ | Sonia Hernández-Díaz¹ | Laurence Meyer⁷ | Remonie Seng⁷ | Daniel R. Drozd⁸ | George R. Seage III¹ | Fabrice Bonnet⁹ | Fabien Le Marec⁹ | Richard D. Moore¹⁰ | Peter Reiss¹¹ | Ard van Sighem¹¹ | William C. Mathews¹² | Inma Jarrín¹³ | Belén Alejos¹³ | Steven G. Deeks¹⁴ | Roberto Muga¹⁵ | Stephen L. Boswell¹⁶ | Elena Ferrer¹⁷ | Joseph J. Eron¹⁸ | John Gill¹⁹ | Antonio Pacheco²⁰ | Beatriz Grinsztejn²⁰ | Sonia Napravnik¹⁸ | Sophie Jose⁴ | Andrew Phillips⁴ | Amy Justice²¹ | Janet Tate²¹ | Heiner C. Bucher²² | Matthias Egger²³ | Hansjakob Furrer²⁴ | Jose M. Miro²⁵ | Jordi Casabona²⁶ | Kholoud Porter⁴ | Giota Touloumi²⁷ | Heidi Crane⁸ | Dominique Costagliola²⁸ | Michael Saag²⁹ | Miguel A. Hernán^{1,3,30}

¹Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

²Department of Population Health, School of Medicine, New York University, New York, New York

³Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

⁴University College London, London, UK

⁵APHP Hôpital Avicenne, Bobigny, France

⁶School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama

⁷Inserm, Université Paris Sud, Orsay, France

⁸University of Washington, Seattle, Washington

⁹Centre Hospitalier Universitaire de Bordeaux, Bordeaux, France

¹⁰School of Medicine, The Johns Hopkins University, Baltimore, Maryland

¹¹Academisch Medisch Centrum Geneeskunde, Amsterdam, The Netherlands

¹²Department of Medicine, University of California San Diego Health, San Diego, California

¹³National Center of Epidemiology, Instituto de Salud Carlos III, Madrid, Spain

Decisions about when to start or switch a therapy often depend on the frequency with which individuals are monitored or tested. For example, the optimal time to switch antiretroviral therapy depends on the frequency with which HIV-positive individuals have HIV RNA measured. This paper describes an approach to use observational data for the comparison of joint monitoring and treatment strategies and applies the method to a clinically relevant question in HIV research: when can monitoring frequency be decreased and when should individuals switch from a first-line treatment regimen to a new regimen?

We outline the target trial that would compare the dynamic strategies of interest and then describe how to emulate it using data from HIV-positive individuals included in the HIV-CAUSAL Collaboration and the Centers for AIDS Research Network of Integrated Clinical Systems. When, as in our example, few individuals follow the dynamic strategies of interest over long periods of follow-up, we describe how to leverage an additional assumption: no direct effect of monitoring on the outcome of interest. We compare our results with and without the “no direct effect” assumption. We found little differences on survival and AIDS-free survival between strategies where monitoring frequency was decreased at a CD4 threshold of 350 cells/ μ l compared with 500 cells/ μ l and where treatment was switched at an HIV-RNA threshold of 1000 copies/ml compared with 200 copies/ml. The “no direct effect” assumption resulted in efficiency

¹⁴School of Medicine, University of California, San Francisco, San Francisco, California

¹⁵Germans Trias Hospital, Barcelona, Spain

¹⁶Fenway Health, Boston, Massachusetts

¹⁷Hospital Universitari de Bellvitge, Barcelona, Spain

¹⁸Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

¹⁹Southern Alberta HIV Program, Calgary, Canada

²⁰Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

²¹School of Public Health, Yale University, New Haven, Connecticut

²²Universitätsspital Basel, Basel, Switzerland

²³Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

²⁴Division of Infectious Diseases, Bern University Hospital, University of Bern, Bern, Switzerland

²⁵University of Barcelona, Barcelona, Spain

²⁶Institut Català d'Oncologia, Barcelona, Spain

²⁷Department of Hygiene, Epidemiology and Medical Statistics, Athens University Medical School, Athens, Greece

²⁸University Pierre and Marie Curie, Paris, France

²⁹University of Alabama at Birmingham, Birmingham, Alabama

³⁰Harvard-MIT Division of Health Sciences and Technology, Boston, Massachusetts

Correspondence

Ellen C. Caniglia, Department of Population Health, School of Medicine, New York University, 227 East 30th Street, New York, NY 10016.

Email: caniglia@nyumc.org

Funding information

National Institutes of Health (NIH), Grant/Award Number: R37 AI102634; National Institute of Allergy and Infectious Diseases (NIAID), Grant/Award Number: T32 AI007433; Center for AIDS Research Network of Integrated Clinical Systems, Grant/Award Number: R24 AI067039; National Heart, Lung and Blood Institute; Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Grant/Award Number: personal 80:20 research grant

improvements for the risk difference estimates ranging from an 7- to 53-fold increase in the effective sample size.

KEYWORDS

causal inference, dynamic regime, joint treatment strategies, marginal structural model, no direct effect

1 | INTRODUCTION

Many clinical guidelines recommend starting, stopping, or switching a therapy when a clinical marker crosses a certain threshold. For example, some guidelines recommend that adults initiate statin therapy if LDL cholesterol is greater than 190 mg/dL,¹ and others recommend that HIV-positive individuals switch treatment (antiretroviral therapy [ART]) if HIV RNA is greater than 500 copies/ml.² These guidelines are examples of dynamic strategies because the decision to start or switch therapy depends on an individual's time-varying covariates (LDL cholesterol or HIV RNA).

When randomized trials are not available to inform guidelines, observational data can be used to try to emulate a hypothetical randomized trial, ie, a target trial,³ of dynamic strategies. For example, previous observational analyses have emulated target trials in which HIV-positive individuals were assigned to different treatment initiation and switching strategies.⁴⁻¹⁰ Adjustment for measured time-varying confounders was achieved via inverse probability weighting^{4,6} or the parametric g-formula.⁸⁻¹⁰

However, recommendations about when to start or switch a therapy generally depend on the frequency with which individuals are monitored or tested.¹¹ For example, the optimal time to switch therapy may be the first time HIV RNA crosses above 500 copies/ml if HIV-positive individuals are monitored every 6 months but at a different threshold, ie, lower than 500 copies/ml, if individuals were monitored every 12 months. Therefore, clinical guidelines for starting or switching a treatment based on the results of a test need to specify both the frequency of monitoring/testing and the threshold at which treatment is started or switched.

In this paper, we extend the methodology to emulate a target trial of joint monitoring and treatment strategies using observational data and describe how to leverage an additional assumption: no effect of monitoring on the outcome except through aiding decisions concerning when to switch ART.¹¹ Exploiting this “no direct effect” assumption may drastically decrease the estimates' variance without requiring additional modeling assumptions. Section 2 outlines the key components of the target trial and how to emulate it using observational data from HIV-positive individuals. In Section 3, we describe how to estimate the per-protocol effect, first in the target trial and then using observational data. Section 4 introduces the “no direct effect” assumption and compares the efficiency of the results estimated with and without the additional assumption.

2 | SPECIFICATION AND EMULATION OF THE TARGET TRIAL

Table 1 summarizes the key components of the protocol of the target trial in which participants are randomly assigned to one of four joint monitoring and treatment strategies, based loosely on current clinical guidelines.^{2,15-17} The goal of the trial is to determine the joint strategy with the greatest 5-year survival and AIDS-free survival for HIV-positive individuals who achieve virologic suppression within 12 months of initiating an eligible ART regimen.

Strategy (1): CD4 threshold 350/tight control. In this strategy, CD4 cell count and HIV RNA are monitored every 3-6 months when CD4 is below a threshold of 350 cells/ μ l and every 9-12 months when CD4 is above the threshold, and individuals switch treatment within 3 months of HIV RNA crossing above 200 copies/ml (tight control) and do not switch again.

Strategy (2): CD4 threshold 350/loose control. This strategy is the same as (1) except that the HIV-RNA threshold is 1000 copies/ml (loose control).

Strategy (3): CD4 threshold 500/tight control. This strategy is the same as (1) except that the CD4 cell count threshold is 500 cells/ μ l.

Strategy (4): CD4 threshold 500/loose control. This strategy is the same as (3) except that the HIV-RNA threshold is 1000 copies/ml.

All four strategies further require individuals to be monitored every 3-6 months when HIV RNA > 200 copies/ml or after diagnosis of an AIDS-defining illness (Figure 1).

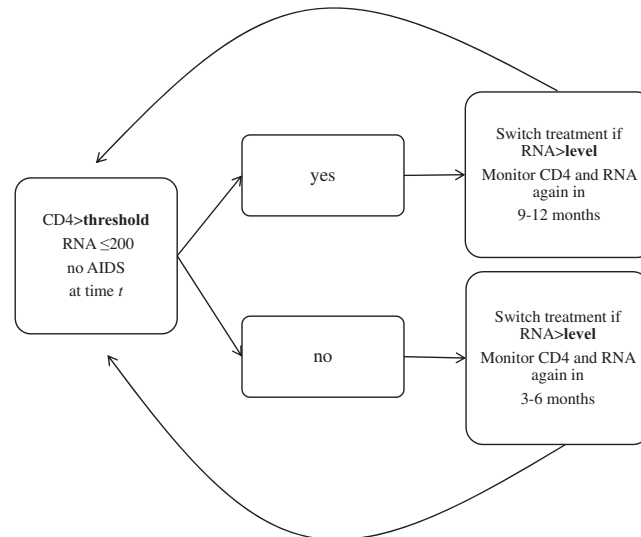
Any treatment change was classified as a nonswitch, ineligible switch, or switch (see Table D1). For example, a change from one protease inhibitor-based regimen to another protease inhibitor-based regimen was not considered a switch, whereas a change from a protease inhibitor-based regimen to an integrase inhibitor-based regimen was considered a switch. Any changes to monotherapy or dual therapy or stopping therapy altogether were considered ineligible switches as they are not consistent with current guidelines.⁷ The switching thresholds were based on current clinical guidelines and to maximize the number of individuals following distinct strategies (in practice, switching also occurs for reasons other than treatment failure).

TABLE 1 Key components of the protocol of the target trial of joint monitoring and treatment strategies

Component	Target Trial	Emulation Using Observational Data
Eligibility criteria	(1) Confirmed virologic suppression (two consecutive HIV RNA \leq 200 copies/ml) within 12 months of initiating an eligible treatment regimen in 2000 or later while remaining on an eligible treatment regimen (Table D1) (2) 18 years or older, no history of AIDS-defining illness, ¹² a CD4 cell count measurement, no pregnancy	Same, except pregnancy information is not available for all individuals.
Joint monitoring and treatment strategies	(1) CD4 threshold 350/tight control: CD4 cell count and HIV RNA are monitored every 3-6 months when CD4 is below a threshold of 350 cells/ μ l and every 9-12 months when CD4 is above the threshold, and individuals switch treatment (Table D1) within 3 months of HIV RNA crossing above 200 copies/ml (2) CD4 threshold 350/loose control: same as (1) except that the HIV-RNA threshold is 1000 copies/ml (3) CD4 threshold 500/tight control: same as (1) except that the CD4 cell count threshold is 500 cells/ μ l (4) CD4 threshold 500/loose control: same as (3) except that the HIV-RNA threshold is 1000 copies/ml All four strategies further require individuals to be monitored once every 3-6 months (\pm 1 month) when HIV RNA $>$ 200 copies/ml or after diagnosis of an AIDS-defining illness. Individuals cannot switch after the initial treatment switch.	Same, except changes to new regimens lasting fewer than 14 days were not considered treatment switches and person-time was assigned to the previous regimen of duration 14 days or longer (sensitivity analyses with periods other than 14 did not materially change the results).
Outcomes	(1) All-cause mortality (2) Combined endpoint of AIDS-defining illness or death	Same. The date of death was identified using a combination of national and local mortality registries and clinical records, as described elsewhere, ^{13,14} and AIDS-defining illnesses were ascertained by the treating physicians.
Follow-up period	Individuals are followed from baseline (confirmed virologic suppression while otherwise eligible, when randomization occurs) until the event of interest, pregnancy, loss to follow-up (12 months after the most recent laboratory measurement), or 5 years of follow-up, whichever occurred earlier.	Same, except that the cohort-specific administrative end of follow-up may be less than 5 years from baseline and pregnancy information is not available for all individuals.
Causal contrast(s) of interest	(1) Intention-to-treat effect (2) Per-protocol effect, ie, the effect that would have been observed if all individuals were monitored and switched treatment as indicated by their randomization arm	Per-protocol effect only. Since all individuals included in our study had data consistent with each of the four joint monitoring and treatment strategies at baseline, an intention-to-treat analysis would compare groups consisting of the same individuals.
Analysis plan	(1) Intention-to-treat analysis (2) Per-protocol analysis (see text)	Same per-protocol analysis, except that individuals were replicated because their baseline data were consistent with more than one strategy.

We consider two clinical endpoints: all-cause mortality and a combined endpoint of AIDS-defining illness or death. Our goal is to estimate the per-protocol effect, that is, the effect that would have been observed if all individuals were monitored and switched treatment as indicated by their assigned strategy.

We emulated this target trial using observational data from two collaborations of prospective studies from high-income countries. The HIV-CAUSAL Collaboration includes prospective cohort studies from Europe and the Americas.¹³ The



Strategy 1: Threshold=350 cells/ μ l; Level=200 copies/ml

Strategy 2: Threshold=350 cells/ μ l; Level=1000 copies/ml

Strategy 3: Threshold=500 cells/ μ l; Level=200 copies/ml

Strategy 4: Threshold=500 cells/ μ l; Level=1000 copies/ml

FIGURE 1 Schematic of four dynamic monitoring and treatment strategies to be followed at each time t of follow-up. By considering the combinations of two CD4 cell count thresholds (350 and 500 cells/ μ l) and two HIV-RNA levels (200 and 1000 copies/ml), four strategies are defined

individual cohort studies are FHDH-ANRSC04 (France), ANRS PRIMO (France), ANRS SEROCO (France), ANRS CO3-Aquitaine (France), UK CHIC (United Kingdom), UK Register of HIV Seroconverts (United Kingdom), ATHENA (the Netherlands), SHCS (Switzerland), PISCIS (Spain), CoRIS/CoRIS-MD (Spain), GEMES (Spain), VACS (United States), AMACS (Greece), IPEC (Brazil), and SAC (Canada). The Center for AIDS Research Network of Integrated Clinical Systems (CNICS) contains clinical data from inpatient and outpatient encounters of HIV-positive individuals at eight US sites: Case Western Reserve University, Fenway Community Health Clinic, Johns Hopkins University, University of Alabama at Birmingham, University of California at San Diego, University of California at San Francisco, University of North Carolina, and University of Washington.¹⁴ All cohorts included in the HIV-CAUSAL and CNICS Collaborations were assembled prospectively and are based on data collected for clinical purposes.

To emulate the target trial, we identified HIV-positive individuals who met the eligibility criteria, classified them into the four strategies (see below), and followed them until the event of interest, pregnancy, loss to follow-up (12 months after the most recent laboratory measurement), 5 years of follow-up, or the cohort-specific administrative end of follow-up, whichever occurred earlier. To allow more individuals to follow each strategy over time, our primary analysis included an additional month before and after each monitoring window (eg, 3-6 \pm 1), so that the monitoring grace period was 5 months. In the next sections, we describe how to estimate the per-protocol effect in the target trial and using observational data.

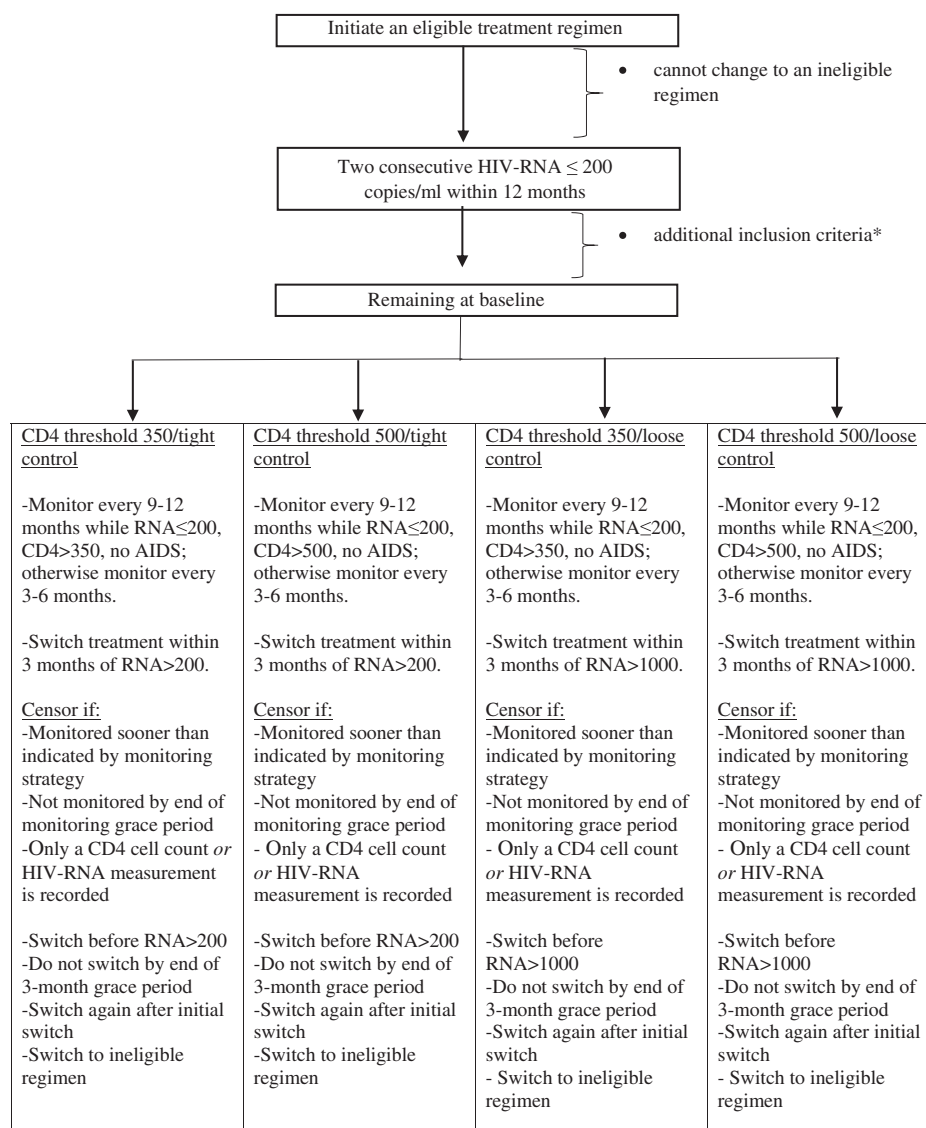
3 | ESTIMATING THE PER-PROTOCOL EFFECT

3.1 | Specification of per-protocol analysis in the target trial

Eligible individuals are randomized at enrollment in the trial to one of the four joint monitoring and treatment strategies. We define the per-protocol effect as the difference in 5-year survival and AIDS-free survival between the four strategies if all individuals have followed their assigned strategies as indicated in the protocol of the target trial. We describe a three-step procedure to estimate the per-protocol effect in the target trial.

First, censor individuals when they deviate from their assigned strategy. Specifically, censor individuals when they are monitored sooner than indicated by their strategy, when they are not monitored soon enough, when they change treatment sooner than indicated by their strategy, when they have not switched at the end of the 3-month treatment-switching grace period, when they switch treatment again after their initial treatment switch, when only a CD4 cell count or HIV-RNA measurement is recorded (uneven monitoring), or when they switch to an ineligible treatment regimen (Figure 2).

Second, fit a discrete-time hazards model and use its predicted values to estimate standardized survival and AIDS-free survival curves.^{6,18,19} For example, the discrete-time hazard at each month t can be estimated by fitting a pooled logistic model such as $\text{logit Pr}(D_{t+1} = 1 | D_t = 0, C_t = 0, X, V) = \theta_0 + h(t) + \theta'_1 V + \theta_2 X_{350\text{-tight control}} + \theta_3 X_{350\text{-loose control}} + \theta_4 X_{500\text{-loose control}} + \theta'_5 X_{350\text{-tight control}} h(t) + \theta'_6 X_{350\text{-loose control}} h(t) + \theta'_7 X_{500\text{-loose control}} h(t)$, where $\theta_0 + h(t)$ is a time-varying intercept, with $h(t)$ defined as a restricted cubic spline for follow-up time (four knots at 1, 6, 12, and 24 months), D_t is an indicator for developing the outcome by month t (1: yes, 0: no), V is a vector of baseline prognostic factors that predict adherence, the X 's are indicators for the corresponding strategy (1: yes, 0: no) with the strategy CD4 threshold 500/tight control as the reference, and $C_t = 0$ is an indicator for the remaining uncensored through t . We use prime notation to



* At baseline, individuals were required to be age 18 years or older, no history of AIDS, a CD4 cell count, no pregnancy (when information was available), and on an eligible treatment regimen.

FIGURE 2 Flow diagram for the eligibility criteria and censoring process (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration, 2000-2015)

denote vectors since V is a vector of baseline covariates and $h(t)$ includes three covariates. We include product terms between $h(t)$ and the X 's so that the estimated hazard ratios can vary over time. The model uses a smooth function to combine information from each strategy, producing more stable estimates than a nonparametric estimate of the survival function (eg, Kaplan-Meier).⁶ This model adjusts for baseline (time-fixed) prognostic factors, but not for post-baseline (time-varying) factors because the inclusion of post-baseline covariates in the model for the outcome may introduce selection bias.²⁰

Third, estimate inverse probability (IP) weights to adjust for post-baseline prognostic factors. To describe the weights, we need to introduce some additional notation. Let $A_t = 2$ indicate that the individual switches to an eligible regimen at time t , $A_t = 1$ indicate that the individual switches to an ineligible regimen during time t , and $A_t = 0$ indicate that the individual does not switch treatment during time t . Let $N_t = 2$ indicate that the individual has both CD4 cell count and HIV-RNA measurements during time t , $N_t = 1$ indicate that the individual has either a CD4 cell count or an HIV-RNA measurement during time t but not both, and $N_t = 0$ indicate that the individual has neither a CD4 cell count nor an HIV-RNA measurement during time t . We use overbars to denote the history of a time-dependent variable: \bar{N}_t is the individual's monitoring history through time t , \bar{A}_t is the individual treatment-switching history through time t , and \bar{L}_t is the individual time-varying covariate history through time t .

The nonstabilized IP treatment-switching weight for each uncensored individual at each time t is

$$W_t^A = \prod_{k=0}^t \frac{1}{\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, \bar{L}_k, \bar{N}_k, \bar{A}_{k-1})}.$$

$\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, \bar{L}_k, \bar{N}_k, \bar{A}_{k-1})$ is equal to $f(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ before and after the treatment-switching grace period, where $f(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ is the conditional probability density function $f_{A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k}(a_k | \bar{a}_{k-1}, d_k = 0, \bar{l}_k, \bar{n}_k)$ with $(a_k | \bar{a}_{k-1}, d_k = 0, \bar{l}_k, \bar{n}_k)$ evaluated at the random argument $(A_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ and is equal to $1 - \Pr(A_k = 1 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ during the treatment-switching grace period.^{6,21}

We estimate $\Pr(A_k = a_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ via a pooled multinomial logistic regression model fit in the original data. Alternatively, we could fit two nested logistic models: a model for ineligible treatment switching and a model for an eligible treatment switch (versus no switch) conditional on not having an ineligible treatment switch (Appendix A).

Similarly, the nonstabilized IP monitoring weights are

$$W_t^N = \prod_{k=0}^t \frac{1}{\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, \bar{L}_{k-1}, \bar{N}_{k-1}, \bar{A}_{k-1})}.$$

$\Pr(C_k = 0 | C_{k-1} = 0, D_k = 0, X = x, \bar{L}_{k-1}, \bar{N}_{k-1}, \bar{A}_{k-1})$ is equal to $f(N_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ before and after the monitoring grace period, where $f(N_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ is the conditional probability mass function $f_{N_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1}}(n_k | \bar{n}_{k-1}, d_k = 0, \bar{l}_{k-1}, \bar{a}_{k-1})$ with $(n_k | \bar{n}_{k-1}, d_k = 0, \bar{l}_{k-1}, \bar{a}_{k-1})$ evaluated at the random argument $(N_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ and is equal to $1 - \Pr(N_k = 1 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ during the monitoring grace period. We estimate $\Pr(N_k = n_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ via a pooled multinomial logistic model fit in the original data. Alternatively, we could fit two nested logistic models: a model for uneven monitoring (only a CD4 cell count or HIV-RNA measurement is recorded) and a model for CD4 and RNA monitoring (versus no monitoring) conditional on not having uneven monitoring (Appendix A).

All models include a time-varying intercept; monitoring history \bar{N}_k summarized by the proportion of months of follow-up from baseline to time k with a CD4 cell count measurement (restricted cubic splines with three knots at 0.2, 0.3, and 0.5), the proportion of months of follow-up from baseline to time k with an RNA measurement (restricted cubic splines with three knots at 0.2, 0.3, and 0.5), months between time k and the last CD4 cell count measurement (restricted cubic splines with three knots at 1, 4, and 7), and months between time k and the last RNA measurement (restricted cubic splines with three knots at 1, 4, and 7); treatment-switching history \bar{A}_k summarized by whether an individual switched treatment from baseline to time k (yes/no) and the number of months between time k and the treatment switch (restricted cubic splines with three knots at 0, 6, and 12); and covariate history \bar{L}_k summarized by V and L_k , which includes the most recently recorded values of CD4 cell count (restricted cubic splines with five knots at 200, 350, 500, 650, and 1000 cells/ μ l), HIV RNA ($\leq 200, 201-999, 1000-9999, \geq 10\,000$ copies/ml), and diagnosis of an AIDS-defining illness (when the outcome was all-cause mortality) at time k .

The nonstabilized IP weight W_t for each person-month is the product of the treatment-switching weight and the monitoring weight $W_t = W_t^A * W_t^N$. We truncate the estimated weights W_t at the 99th percentile to protect against potential model misspecification and near violations of positivity.

Under the assumptions of sequential exchangeability, positivity, and consistency for monitoring and treatment switching conditional on the measured time-fixed and time-varying covariates,¹¹ the parameters of the IP-weighted discrete-time hazards model consistently estimate the parameters of a dynamic marginal structural discrete-time hazards model, ie,

$$\begin{aligned} \text{logit Pr}(D_{t+1}^x = 1 | D_t^x = 0, V) = & \beta_0 + h(t) + \beta_1' V + \beta_2 x_{350\text{-tight control}} + \beta_3 x_{350\text{-loose control}} \\ & + \beta_4 x_{500\text{-loose control}} + \beta_5' x_{350\text{-tight control}} h(t) + \beta_6' x_{350\text{-loose control}} h(t) \\ & + \beta_7' x_{500\text{-loose control}} h(t), \end{aligned}$$

where D_t^x is, for each individual, a (counterfactual) indicator for developing the outcome by month t (1: yes, 0: no) under strategy $X = x$ for monitoring and treatment. The validity of this procedure also requires no misspecification of the models used to estimate the weights and of the discrete-time hazards model.

The nonstabilized IP weights defined above imply strategies under which individuals who were not monitored or did not switch during the corresponding grace period are forced to be monitored or switch treatment at the end of the grace period.⁶ Since this can lead to unstable estimates and may not be consistent with clinical practice, we also consider IP weights that estimate strategies under which individuals are monitored with a uniform probability during the monitoring grace period and switch treatment with a uniform probability during the treatment-switching grace period.⁶ Appendix A shows the contributions to the monitoring and treatment-switching weights at different time points for both the nonstabilized and uniform IP weights. We used nonparametric bootstrapping with 500 samples to compute 95% confidence intervals around our estimates.

Even if the data on monitoring and treatment strategy assignment had been inadvertently erased from the analysis file in the target trial, one can still construct consistent estimators of the per-protocol effects under exchangeability, consistency, and positivity. In fact, Robins shows that the most efficient estimator of the per-protocol effect ignores data on assignment even when available.²² The following section describes estimators that ignore data on assignment to emulate the per-protocol analysis of a target trial using observational data.

3.2 | Emulation of the per-protocol analysis using observational data

The observational per-protocol analysis is the same as described above for the target trial except that data on monitoring and treatment strategy assignment are absent as no such assignment occurred. In fact, an individual's data at baseline may be consistent with more than one of the four strategies of interest. As previously described in detail,^{6,7,23} we solved this problem by creating an expanded data set with four exact replicates of each individual (one per strategy), each following one of the strategies of interest. We censored each replicate, as described above, when the individual's data were no longer consistent with the strategy assigned to the replicate. Appendix B describes data for three hypothetical individuals and the strategies they followed over 24 months of follow-up time. To estimate the per-protocol effect, we used the same IP-weighted pooled logistic model described for the target trial, except that we fitted the model to the expanded data set. The models for the weights were fit in the original unexpanded data set.

Data from each participating study can be requested from the sponsoring institution in accordance with the applicable laws or regulations in each country. SAS programs are available at <https://www.hsph.harvard.edu/causal/hiv>.

3.3 | Data analysis results

In our analysis, 41 724 individuals met the eligibility criteria and were included. After 2 years of follow-up, 1006 individuals were following the CD4 threshold 350/tight control strategy, 2634 individuals were following the CD4 threshold 500/tight control strategy, 1050 individuals were following the CD4 threshold 350/loose control strategy, and 2741 individuals were following the CD4 threshold 500/loose control strategy. After 5 years of follow-up, these numbers were 45, 152, 47, and 164, respectively (Figure 3). Over the 5-year follow-up, there were 455 deaths and 1151 cases of AIDS-defining illness or death (Table 2). The median (interquartile range) time from baseline to death among individuals who died was 5 (2, 10) months. Figure 4 plots the estimated 5-year survival and 5-year AIDS-free survival. Compared with the CD4 threshold 500/tight control strategy, the 5-year survival difference was 0.01 (−0.01, 0.04) for the CD4 threshold 350/tight control strategy, 0.01 (−0.02, 0.03) for the CD4 threshold 350/loose control strategy, and 0.00 (−0.01, 0.01) for the CD4 threshold

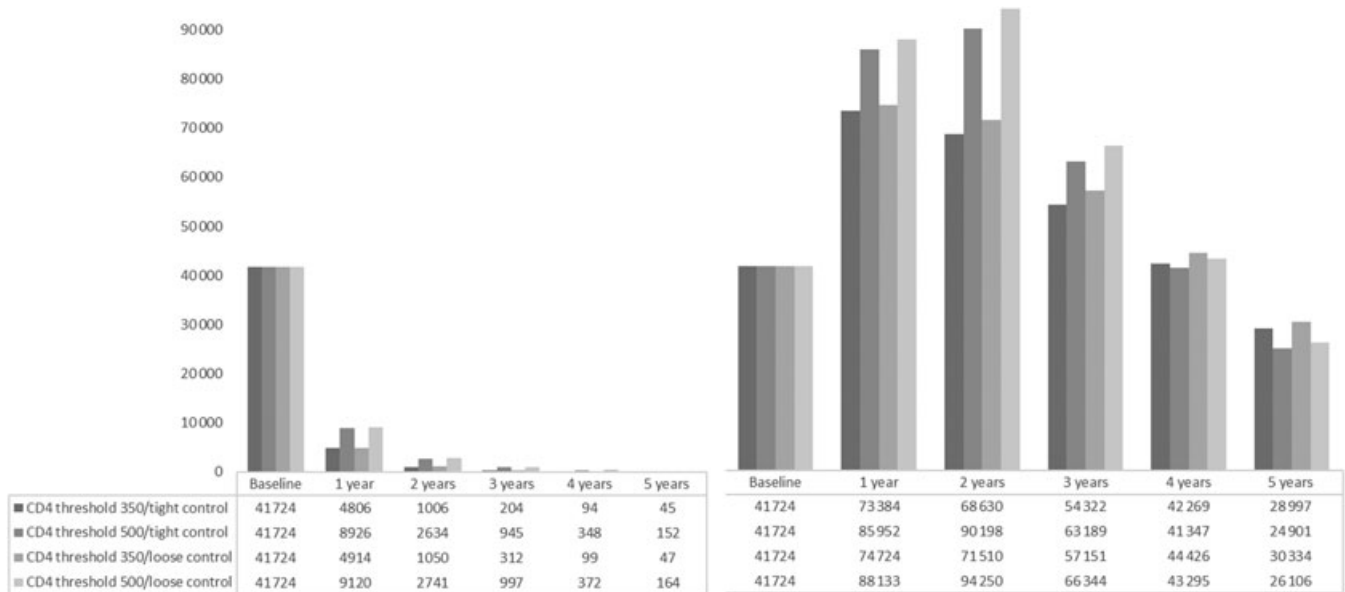


FIGURE 3 Number of replicates following each monitoring and treatment strategy over follow-up time, without (left) and with (right) the “no direct effect” assumption (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration, 2000-2015)

500/loose control. The 5-year AIDS-free survival difference was 0.00 (−0.02, 0.03) for the CD4 threshold 350/tight control strategy, 0.00 (−0.02, 0.02) for the CD4 threshold 350/loose control strategy, and 0.00 (0.00, 0.01) for the CD4 threshold 500/loose control strategy, compared with the CD4 threshold 500/tight control strategy (Table 2).

4 | ESTIMATING THE PER-PROTOCOL EFFECT WITH A “NO DIRECT EFFECT” ASSUMPTION

In many settings, it can be argued that monitoring has no direct effect on the outcome except through aiding decisions regarding switching treatment. For example, in our study, we can assume that having a laboratory measurement can only affect the risk of AIDS or death by triggering treatment changes. More precisely, consider a target trial with two arms in which both arms are assigned the same static treatment strategy but different monitoring strategies. Then, the “no direct effect” assumption says that the two arms will have the same survival curves. By emulating this type of target trial with observational data, it is, in principle, possible to test the “no direct effect” assumption under sequential exchangeability, positivity, and consistency.²⁴

Under the assumption of no direct effect of monitoring, measurements inconsistent with the monitoring strategy have no effect on survival. Therefore, CD4 and HIV-RNA measurements at months not consistent with the monitoring strategy can be ignored, and replicates need not be censored at those times. Because these individuals are not censored, we will have more individuals continuing to follow the strategies of interest at any given time and, therefore, more precise effect estimates. Under the “no direct effect” assumption, individuals can no longer be censored for being monitored too frequently but can still be censored for not being monitored frequently enough. In fact, it is possible for the counterfactual survival curve under a particular monitoring and treatment strategy to be identified under the “no direct effect” assumption but to be unidentified (due to lack of positivity) without the assumption.¹¹ For example, if all subjects in the observational data were monitored every month, it would not be possible to identify the effect of any less frequent monitoring strategy without the “no direct effect” assumption.

To implement the modified per-protocol analysis that incorporates the “no direct effect” assumption, we construct a “no direct effect” version of the data using the following algorithm, formalized previously by Robins et al. Specifically, the algorithm described next implements Equation 31 on page 4711, extended to allow for random treatment regimes as well as random monitoring regimes¹¹: (i) recode the monitoring indicator at each month t to 0 when an individual is monitored at a time t inconsistent with his or her randomization arm's monitoring strategy (including when only CD4 cell count or HIV RNA is measured), (ii) delete the CD4 cell count and HIV RNA recorded at month t whenever the monitoring indicator at time t has been recoded to 0, and (iii) carry forward the previous CD4 cell count and HIV RNA until the next

TABLE 2 Estimated 5-year survival and AIDS-free survival* under each monitoring and treatment strategy (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration, 2000-2015)

“No Direct Effect” Assumption	Strategy	No. of Deaths	5-Year Survival (95% CI)	5-Year Survival Difference (95% CI)	No. of Deaths/AIDS	5-Year AIDS-Free Survival (95% CI)	5-Year AIDS-Free Survival Difference (95% CI)
No	CD4 threshold 350/tight control	101	0.99 (0.99, 1.00)	0.01 (-0.01, 0.04)	262	0.98 (0.97, 0.99)	0.01 (-0.02, 0.03)
	CD4 threshold 500/tight control	123	0.98 (0.95, 1.00)	0 (reference)	309	0.98 (0.95, 1.00)	0 (reference)
	CD4 threshold 350/loose control	105	0.99 (0.98, 1.00)	0.01 (-0.02, 0.03)	267	0.98 (0.97, 0.99)	0.00 (-0.02, 0.02)
	CD4 threshold 500/loose control	126	0.98 (0.96, 1.00)	0.00 (-0.01, 0.01)	313	0.98 (0.96, 1.00)	0.00 (0.00, 0.01)
Yes	CD4 threshold 350/tight control	1375	0.98 (0.97, 0.98)	0.00 (0.00, 0.00)	2224	0.96 (0.96, 0.97)	0.00 (0.00, 0.01)
	CD4 threshold 500/tight control	1478	0.98 (0.97, 0.98)	0 (reference)	2514	0.96 (0.95, 0.97)	0 (reference)
	CD4 threshold 350/loose control	1455	0.98 (0.97, 0.98)	0.00 (0.00, 0.00)	2353	0.96 (0.95, 0.97)	0.00 (-0.01, 0.01)
	CD4 threshold 500/loose control	1578	0.98 (0.97, 0.99)	0.00 (0.00, 0.00)	2667	0.96 (0.95, 0.97)	0.00 (-0.01, 0.00)

These estimates were standardized by the baseline covariates: sex, CD4 cell count (≤ 200 , 201-350, ≥ 351 -500, ≥ 501 cells/ μ l), years since HIV diagnosis (< 1 , 1 to 4, ≥ 5 years, unknown), race (white, black, other, or unknown), geographic origin (N. America/W. Europe, Sub-Saharan Africa, other, unknown), acquisition group (heterosexual, homosexual, or bisexual; injection drug use; other; or unknown), calendar year (restricted cubic splines with three knots at 2001, 2007, and 2011), age (restricted cubic splines with three knots at 25, 39, and 60), cohort, and months from treatment initiation to virologic suppression (2-4, 5-8, ≥ 9).

* All estimates rounded to the nearest hundredth

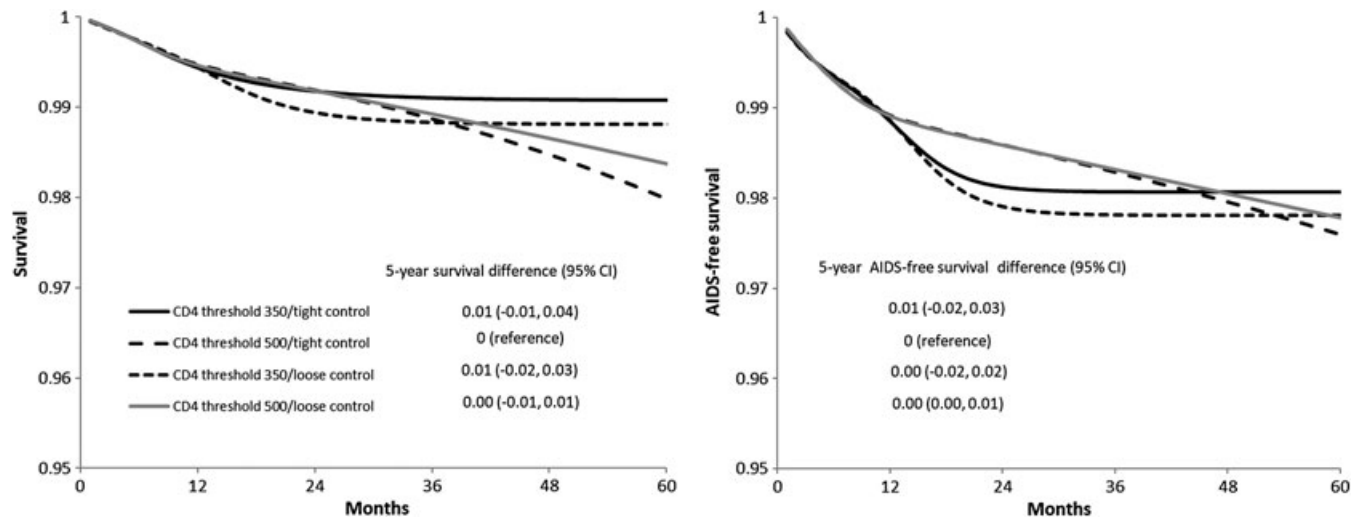


FIGURE 4 Survival (left) and AIDS-free survival (right) under each monitoring and treatment strategy (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration, 2000-2015). The curves are standardized by the baseline covariates: sex, CD4 cell count (≤ 200 , 201-350, 351-500, ≥ 501 cells/ μl), years since HIV diagnosis (<1 , 1 to 4, ≥ 5 years, unknown), race (white, black, other, or unknown), geographic origin (N. America/W. Europe, Sub-Saharan Africa, other, unknown), acquisition group (heterosexual, homosexual, or bisexual; injection drug use; other; or unknown), calendar year (restricted cubic splines with three knots at 2001, 2007, and 2011), age (restricted cubic splines with three knots at 25, 39, and 60), cohort, and months from treatment initiation to virologic suppression (2-4, 5-8, ≥ 9)

time that individual is monitored (Appendix C). In this “no direct effect” data set, we then proceed to censor individuals when they deviate from their assigned strategy and to estimate the survival and AIDS-free survival for each strategy, as described above. To emulate the target trial using the observational data, we modify the expanded data set in the same way to construct a “no direct effect” data set.

The algorithm described above can be extended to strategies with grace periods, which allows even more individuals to follow the strategies of interest over a long period of time. The algorithm can be adapted to handle more than one monitoring time during a grace period. To extend the algorithm, replicates monitored during the monitoring grace period are further replicated, as a way of simulating monitoring trajectories where replicates are monitored at different times during the monitoring grace period.¹¹ Specifically, a replicate monitored during the grace period is cloned at the point of their first measurement in the grace period into two new replicates 1_t and 2_t , where t denotes the time in the grace period the cloning occurred. For replicate 1_t , the new CD4 cell count and HIV-RNA measurements are revealed (recorded for data analysis), and the replicate exits the monitoring grace period in the usual way. For replicate 2_t , the new CD4 cell count and HIV-RNA measurements are ignored, ie, the monitoring indicator is recoded to 0, the CD4 and HIV-RNA measurements are deleted, the previous CD4 and RNA measurements are carried forward, and replicate 2_t moves to the next month of the grace period. Replicates are only censored if they are not monitored at least as frequently as required by the strategy. As an example, replicate 2_t will be censored if she receives no further monitoring during the grace period. Appendix C describes the “no direct effect” data set for one hypothetical individual, first under strategies that require individuals to be monitored exactly every 6 months if their CD4 cell count falls below the strategy’s threshold and exactly every 12 months otherwise and, second, under strategies with grace periods.

Under the “no direct effect” assumption, the monitoring weights W_t^N are equal to $\prod_{k: N^*(k)=2}^t \frac{1}{f(N_k | \bar{N}_{k-1}, D_k=0, \bar{I}_{k-1}, \bar{A}_{k-1})}$, where $N^*(k) = 2$ denotes times k when a replicate’s CD4 and HIV-RNA values are revealed. The factors in the denominator of W_t^N are 1 at all times when a replicate’s CD4 and HIV-RNA values are not revealed (Appendix A).¹¹ The treatment weights remain as before. Note that different replicates from a single person will have different weights. Below, we compare the efficiency of the estimates with and without the “no direct effect” assumption when using uniform IP weights (Table 2).

4.1 | Data analysis results

After 5 years of follow-up, 28 997 individuals were following the CD4 threshold 350/tight control strategy, 24 901 individuals were following the CD4 threshold 500/tight control strategy, 30 334 individuals were following the CD4 threshold

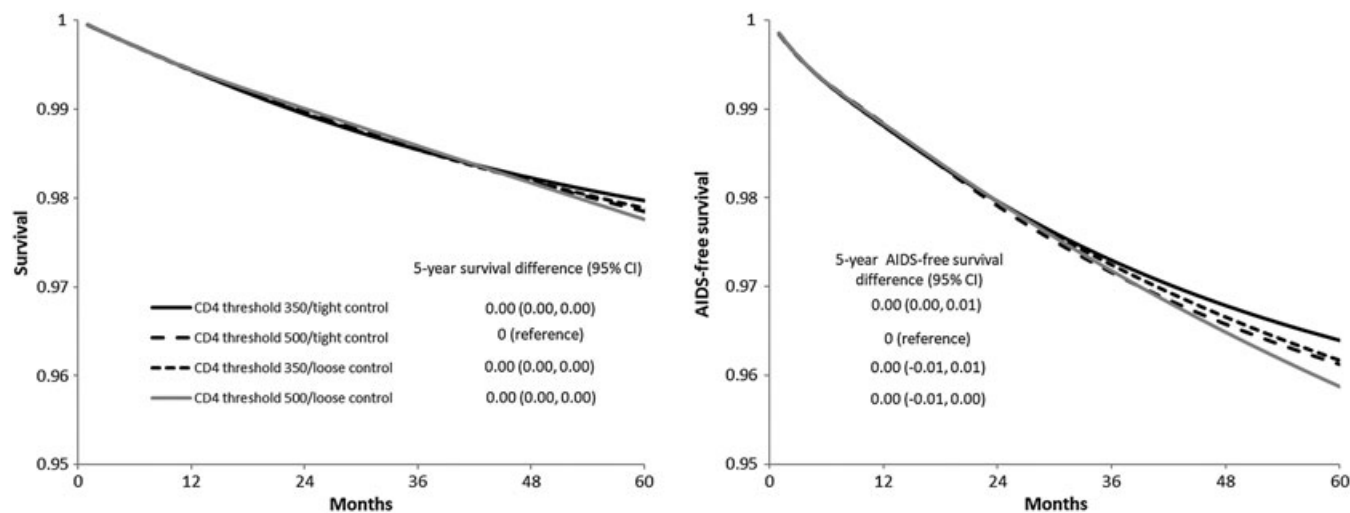


FIGURE 5 Survival (left) and AIDS-free survival (right) under each monitoring and treatment strategy with the “no direct effect” assumption (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration, 2000-2015). The curves are standardized by the baseline covariates: sex, CD4 cell count (≤ 200 , 201-350, 351-500, ≥ 501 cells/ μl), years since HIV diagnosis (<1 , 1 to 4, ≥ 5 years, unknown), race (white, black, other, or unknown), geographic origin (N. America/W. Europe, Sub-Saharan Africa, other, unknown), acquisition group (heterosexual, homosexual, or bisexual; injection drug use; other; or unknown), calendar year (restricted cubic splines with three knots at 2001, 2007, and 2011), age (restricted cubic splines with three knots at 25, 39, and 60), cohort, and months from treatment initiation to virologic suppression (2-4, 5-8, ≥ 9)

350/loose control strategy, and 26 106 individuals were following the CD4 threshold 500/loose control strategy (Figure 3). Over the 5-year follow-up, there were 5886 deaths and 9758 cases of AIDS-defining illness or death (Table 2). The median (interquartile range) time to death among individuals who died was 14 (6, 28) months. Figure 5 plots the estimated 5-year survival and 5-year AIDS-free survival. Compared with the CD4 threshold 500/tight control strategy, the 5-year survival difference was 0.00 (0.00, 0.00) for the CD4 threshold 350/tight control strategy, 0.00 (0.00, 0.00) for the CD4 threshold 350/loose control strategy, and 0.00 (0.00, 0.00) for the CD4 threshold 500/loose control strategy. The 5-year AIDS-free survival difference was 0.00 (0.00, 0.01) for the CD4 threshold 350/tight control strategy, 0.00 (-0.01, 0.01) for the CD4 threshold 350/loose control strategy, and 0.00 (-0.01, 0.00) for the CD4 threshold 500/loose control strategy, compared with the CD4 threshold 500/tight control strategy (Table 2).

These estimates under the “no direct effect” assumption were more precise than those in the previous section. For example, the standard errors for the 5-year survival difference estimates ranged from 0.0046 to 0.0131 without the “no direct effect” assumption and from 0.0017 to 0.0019 with the “no direct effect” assumption, implying the ratio of the effective sample size under the “no direct effect” assumption to that without the assumption ranging from 7 to 53.

5 | DISCUSSION

This paper describes the use of observational data to emulate a target trial of joint monitoring and treatment strategies. We applied the method to strategies for the management of HIV-positive individuals and found no differences on survival and AIDS-free survival between strategies with monitoring at a CD4 threshold of 350 cells/ μl compared with 500 cells/ μl and with treatment switching at an HIV-RNA threshold of 1000 copies/ml compared with 200 copies/ml.

Like for any other observational study, the validity of our estimates relies on the untestable assumption that the measured covariates were sufficient to adjust for confounding and selection bias. In our analysis, we adjusted for several joint predictors of monitoring and the outcome as well as of treatment switching and the outcome. If physicians monitor individuals perceived to have lower adherence with greater frequency or make different decisions about treatment switching based on perceived adherence, which we did not directly adjust for, this assumption may not hold. However, we were able to adjust for several potential proxies of adherence, such as HIV RNA. In addition, our results could be biased if both the monitoring frequency and survival vary by site. However, the monitoring frequency was similar between the countries included in our analysis (data not shown).

One challenge in estimating the effect of complex treatment strategies using observational data is that few individuals may have data consistent with the strategies of interest over an extended period of follow-up. We described how to ameliorate this by incorporating the often plausible assumption that monitoring has no direct effect on the outcome, except through aiding decisions regarding when to switch treatment. The “no direct effect” assumption is advantageous because it increases the number of individuals whose data are consistent with the strategies of interest and does not require additional modeling assumptions. In fact, our survival and AIDS-free survival estimates were similar with and without the “no direct effect” assumption, but the estimates under the “no direct effect” assumption were more precise.

The “no direct effect” assumption may not be met if contact with health facilities improves outcomes through interventions that are either unrecorded in the database or not included as additional types of “treatments” in addition to treatment switching. For example, if HIV care is integrated with other services like screening for cancer or cardiovascular disease, or if physicians use the results of a CD4 cell test to initiate treatments other than that under consideration (eg, prophylaxis for opportunistic infections), the assumption may not hold (unless the strategy includes screening and/or prophylaxis as additional “treatments”). However, under the strategies of interest, individuals with low CD4 cell counts were monitored frequently (every 2-7 months), and so, few CD4 tests were deleted for these individuals when creating the modified data set (in the analysis without the “no direct effect” assumption, only 4% of replicates censored for being monitored too frequently had a CD4 cell count ≤ 200 cells/ μ l at the time they were censored).

The monitoring and treatment strategies in our primary analysis did not consider treatment switches after the initial treatment switch. While strategies that allow arbitrary treatment switches would be more realistic,²⁵ they are also computationally harder to implement under the “no direct effect” assumption. In a sensitivity analysis without the “no direct effect” assumption, we considered modified strategies that allowed treatment switches both before virologic failure and after the initial treatment switch. This analysis yielded similar results (5-year survival difference compared with the CD4 threshold 500/tight control strategy: 0.01 (−0.01, 0.02) for the CD4 threshold 350/tight control strategy, 0.01 (−0.01, 0.04) for the CD4 threshold 350/loose control strategy, and 0.01 (0.00, 0.01) for the CD4 threshold 500/loose control strategy).

The methods described in this paper can be extended to other joint monitoring and treatment strategies analyzed in health research. The “no direct effect” assumption can be a useful tool in settings where individuals are monitored or tested often, but inferences about less frequent monitoring or testing are desired. For example, this approach may be particularly useful when data from a high-resource population with frequent testing are available but researchers want to apply the estimates to a low-resource population with infrequent testing.¹¹

DISCLOSURES

The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

ORCID

Ellen C. Caniglia  <https://orcid.org/0000-0002-7567-8428>

REFERENCES

1. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. *Circulation*. 2014;129(25, suppl 2):S1-S45.
2. European AIDS Clinical Society. EACS Guidelines Version 8.0. 2015.
3. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764.
4. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol*. 2006;98(3):237-242.
5. Cain LE, Logan R, Robins JM, et al. When to initiate combined antiretroviral therapy to reduce mortality and AIDS-defining illness in HIV-infected persons in developed countries: an observational study. *Ann Intern Med*. 2011;154(8):509-515.
6. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernán MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat*. 2010;6(2):18

7. Cain LE, Saag MS, Petersen M, et al. Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy. *Int J Epidemiol*. 2015;45:2038-2049.
8. Lodi S, Costagliola D, Sabin C, et al. Effect of immediate initiation of antiretroviral treatment in HIV-positive individuals aged 50 years or older. *J Acquir Immune Defic Syndr*. 2017;76(3):311-318.
9. Lodi S, Gunthard HF, Dunn D, et al. Effect of immediate initiation of antiretroviral treatment on the risk of acquired HIV drug resistance. *AIDS*. 2018;32(3):327-335.
10. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Stat Biosci*. 2011;3(1):119-143.
11. Robins J, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. *Stat Med*. 2008;27(23):4678-4721.
12. Centers for Disease Control and Prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morb Mortal Wkly Rep*. 1992;41:1-9.
13. Ray M, Logan R, Sterne JA, et al. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. *AIDS*. 2010;24(1):123-137.
14. Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of integrated clinical systems. *Int J Epidemiol*. 2008;37(5):948-955.
15. World Health Organization. *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach*. Geneva, Switzerland: WHO Press; 2016.
16. Asboe D, Aitken C, Boffito M, et al. British HIV Association guidelines for the routine investigation and monitoring of adult HIV-1-infected individuals 2011. *HIV Med*. 2012;13(1):1-44.
17. United States Department of Health and Human Services. *Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents*. 2016.
18. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21(1):13-15.
19. Hernán MA, Robins JM. *Causal Inference*. Boca Raton, FL: Chapman & Hall/CRC; 2019:73-86. Forthcoming.
20. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-625.
21. Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Statist Med*. 2002;21(12):1689-1709.
22. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect. *Math Model*. 1986;7:1393-1512.
23. Caniglia EC, Sabin C, Robins JM, Logan R, Cain LE, Al E. When to monitor CD4 cell count and HIV RNA to reduce mortality and AIDS-defining illness in virologically suppressed HIV-positive persons on antiretroviral therapy in high-income countries: a prospective observational study. *J Acquir Immune Defic Syndr*. 2016;72:214.
24. Robins JM. Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models. In: Glymour C, Cooper G, eds. *Computation, Causation, and Discovery*. Menlo Park, CA/Cambridge, MA: AAAI Press/The MIT Press; 1999:349-405.
25. Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med*. 2017;377(14):1391-1398.

How to cite this article: Caniglia EC, Robins JM, Cain LE, et al. Emulating a trial of joint dynamic strategies: An application to monitoring and treatment of HIV-positive individuals. *Statistics in Medicine*. 2019;1-19. <https://doi.org/10.1002/sim.8120>

APPENDIX A

ESTIMATION OF IP WEIGHTS

We estimated $\Pr(A_k = a_k | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ and $\Pr(N_k = n_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ via pooled multinomial logistic regression models. The corresponding contributions to the IP weights are outlined in Table A1.1 and A1.2. Equivalently, we could have fit two sets of nested logistic models: a model for ineligible treatment switching $\Pr(A_k = 1 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$ and a model for an eligible treatment switch (versus no switch) conditional on not having an ineligible treatment switch $\Pr(A_k = 2 | A_k \neq 1, \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$. To estimate $\Pr(N_k = n_k | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$, we fit a model for uneven monitoring $\Pr(N_k = 1 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$ and a model for CD4 and RNA monitoring (versus no monitoring) conditional on not having uneven monitoring $\Pr(N_k = 2 | N_k \neq 1 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$. The corresponding contributions to the IP weights are outlined in Table A2.1 and A2.2 The contributions to the IP weights outlined in

TABLE A1.1 Contribution to monitoring weights at different time points

Time Point	Monitor	Type of Weight		
		Nonstabilized	Uniform	Uniform w/NDE
Before grace period	None	$\frac{1}{\Pr(N0)}$	$\frac{1}{\Pr(N0)}$	$\frac{1}{1}$
	CD4 + RNA	0 (censored)	0 (censored)	
During grace period $0 \leq j < m$	CD4 + RNA	$\frac{1}{1-\Pr(N1)}$	$\frac{1/(m+1-j)}{\Pr(N2)}$	Replicate 1 _t : $\frac{1/(m+1-j)}{\Pr(N2)}$ Replicate 2 _t : $\frac{1-\left[\frac{1}{m+1-j}\right]}{1}$
	None	$\frac{1}{1-\Pr(N1)}$	$\frac{1-\left[\frac{1}{m+1-j}\right]}{\Pr(N0)}$	$\frac{1-\left[\frac{1}{m+1-j}\right]}{1}$
End of grace period $j = m$	CD4 + RNA	$\frac{1}{\Pr(N2)}$	$\frac{1/(m+1-j)}{\Pr(N2)}$	$\frac{1/(m+1-j)}{\Pr(N2)}$

m , length of monitoring grace period in months; j , position in monitoring grace period (month)

N0: neither CD4 or RNA measured; N1: either CD4 or RNA measured but not both; N2: CD4 and RNA measured

$\Pr(N0)$, estimated by $[\Pr(N_k = 0 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})]$

$\Pr(N1)$, estimated by $[\Pr(N_k = 1 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})]$

$\Pr(N2)$, estimated by $[\Pr(N_k = 2 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})]$

TABLE A1.2 Contribution to treatment switching weights at different time points

Time Point	Treatment Switch	Type of Weight		
		Nonstabilized	Uniform	Uniform w/NDE
Before grace period*	Do not switch	$\frac{1}{\Pr(A0)}$	$\frac{1}{\Pr(A0)}$	$\frac{1}{\Pr(A0)}$
During grace period $0 \leq k < n$	Do not switch	$\frac{1}{1-\Pr(A1)}$	$\frac{1-\left[\frac{1}{n+1-k}\right]}{\Pr(A0)}$	$\frac{1-\left[\frac{1}{n+1-k}\right]}{\Pr(A0)}$
	Switch to eligible	$\frac{1}{1-\Pr(A1)}$	$\frac{1/(n+1-k)}{\Pr(A2)}$	$\frac{1/(n+1-k)}{\Pr(A2)}$
End of grace period $k = n$	Switch to eligible	$\frac{1}{\Pr(A2)}$	$\frac{1/(n+1-k)}{\Pr(A2)}$	$\frac{1/(n+1-k)}{\Pr(A2)}$
After grace period	Do not switch	$\frac{1}{\Pr(A0)}$	$\frac{1}{\Pr(A0)}$	$\frac{1}{\Pr(A0)}$

n , length of treatment-switching grace period in months

k , position in treatment-switching grace period (month)

A0: no switch; A1: switch to ineligible regimen; A2: switch to eligible regimen

$\Pr(A0)$, estimated by $[\Pr(A_k = 0 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)]$

$\Pr(A1)$, estimated by $[\Pr(A_k = 1 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)]$

$\Pr(A2)$, estimated by $[\Pr(A_k = 2 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)]$

*Corresponds to time before the first RNA > 200 copies/ml, as well as the time in the loose-control strategies where RNA > 200 but has not yet increased above 1000 copies/ml.

TABLE A2.1 Contribution to monitoring weights at different time points using nested models

Time Point	Monitor	Type of Weight		
		Nonstabilized	Uniform	Uniform w/NDE
Before grace period	None	$\frac{1}{\Pr(N0 N1^c) \Pr(N1^c)}$	$\frac{1}{\Pr(N0 N1^c) \Pr(N1^c)}$	$\frac{1}{1}$
	CD4 + RNA	0 (censored)	0 (censored)	
During grace period $0 \leq j < m$	CD4 + RNA	$\frac{\Pr(N2 N1^c)}{\Pr(N2 N1^c) \Pr(N1^c)}$	$\frac{1/(m+1-j)}{\Pr(N2 N1^c) \Pr(N1^c)}$	Replicate 1 _t : $\frac{1/(m+1-j)}{\Pr(N2 N1^c) \Pr(N1^c)}$ Replicate 2 _t : $\frac{1-\left[\frac{1}{m+1-j}\right]}{1}$
	None	$\frac{\Pr(N0 N1^c)}{\Pr(N0 N1^c) \Pr(N1^c)}$	$\frac{1-\left[\frac{1}{m+1-j}\right]}{\Pr(N0 N1^c) \Pr(N1^c)}$	$\frac{1-\left[\frac{1}{m+1-j}\right]}{1}$
End of grace period $j = m$	CD4 + RNA	$\frac{1}{\Pr(N2 N1^c) \Pr(N1^c)}$	$\frac{1/(m+1-j)}{\Pr(N2 N1^c) \Pr(N1^c)}$	$\frac{1/(m+1-j)}{\Pr(N2 N1^c) \Pr(N1^c)}$

m , length of monitoring grace period in months; j , position in monitoring grace period (month)

N0: neither CD4 or RNA measured; N1: either CD4 or RNA measured but not both; N2: CD4 and RNA measured

$\Pr(N1^c)$, estimated by $1 - [\Pr(N_k = 1 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})]$

$\Pr(N2|N1^c)$, estimated by $\Pr(N_k = 2 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})$

$\Pr(N0|N1^c)$, equal to $1 - \Pr(N2|N1^c)$, estimated by $1 - [\Pr(N_k = 2 | \bar{N}_{k-1}, D_k = 0, \bar{L}_{k-1}, \bar{A}_{k-1})]$

Tables A1 and A2 are mathematically equivalent.

TABLE A2.2 Contribution to treatment switching weights at different time points using nested models

Time Point	Treatment Switch	Type of Weight		
		Nonstabilized	Uniform	Uniform w/NDE
Before grace period*	Do not switch	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$
During grace period $0 \leq k < n$	Do not switch	$\frac{\Pr(A0 A1^c)}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1 - \left[\frac{1}{n+1-k}\right]}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1 - \left[\frac{1}{n+1-k}\right]}{\Pr(A0 A1^c) \Pr(A1^c)}$
	Switch to eligible	$\frac{\Pr(A2 A1^c)}{\Pr(A2 A1^c) \Pr(A1^c)}$	$\frac{1/(n+1-k)}{\Pr(A2 A1^c) \Pr(A1^c)}$	$\frac{1/(n+1-k)}{\Pr(A2 A1^c) \Pr(A1^c)}$
End of grace period $k = n$	Switch to eligible	$\frac{1}{\Pr(A2 A1^c) \Pr(A1^c)}$	$\frac{1/(n+1-k)}{\Pr(A2 A1^c) \Pr(A1^c)}$	$\frac{1/(n+1-k)}{\Pr(A2 A1^c) \Pr(A1^c)}$
After grace period	Do not switch	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$	$\frac{1}{\Pr(A0 A1^c) \Pr(A1^c)}$

n , length of treatment-switching grace period in months

k , position in treatment-switching grace period (month)

A0: no switch; A1: switch to ineligible regimen; A2: switch to eligible regimen

$\Pr(A1^c)$, estimated by $1 - [\Pr(A_k = 1 | \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)]$

$\Pr(A2|A1^c)$, estimated by $\Pr(A_k = 2 | A_k \neq 1, \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)$

$\Pr(A0|A1^c)$, equal to $1 - \Pr(A2|A1^c)$, estimated by $1 - [\Pr(A_k = 2 | A_k \neq 1, \bar{A}_{k-1}, D_k = 0, \bar{L}_k, \bar{N}_k)]$

*Corresponds to time before the first RNA > 200 copies/ml, as well as the time in the loose-control strategies where RNA > 200 but has not yet increased above 1000 copies/ml.

APPENDIX B

EXAMPLE DATA WITHOUT THE “NO DIRECT EFFECT” ASSUMPTION

Table B1 shows data for three hypothetical individuals following each of the four joint monitoring and treatment strategies over 24 months of follow-up. All three individuals had a baseline CD4 cell count of 300 cells/ μ l, an HIV RNA of

TABLE B1 Three hypothetical individuals following each of the four joint monitoring and treatment strategies over 24 months of follow-up

Individual	Time (Months)	Monitor (1: Yes, 0: No)	CD4 Cell Count	HIV RNA	Switch Treatment(1: Yes, 0: No)	Strategies Following*
1	0	1	300	50	0	350-T, 500-T, 350-L, 500-L
1	3	0	300	50	0	350-T, 500-T, 350-L, 500-L
1	6	1	380	50	0	350-T, 500-T, 350-L, 500-L
1	9	0	380	50	0	350-T, 500-T, 350-L, 500-L
1	12	1	400	50	0	500-T, 500-L
1	15	0	400	50	0	500-T, 500-L
1	18	1	430	50	0	500-T, 500-L
1	21	0	430	50	0	500-T, 500-L
1	24	1	500	50	0	500-T, 500-L
2	0	1	300	50	0	350-T, 500-T, 350-L, 500-L
2	3	0	300	50	0	350-T, 500-T, 350-L, 500-L
2	6	1	380	500	1	350-T, 500-T
2	9	0	380	500	0	350-T, 500-T
2	12	1	400	500	0	350-T, 500-T
2	15	0	400	500	0	350-T, 500-T
2	18	1	430	100	0	350-T, 500-T
2	21	0	430	100	0	350-T, 500-T
2	24	1	500	100	0	500-T
3	0	1	300	50	0	350-T, 500-T, 350-L, 500-L
3	3	0	300	50	0	350-T, 500-T, 350-L, 500-L
3	6	1	380	500	0	350-T, 500-T, 350-L, 500-L
3	9	0	380	500	0	350-L, 500-L
3	12	1	400	1200	1	350-L, 500-L
3	15	0	400	1200	0	350-L, 500-L
3	18	1	430	500	0	350-L, 500-L
3	21	0	430	500	0	350-L, 500-L
3	24	1	500	150	0	350-L, 500-L

*350-T: threshold 350-tight control; 500-T: threshold 500-tight control; 350-L: threshold 350-loose control; 500-L: threshold 500-loose control

50 copies/ml, and were monitored every 6 months. Their CD4 cell count increased from 300 to 380, then to 400, then to 430, and, finally, to 500. The difference between the three individuals was that individual 1 never experienced virologic failure (HIV RNA > 200 copies/ml), individual 2 experienced virologic failure at month 6 and switched treatment immediately, and individual 3 experienced virologic failure at month 6 but did not switch treatment until month 12. Individual 2 achieved virologic suppression (HIV RNA \leq 200 copies/ml) again at month 18, and individual 3 achieved virologic suppression at month 24. Individual 1 was censored from the threshold 350 strategies at month 12 because she was monitored at an interval shorter than 9-12 months while her CD4 cell count was above the 350 threshold. Individual 2 was censored from the loose-control strategies at month 6 when she switched treatment before her HIV RNA crosses above 1000 copies/ml and was censored from the remaining 350 threshold strategy at month 24 because she was monitored at an interval shorter than 9-12 months while her CD4 cell count was above the 350 threshold and her HIV RNA \leq 200 copies/ml. Individual 3 was censored from the tight-control strategies at month 9 when she did not switch treatment within 3 months of her HIV RNA crossing above 200 copies/ml.

APPENDIX C

EXAMPLE DATA WITH THE “NO DIRECT EFFECT” ASSUMPTION

Table C1.1 shows how the data for hypothetical individual 1 from Table B1 can be modified to create a “no direct effect” data set.

First, we show individual 1’s “no direct effect” data under strategies that require individuals to be monitored exactly every 6 months if their CD4 cell count falls below the strategy’s threshold and exactly every 12 months otherwise. In the

TABLE C1.1 One hypothetical individual following all of the four joint monitoring and treatment strategies over 24 months of follow-up under the “no direct effect” assumption, **without grace periods**

Time (Months)	Monitor (1: Yes, 0: No)	CD4 Cell Count	“No Direct Effect” Data for Threshold 350 Strategies		“No Direct Effect” Data for Threshold 500 Strategies	
			Monitor (1: Yes; 0: No)	CD4 Cell Count	Monitor (1: Yes; 0: No)	CD4 Cell Count
0	1	300	1	300	1	300
1	0	300	0	300	0	300
2	0	300	0	300	0	300
3	0	300	0	300	0	300
4	0	300	0	300	0	300
5	0	300	0	300	0	300
6	1	380	1	380	1	380
7	0	380	0	380	0	380
8	0	380	0	380	0	380
9	0	380	0	380	0	380
10	0	380	0	380	0	380
11	0	380	0	380	0	380
12	1	400	0	380	1	400
13	0	400	0	380	0	400
14	0	400	0	380	0	400
15	0	400	0	380	0	400
16	0	400	0	380	0	400
17	0	400	0	380	0	400
18	1	430	1	430	1	430
19	0	430	0	430	0	430
20	0	430	0	430	0	430
21	0	430	0	430	0	430
22	0	430	0	430	0	430
23	0	430	0	430	0	430
24	1	500	0	430	1	500

original per-protocol analysis, individual 1 was censored from the threshold 350 strategies at month 12 because she was monitored at an interval shorter than 9-12 months while her CD4 cell count was above the 350 threshold. In the “no direct effect” data set, we changed the observations for the two replicates following the threshold 350 strategies as follows: we (i) recoded the monitoring indicator at month 12 to 0, (ii) discarded the CD4 cell count and HIV-RNA measurement recorded at month 12, and (iii) carried forward the CD4 cell count and HIV RNA from month 11 until the next time the individual was monitored at month 18. The monitoring indicator was again recoded to 0 at month 24. Monitoring is recoded to 0 at months 12 and 24 for the replicates following the threshold 350 strategies because they are monitored sooner than indicated by the strategy (every 6 months rather than every 12 months). Under the “no direct effect” assumption, individual 1 followed all four strategies for the first 24 months of follow-up. This “no direct effect” data set is shown in Table C1.1. Note that we exclude HIV RNA from the Table for simplicity.

Table C1.2 shows the “no direct effect” data for individual 1 under strategies with grace periods (monitor every 2-7 months if CD4 is above the threshold and every 8-13 months if CD4 is below the threshold). Consider the replicate following the threshold 350–tight control strategy. When this replicate is monitored at month 6, during a monitoring grace period, the replicate is split into two new replicates 1_6 and 2_6 . New replicate 1_6 has her new CD4 cell count and HIV-RNA measurements revealed and reenters the grace period in the usual way. This new replicate 1_6 is again split into two new replicates when monitored during another monitoring grace period at month 18. New replicate 2_6 has her monitoring at month 6 ignored (because she could still be monitored at month 7), moves to the next month of the monitoring grace period, and is censored at month 7 for not being monitored at the end of the grace period. The replicate following the threshold 350–loose control strategy is split at the same time points as the replicate following the threshold 350–tight control strategy. The two replicates following the threshold 500 strategies are split when monitored during grace periods at months 6, 12, 18, and 24. By leveraging the grace periods and the “no direct effect” assumption, six replicates corresponding to this hypothetical individual remain under follow-up at 24 months of follow-up: one follows the threshold 350–tight control strategy, one follows the threshold 350–loose control strategy, two follow the threshold 500–tight control strategy, and two follow-up the threshold 500–loose control strategy.

APPENDIX D

TREATMENT SWITCH

A summary of classifications of a treatment change as a nonswitch, ineligible switch, or switch is shown in Table D1, as follows.

TABLE D1 Changes from initial eligible^a regimens to new regimens that are considered switches^b (Center for AIDS Research Network of Integrated Clinical Systems and HIV-CAUSAL Collaboration)

Regimen Classification	Switch From (PI + ≥ 2 NRTI)?	Switch From (bPI + ≥ 2 NRTI)?	Switch From (NNRTI + ≥ 2 NRTI)?	Switch From (≥ 2 NRTI + FI/INSTI/EI)?
PI + ≥ 2 NRTI	No	No	Yes	Yes
bPI + ≥ 2 NRTI	Yes	Yes if PI changes	Yes	Yes
NNRTI + ≥ 2 NRTI	Yes	Yes	Yes if NNRTI to etravirine	Yes
bPI + PI/NNRTI (+ other)	Yes	Yes	Yes if NNRTI to etravirine	Yes
≥ 2 NRTI + FI/INSTI/EI	Yes	Yes	Yes	Yes if FI/INSTI/EI changes or addition of a FI/INSTI/EI

[Adapted from Cain et al, 2015]

^aEligible cART regimens consist of (i) at least two nucleoside reverse transcriptase inhibitors (NRTIs) plus one protease inhibitor (PI), (ii) at least two NRTIs plus one PI boosted with ritonavir (bPI), (iii) at least two NRTIs plus one nonnucleoside reverse transcriptase inhibitor (NNRTI), or (iv) at least two NRTIs plus a fusion inhibitor (FI), an integrase strand inhibitor entry (INSTI), or an entry inhibitor (EI).

^bWe define a treatment switch as a change of treatment regimen to any new regimen, as described previously.⁸ Each treatment switch is classified as a switch to an eligible or an ineligible regimen. In the Table above, switches to eligible regimens are labeled “Yes,” and changes in treatment that are not considered a switch are labeled “No.” Ineligible switches include any change not outlined in this Table, such as changes to monotherapy or dual therapy or stopping therapy all together.