STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:4678–4721
Published online 21 July 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3301

Estimation and extrapolation of optimal treatment and testing strategies

James Robins[1,3,*,†], Liliana Orellana[2] and Andrea Rotnitzky[3,4]

[1]*Department of Epidemiology, Harvard School of Public Health, Boston, MA, U.S.A.*
[2]*FCEyN, Universidad de Buenos Aires, Buenos Aires, Argentina*
[3]*Department of Economics, Di Tella University, Buenos Aires, Argentina*
[4]*Department of Biostatistics, Harvard School of Public Health, Boston, MA, U.S.A.*

SUMMARY

We review recent developments in the estimation of an optimal treatment strategy or regime from longitudinal data collected in an observational study. We also propose novel methods for using the data obtained from an observational database in one health-care system to determine the optimal treatment regime for biologically similar subjects in a second health-care system when, for cultural, logistical, or financial reasons, the two health-care systems differ (and will continue to differ) in the frequency of, and reasons for, both laboratory tests and physician visits. Finally, we propose a novel method for estimating the optimal timing of expensive and/or painful diagnostic or prognostic tests. Diagnostic or prognostic tests are only useful in so far as they help a physician to determine the optimal dosing strategy, by providing information on both the current health state and the prognosis of a patient because, in contrast to drug therapies, these tests have no direct causal effect on disease progression. Our new method explicitly incorporates this no direct effect restriction. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: dynamic regime; causal inference; marginal structural model; value of information

1. INTRODUCTION

The goal of this paper is to describe recent developments in the estimation of an optimal treatment strategy or regime from longitudinal data collected in an observational study. Estimation of the optimal time for an asymptomatic HIV-infected subject to start highly active retroviral therapy (HAART) will serve as a paradigmatic example. The following is a loose approximation to the

*Correspondence to: James Robins, Department of Epidemiology, Harvard School of Public Health, Boston, MA, U.S.A.
†E-mail: robins@hsph.harvard.edu

Contract/grant sponsor: NIH; contract/grant numbers: R01 GM48704, R0132475

Received 10 March 2008
Accepted 10 March 2008

current International Aids Society (IAS) 'when to start' recommendations:

- If plasma HIV RNA exceeds 150 000 copies/mL and CD4 decline has exceeded 100 cells/μl per year, start HAART when the CD4 first falls below 500.
- If HIV RNA is persistently low (<10 000 copies/mL) and CD4 decline is less than 50 cells/μl per year, start when the CD4 first falls below 200.
- Otherwise, start when the CD4 count first drops below 350.

The recommended strategy is an example of a deterministic dynamic treatment regime, that is, a sequential decision strategy in which the treatment to be given at time t is a deterministic function of the subject's measured time-dependent covariate (and possibly treatment history) up to t.

Some of the biological considerations and empirical data on which the IAS recommendations were based are as follows. The current understanding of the biology of the effect of HAART on HIV infection is not sufficient to determine the optimal time to start therapy because early initiation of HAART has both risks and benefits; early initiation can prevent viral-induced decline of, and permanent damage to, the immune system, but also allows more time for drug resistance and side effects to develop. Thus, the 'when to start' question must be addressed with empirical data. Randomized trial data have shown that for asymptomatic subjects with CD4 count <200, it is better to initiate HAART than to delay further. However, current recommendations for subjects with CD4 counts exceeding 200, such as the above IAS recommendations, are based solely on the analyses of observational data. Unfortunately, we show in Section 2 that the appropriate use of observational data to determine 'when to start' poses a nontrivial methodologic challenge. Specifically, we provide a somewhat simplified description of one of the observational analyses [1] that formed the basis of the 2004 IAS recommendations and show that the published analysis has potential for severe bias.

In Section 3, we discuss assumptions under which it is possible to use observational data to estimate the optimal treatment regime in a class of prespecified, logistically feasible dynamic regimes and describe an analytic approach based on dynamic marginal structural models (MSMs), which recovers the optimal regime in the class. Dynamic MSMs were introduced by Orellana *et al.* [2] and independently by Van der Laan [3]. In Section 4 we give a more precise mathematical formulation of the problem, the identifying assumptions, and our analytic methodology.

In Section 5, we describe an alternative approach based on doubly robust g-estimation of optimal regime structural nested mean models (SNMMs) that can optimize over a much bigger class of candidate regimes than the class that can be optimized over when using dynamic MSMs. This method was introduced by Robins [4] as a generalization of a closely related approach by Murphy [5]. See Moodie *et al.* [6] for additional discussion. G-estimation of optimal regime SNMMs is a robust twist on the classic method of dynamic programming (backward induction) for sequential decision making under uncertainty. Because optimal regime SNMM methods generally optimize over a larger class of regimes than do dynamic MSM methods, the optimal treatment strategy estimated using an optimal regime SNMM should generally have a higher expected utility than the optimal strategy estimated with a dynamic MSM. However, the estimated optimal decision strategy based on optimal regime SNMMs may be such a complicated function of past covariate history that in many health-care settings, particularly those found in developing countries, it may be logistically difficult or impossible to implement. In that case, dynamic MSM would be preferable. An example of the complex estimated optimal regimes that could be produced by fitting an optimal regime SNMM would be the following: Begin HAART the first time that the following quantity

DOI: 10.1002/sim

exceeds zero:

$17.2-4.1\times$ current CD4 count $-3.2\times$ the rate of decline of CD4 $+5.6\times$ the square of the rate of decline of CD4 $+2.1\times$ the current CD4 count $\times$ the rate of decline CD4 $-4.6\times$ the rate of decline of CD4 $\times$ the square of the current HIV RNA.

This regime would be logistically impossible to implement in settings where, because of poor record keeping or fragmented patient care, complete data on past CD4-count measurements do not exist.

In Section 6, we develop novel methods to use data obtained from an observational database in one health-care system to determine the optimal treatment regime for biologically similar subjects in a second health-care system when, for cultural, logistical, and financial reasons, the two health-care systems differ (and will continue to differ) in the frequency of, and reasons for, laboratory tests of diagnostic or prognostic value. In Section 7, we derive conditions under which the determination of the optimal treatment regime remains possible even though the observation processes (i.e. visit processes) in both health-care systems are nonignorable, provide a method of estimation under these conditions, and discuss the substantive plausibility of the derived conditions.

In Section 6, we also develop a novel method to estimate the optimal timing of expensive and/or painful diagnostic tests. Diagnostic tests, in contrast to drug therapies, have no direct causal effect on disease progression. That is, tests are only useful in so far as they help a physician to determine the optimal dosing strategy, by providing information on both the current health state and the prognosis of a patient. Our new method explicitly incorporates this no direct effect restriction.

2. POTENTIAL BIASES IN OBSERVATIONAL ANALYSES

A greatly simplified version of the design and analysis reported in Palella *et al.* [1] can be described as follows. Asymptomatic treatment-naive HIV-infected subjects were enrolled on January 1, 1996 and their CD4 count was measured. Subjects with CD4 counts from 325 to 350 at enrollment were followed until the minimum of time to death and clinical AIDS, with time measured as time since enrollment. Subjects with CD4 count above 300 at the time of HAART initiation were designated as group 1, and those with CD4 count below 200 at the time of HAART initiation were designated as group 2. Subjects who either never started HAART or started it when their CD4 count was between 200 and 300 were excluded from the analysis. Subjects in groups 1 and 2 were compared using a time-independent Cox proportional hazards model with an indicator variable for the group as the covariate and with time measured as time since enrollment.

Before we critique this approach, we begin by observing that it is better than a common alternative approach in which time in the Cox model is not time since enrollment on January 1, 1996 but time since initiation of HAART. Under this alternative approach, the survival time of those starting with CD4$>$300 (group 1) will be better than the survival time of those starting when CD4$<$200 (group 2) even under the null hypothesis that HAART has no causal effect on any subject's time to clinical AIDS or death. In fact, this is an instance of classic lead-time bias. Clinical AIDS or death rarely occurs before the CD4 count drops below 200. The time it takes a group 1 subject's CD4 count to fall from 300 to 200 is the lead time. We, therefore, see that a recommendation based on this alternative analysis would incorrectly recommend to 'start HAART when CD4 is above 300' when, indeed, HAART has no causal effect.

The approach based on time since enrollment also has the potential for severe bias when prescribing trends' change with calendar time. To make clear why this is so, we consider an example

DOI: 10.1002/sim

with extreme trends. Suppose that during 1996, no subject with a CD4 above 300 was treated with HAART because physicians were worried about using a newly introduced treatment regimen on patients who were doing reasonably well. However, all subjects with CD4 counts falling below 200 were treated. Suppose that during 1997 every subject with a CD4 below 350 was treated with HAART because doctors had come to believe that HAART was beneficial for most patients. Subjects who started HAART when their CD4 fell below 200 constitute a selected subgroup of very sick subjects with rapid CD4-count decline. Many of these subjects would probably have died before the end of 1996. In contrast, all subjects who started HAART when their CD4 count was above 300 must have started during 1997 and, consequently, must have had (i) at least one year of survival and (ii) a slow rate of CD4 decline, an indication of good health.

It follows that under the null hypothesis that HAART has no causal effect on any subject's time to clinical AIDS or death, the survival time in group 1 will be better than the survival time in group 2 even under the null hypothesis that HAART has no causal effect on any subject's time to clinical AIDS or death, so a recommendation based on this analysis will also incorrectly recommend to 'start HAART when CD4 is above 300'. The reason for this bias (which is sometimes referred to as 'immortal person time' bias) is that starting treatment with a CD4 above 300 is a surrogate for slowly declining counts and, in fact, for survival itself. In fact, the extreme prescribing trends of the example are but an exaggeration of the actual prescribing trends in 1996–1997, so the possibility of bias is real.

In fact, the approach based on time since enrollment can result in bias even in the absence of prescribing trends, because a subject's group status is determined by events that occur after start of follow-up and are prognostic for survival. Below we show how to eliminate this type of bias.

3. APPROPRIATE OBSERVATIONAL ANALYSES

3.1. A randomized clinical trial comparing two treatment strategies

Suppose we wanted to compare two simple strategies or regimes: strategy 1: 'start HAART when the CD4 count first falls below 500' *versus* strategy 2: 'start HAART when CD4 count first falls below 200' among the subset of people who at the time of diagnosis have CD4 count above 500. If we had the resources to conduct a randomized trial, we might proceed as follows. We enroll, at the time of HIV diagnosis, asymptomatic HIV-infected subjects with CD4 counts greater than 500 at diagnosis with HIV. We do not give HAART while a subject's CD4 count is above 500. We designate as start of follow-up (i.e time 0) the time when a subject's CD4 count first falls below 500 and at that time we randomize each subject to either 'start HAART when the CD4 count first falls below 500' (group 1) or to 'start HAART when the CD4 count first falls below 200' (group 2). We compare the survival of the two groups using a Cox proportional hazards model for the minimum of time to clinical AIDS or death with an indicator variable for group as the covariate and with time measured as time since start of follow-up (i.e. randomization). As such a trial has not been conducted, we must rely on observational data.

3.2. Observational analogue of a randomized trial

Our goal is to implement a design and analysis plan for observational data, which is the observational analogue of the above randomized trial. We follow an approach similar to the one described

DOI: 10.1002/sim

in Hernán *et al.* [7]. As in the randomized trial, we find asymptomatic HIV-infected subjects with CD4 counts greater than 500 at the time of diagnosis. Suppose, for now, that, in our data, no subject with a CD4 count exceeding 500 initiates HAART. Then we designate as the subject's start of follow-up time (i.e. time 0) the time at which the subject's CD4 count first falls below 500. Every subject who initiates HAART at time 0 is assigned to group 1. All subjects who do not start HAART at time 0 are assigned to group 2. The two groups are then compared using a time-independent Cox proportional hazards model with an indicator variable for group as a covariate and with time measured as time since start of follow-up. Because treatment was not randomized, we must also adjust for potential baseline confounding factors (measured at or before time 0) such as HIV RNA, calendar date of entry, pre-baseline rate of CD4-cell-count decline, and various higher-order interactions. We can adjust either by matching or stratifying on these baseline factors or by including them as covariates in the Cox model.

By definition, all group 1 subjects successfully follow the protocol 'start HAART when the CD4 count first falls below 500'. However, any group 2 patient who either (i) begins HAART at a time t at which his CD4 count still exceeds 200 or (ii) fails to start HAART at the first time that his CD4 count falls below 200 must be regarded as censored at time t because he has failed to follow the group 2 protocol we wish to test— 'start HAART when the CD4 count first falls below 200'. In our experience, essentially all untreated subjects whose CD4 count falls below 200 are started on HAART, so only censoring for reason (i) remains an issue.

Now, in a time independent Cox model, censoring is ordinarily handled simply by excluding a subject censored at t from all risk sets subsequent to t. However, this standard approach is statistically valid only if, conditional on the baseline covariates in the model, censoring and failure (the minimum of time to clinical AIDS or death) are independent. However, as we now argue, censoring for reason (i) is dependent.

To understand why, consider all group 2 subjects with CD4 cell count still greater than 200 cells/mL at time t after entry. Conditional on the baseline covariates, the subgroup who initiates HAART (and thus becomes censored) at t will have, on average, a faster CD4-cell-count decline from 0 to t and higher HIV RNA at t than those who do not initiate HAART (and thus remain uncensored) at time t. This reflects the fact that physicians preferentially prescribe HAART to patients doing poorly. Thus, within any stratum of baseline covariates, individuals censored at t will have worse prognosis than those who remain uncensored.

We conclude that a standard Cox analysis (which assumes independent censoring) will show the failure rate of the group 2 subjects artificially lowered compared with what would have been seen in the above ideal randomized trial in which all group 2 subjects wait until a CD4 of 200 to start. Thus, if in truth, it is better to start at 500 than at 200, a standard Cox analysis may fail to detect this fact due to selection bias.

We can try to correct for bias due to dependent censoring by inverse probability of censoring weighting (IPCW) of each subject in each risk set. That is, each group 2 subject in a risk set at time u who remains uncensored at u is given a weight equal to the inverse of the conditional probability of his having remained uncensored up to time u, given both his baseline covariates and his history up to u of post-baseline time-dependent prognostic covariates (such as HIV RNA and CD4 cell count). Specifically, suppose a group 2 subject at risk and uncensored at u with a CD4 count falling linearly from 500 to 250 from 0 to u has a probability of $\frac{1}{4}$ that he would not have started treatment by u. Then he counts for four people: himself and the three other similar people, i.e. with the same rate of CD4-count decline, who did start therapy. That is, his weight is 4 in the risk set at u. On the other hand, suppose that an uncensored subject with a CD4 falling from 500

DOI: 10.1002/sim

to 300 from 0 to u has probability of $\frac{1}{2}$ of having remained uncensored. Then, he counts for only two persons with a similar CD4-count decline rate.

Formally, let C be the discrete random variable encoding time to censoring measured in weeks. Let $\text{pr}[C \neq j | C > j-1, \text{Past}(j)]$ be the discrete hazard of remaining uncensored at week j given all the past measured (time-dependent and baseline) covariate history Past(j), where we include the indicator $I(T>j)$ of nonfailure in Past(j) (that is, $I(T>j)=1$ if T, the minimum of time to AIDS or death exceeds j and is zero otherwise). Then, to estimate the parameters of the Cox model for failure, we compute the usual partial likelihood estimator except that for each subject in a risk set at time u, we multiply his contribution to the risk set by an estimate $\widehat{W}(u)$ of

$$W(u) = 1 \Big/ \prod_{j=1}^{u} \text{pr}[C \neq j | C > j-1, \text{Past}(j)] \tag{1}$$

the inverse of the probability of having remained uncensored to time u. We refer to such a procedure as an IPCW-adjusted Cox analysis.

In our discussion, it will prove useful to have an alternative expression for the weight $W(u)$. Let $A(t)=1$ if a subject has started HAART by week t and $A(t)=0$ otherwise. Let $\overline{A}(k-1)=\{A(0), A(1), \ldots, A(k-1)\}$ be treatment history through $k-1$. Then define $f[a(k)|\overline{a}(k-1), \text{past}(k)] = \text{pr}(A(k)=a(k)|\overline{A}(k-1)=\overline{a}(k-1), \text{Past}(k)=\text{past}(k))$. Define

$$W(u) = 1 \Big/ \prod_{k=1}^{u} f[A(k)|\bar{A}(k-1), \text{Past}(k)] \tag{2}$$

where, for any subject, $f[A(k)|\bar{A}(k-1), \text{Past}(k)]$ is the density $f[a(k)|\overline{a}(k-1), \text{past}(k)]$ evaluated at the subject's observed data. The denominator of $W(u)$ is informally the probability that a subject had his observed HAART history through u. Note, by definition, if $A(k-1)=1$, then $A(k)=1$ so $f[A(k)|\bar{A}(k-1), \text{Past}(k-1)]=1$ unless past HAART history is identically 0, i.e. $\bar{A}(k-1)=0$. For uncensored subjects in a risk set at u, this alternative definition of $W(u)$ agrees with the previous definition. To see why, note that all group 1 subjects started HAART at time 0 so $A(0)=1$ and thus $W(u)=1$. Further, by definition, any group 2 subject who is uncensored at u must have an observed HAART history that agrees with that specified by the regime 'start HAART when CD4 count first falls below 200.' Thus, the probability that the subjects remained uncensored is precisely the probability that the subjects had their observed HAART history.

Suppose we specify the discrete time logistic model

$$\text{logit}\{\text{pr}(A(k)=1|\overline{A}(k-1)=0, \text{Past}(k), I(T>k)=1)\} = \alpha^{\text{T}} Q(k)$$

where $Q(k)^{\text{T}} = (Q_1(k), Q_2(k), Q_3(k))$, $Q_1(k)$ is HIV RNA at week k, $Q_2(k)$ is the lowest recorded CD4 count up to k, and $Q_3(k)$ is the rate of CD4-count decline from 0 to k. Let

$$\widehat{W}(u) = 1 \Big/ \prod_{k=1}^{u} f[A(k)|\bar{A}(k-1), \text{Past}(k); \widehat{\alpha}]$$

where $\widehat{\alpha}$ is the partial maximum likelihood estimator (MLE) of α. For uncensored subjects at risk at u, $\widehat{W}(u)=1$ for all group 1 subjects since $A(0)=1$, while for uncensored group 2 subjects, $f[A(k)|\bar{A}(k-1), \text{Past}(k); \widehat{\alpha}]=1$ if $A(k-1)=1$, $f[A(k)|\bar{A}(k-1), \text{Past}(k); \widehat{\alpha}] = e^{\widehat{\alpha}^{\text{T}} Q(k)}/\{1+e^{\widehat{\alpha}^{\text{T}} Q(k)}\}$ if $Q_2(k)<200$ and $A(k-1)=0$ (since then $A(k)=1$), and $f[A(k)|\bar{A}(k-1), \text{Past}(k); \widehat{\alpha}] = 1/\{1+e^{\widehat{\alpha}^{\text{T}} Q(k)}\}$ if $Q_2(k) \geqslant 200$ and $A(k-1)=0$ (since then $A(k)=0$).

DOI: 10.1002/sim

This IPCW-adjusted Cox analysis produces valid inferences analogous to those that would be found with our idealized randomized trial if (a) the Cox model for failure and the logistic model for HAART initiation are correctly specified, (b) there is no residual confounding by any unmeasured baseline covariates, and (c) Past(j) contains sufficient time-dependent and baseline covariates so that censoring and failure are conditionally independent at each j within group 2. In contrast to assumption (a), assumptions (b) and (c) are not empirically testable. Assumptions (b) and (c) can be incorporated in the single assumption of no unmeasured confounders or sequential randomization that specifies

$$T_x \amalg A(t) | \bar{A}(t-1) = 0, \text{Past}(t), \quad t = 0, 1, \ldots \tag{3}$$

for $x = 200$ and 500, where T_x is a subject's potential or counterfactual failure time had the subject followed the dynamic regime in which HAART is initiated the first time the subject's measured CD4 count falls below x. Here Past(0) are the baseline covariates.

Henceforth, we suppose that there exist subjects in our data set who begin HAART before their CD4 count drops to 500. The preceding IPCW-adjusted Cox analysis can be easily modified to account for this additional complication. Specifically, we proceed as above except with the following modifications: (i) all subjects who initiate HAART with a CD4 count exceeding 500 are excluded from the Cox model analysis comparing groups 1 and 2 and (ii) each subject included in the Cox analysis who is in a risk set at u receives a weight $\{\widehat{W}(u)\widehat{W}^*\}$, where $\widehat{W}(u)$ is as previously defined and $\widehat{W}^*$ is an estimate of W^*, the inverse of the conditional probability that a subject did not initiate therapy during the interval from time of diagnosis of infection with HIV to the time his CD4 count was first observed to fall below 500. That is, if we let T_{500} denote the time a subject's CD4 count first fell below 500 with time measured as time since HIV diagnosis, $W^* = 1/\prod_{t=1}^{T_{500}} f[A(t)|\bar{A}(t-1), \text{Past}(t)]$ if $T_{500} \neq 0$ and $W^* = 1$ if $T_{500} = 0$. As an example, consider a subject who started HAART the first time his CD4 fell below 500 and so is in group 1 with $\widehat{W}(u) = 1$. Suppose the subject's estimated probability of having started HAART before his CD4 fell to 500 was $\frac{1}{2}$. Then $\widehat{W}^* = 2$ and the subject is given a weight of 2 whenever he contributes to a risk set.

In contrast with our IPCW analyses, if one were to add post-baseline CD4 count and HIV RNA as a time-dependent covariate in the Cox model for failure (rather than using IPCW weights), bias due to dependent censoring would not be eliminated; furthermore, new bias could be introduced from regression adjustment for a post-baseline variable if that variable was affected by treatment group. In summary, correction for dependent censoring can be accomplished by using post-baseline time-dependent covariates, not as covariates in a time-dependent Cox model, but rather to estimate the IPC weights for a time-independent Cox model.

3.2.1. Data analysis. Hernán *et al.* [7] conducted an analysis of the 2344 HIV-infected subjects included in the French Hospital Database on HIV (FHDH) [8] who had their first CD4 cell count measurement below 500 cells/mL between January 1, 1996 and June 30, 2004 while being treatment-naive. They followed these subjects from their first CD4 cell count measurement below 500 cells/mL (baseline) until a diagnosis of AIDS, death, or June 2004, whichever occurred earlier. Data on HAART use, as well as on time-dependent covariates (e.g. CD4 cell count), were recorded throughout the follow-up. Groups 1 and 2 were defined as in the previous subsection except that they followed subjects from entry into the FHDH cohort because time of diagnosis was often not recorded in the database. There were 131 subjects in group 1 and 2217 in group 2. Six hundred and

DOI: 10.1002/sim

fifty subjects in group 2 were censored from starting HAART when their CD4 count still exceeded 200 and five were censored for failing to start HAART the first time their CD4 was below 200.

A Cox analysis found a hazard ratio of 0.9 (95 per cent confidence interval: 0.4, 1.8) comparing groups 1 and 2 with a standard Cox analysis that adjusted for baseline potential confounders but assumed independent censoring. In contrast an IPCW Cox analysis found a hazard ratio of 0.5 (95 per cent confidence interval: 0.2,1.1), so as expected on theoretical grounds, the standard Cox analysis underestimated the benefit of starting HAART at a CD4 of 500.

3.2.2. Time since diagnosis as an alternative analytic time scale. In the hypothetical randomized trial described earlier, it makes no logical difference whether we randomize each subject to one of the two regimes—'start HAART when the CD4 count first falls below 500' (group 1) or to 'start HAART when the CD4 count first falls below 200' (group 2) —at time of enrollment (i.e. time of diagnosis of HIV infection) *versus* at the time when the CD4 count first falls below 500. Had we done the former, time since randomization would denote time since the diagnosis of HIV infection rather than time since a subject's CD4 count first fell below 500. It follows that had we compared the survival of the two groups using a Cox model for the minimum of time to clinical AIDS or death with an indicator variable for group as the covariate, an interaction between time and group, and with time measured as time since the diagnosis of HIV infection, the rate ratios of our Cox model would still have a valid causal interpretation. However, the rate ratios would have a different causal interpretation (and different magnitude) than the rate ratios in our earlier Cox model that used time since the CD4 count first fell below 500. We next describe the observational analogue of the Cox model analysis of our randomized trial with time now measured as time since HIV diagnosis.

Specifically in our observational analogue, we again restrict to subjects whose CD4 count exceeded 500 at time of diagnosis. However, baseline covariates would now be covariates recorded at the time of HIV diagnosis and the time in the Cox model would become time from diagnosis. Each subject would be placed in both groups 1 and 2 at time of diagnosis. We use a robust variance estimator and the Breslow estimator to handle the correlations and ties thereby induced. In each treatment group, we censor a subject the first time they fail to follow the group-specific regime. Among uncensored subjects in a risk set at time u, we weight each subject by $\widehat{W}(u)$ calculated as earlier, except that u now indicates time since diagnosis. Subjects in group 1 are now subject to censoring. In particular, group 1 subjects are censored if they start HAART when their CD4 is above 500 or fail to start HAART when their CD4 first drops below 500. For concreteness and without loss of generality, in the remainder of this paper we will assume that time is measured as time from diagnosis of HIV infection.

3.3. Choosing the optimal regime among more than two regimes

In this subsection we consider two important issues we have yet to address. First, in practice we want to choose between more than two candidate regimes. For example, if we wish to find the optimal CD4 count x at which to start HAART, we would want to compare all x's in the candidate set $\{500, 499, \ldots, 200\}$ rather than just two dynamic regimes.

Second, to determine which x is optimal, we need a well-defined numerical measure to rank regimes. For example, one might wish to choose x to maximize the expected 'years (or quality-adjusted years) of life'. We refer to the random variable Y whose expectation we wish to maximize as a (subject-specific) utility function. Our goal is then to find the regime x that maximizes expected

DOI: 10.1002/sim

utility. Expected years (or quality-adjusted years) of life measures have a much more natural and useful public health and policy interpretation than the hazard ratio measure that is typically reported after fitting a Cox model.

However 'years (or quality-adjusted years) of life' cannot be assessed when follow-up of the cohort is not to extinction but rather only to a fixed administrative end-of-follow-up date $K+1$ weeks from start of a subject's follow-up. Our methods can be extended to settings in which the potential length of follow-up varies with the subject, but, for simplicity, we do not do so in this paper. Clearly, among survivors at the administrative end of follow-up, the healthier ones (according to an agreed on measure of current health) have a greater post-study expected survival (and thus warrant a higher utility Y) than the less healthy survivors and a much greater expected survival (and thus warrant a much higher utility) than the nonsurvivors who died just before $K+1$. For example, in the context of the FHDH data, a possible choice of utility Y would be

$$Y = \text{time to death if death occurred before } K+1$$

$$Y = K+4\frac{\text{CD4}}{500} \quad \text{if alive at } K+1$$

Henceforth, we assume that some agreement has been reached on the choice of utility Y so data on Y is available for all subjects, if as we assume throughout, there is no drop-out.

To formalize the problem, we let Y_x be a subject's potential or counterfactual utility had the subject followed the dynamic regime in which HAART *is to be* initiated the first time the subject's measured CD4 count falls below x. Our goal is to find $x_{\text{opt}} \in \{500, 499, \ldots, 201\}$ for which the expected utility $E(Y_x)$ is a maximum. We first show how to design and analyze a randomized trial to determine x_{opt}. We then consider how to mimic the trial using observational data. Throughout we will assume that there exists a unique x_{opt}.

3.3.1. A randomized trial. Consider a randomized clinical trial (RCT) with full compliance with X, taking values in $\{500, 499, \ldots, 201\}$, the random variable recording the assigned dynamic regime. Because X was randomly assigned, it is independent of the counterfactuals Y_x. Because if $X=x$ then $Y=Y_x$, the observed utility Y is a function of the set of counterfactuals $\{Y_x; x = 500, 499, \ldots, 201\}$ and the treatment assignment X. Suppose the data on each subject are (Y, X). Then the average $\widehat{E}[Y|X=x]$ of Y among subjects randomized to regime x is an unbiased estimator of $E[Y_x]$. [This follows from the fact that (i) $\widehat{E}[Y|X=x]$ is unbiased for $E[Y|X=x]$ and (ii) $E[Y|X=x]=E[Y_x|X=x]=E[Y_x]$ by consistency and randomization.] Thus, the natural estimate of x_{opt} is the value of x for which $\widehat{E}[Y|X=x]$ is a maximum. However, even if, say, 900 subjects are enrolled in the trial, the number of subjects randomized to treatment arm x must be three or less for the majority of the 300 treatment arms. For these arms, $\widehat{E}[Y|X=x]$ will be highly variable. As a consequence, the natural estimator of x_{opt} will be so variable as to be useless. Thus, even in a randomized trial, to obtain a reasonably efficient estimator of x_{opt}, it is necessary to use our prior biological knowledge that $E[Y_x]$ is a smooth function of x. One, among many, possible approaches to exploiting this smoothness is to assume that $E[Y_x]$ is a flexible polynomial, for example, a polynomial of degree 5. If we assume

$$E[Y_x]=\beta_0+\sum_{k=1}^{5}\beta_k x^k$$

DOI: 10.1002/sim

then, by $E[Y_x]=E[Y|X=x]$, we have that the estimator $\widehat{\beta}$ of $\beta=(\beta_0,\ldots,\beta_5)$ obtained from fitting by ordinary least squares (OLS) the regression model $Y=\beta_0+\sum_{k=1}^5\beta_k X^k+\varepsilon$ to the observed data $(Y_i, X_i), i=1,\ldots,n$, on the n study subjects is unbiased for $\beta=(\beta_0,\ldots,\beta_5)^{\mathrm{T}}$. It follows that an unbiased estimator of $\widehat{E}[Y_x]$ is $\widehat{\beta}_0+\sum_{k=1}^5\widehat{\beta}_k x^k$. We then use first year calculus to find the value $\widehat{x}_{\mathrm{opt}}$ of x at which the fitted polynomial $\widehat{E}[Y_x]=\widehat{\beta}_0+\sum_{k=1}^5\widehat{\beta}_k x^k$ attains its maximum on the closed interval $[201, 500]$. Of course, the $\widehat{x}_{\mathrm{opt}}$ one obtains in this manner will not generally be an integer, but can always be rounded off to the closest integer in $\{500, 499, \ldots, 201\}$. For simplicity, we assume that the maximum is attained at a unique $\widehat{x}_{\mathrm{opt}}$.

The quintic model for $E[Y_x]$ will, of course, not be exactly correct. If incorrect, the estimator $\widehat{x}_{\mathrm{opt}}$ will be biased as an estimator of the true x_{opt}. Using a higher-degree polynomial will decrease bias but increase variance. Cross-validation might be used to select the degree of the flexible polynomial.

By including interactions with the measured pretreatment variables V, such as gender, ethnicity, HIV risk group (e.g. IV drug users versus homosexual contact), and genetic profile (single nucleotide polymorphisms), we can allow for the fact that the optimal treatment regime in our candidate set may differ depending on a subject's measured pretreatment variables V.

For instance, consider the model $E[Y_x|V]=\beta_0+\sum_{k=1}^5\beta_k x^k+\eta_0^{\mathrm{T}}V+\sum_{k=1}^5\eta_k^{\mathrm{T}}Vx^k$. Let $(\widehat{\beta},\widehat{\eta})$ be the OLS estimates from the fit of $Y=\beta_0+\sum_{k=1}^5\beta_k X^k+\eta_0^{\mathrm{T}}V+\sum_{k=1}^5\eta_k^{\mathrm{T}}VX^k+\varepsilon$ to the observed data $(Y_i, X_i, V_i), i=1,\ldots,n$. The optimal CD4 count $x_{\mathrm{opt}}(V)$ at which a subject with baseline covariates V should start is given by the value $x_{\mathrm{opt}}(V)$ that maximizes $E[Y_x|V]$, over all allowable x, which is also the $x(V)$ that maximizes $E[Y_x|V]-E[Y_{x_0}|V]$ at any fixed value x_0. Choosing $x_0=0, x_{\mathrm{opt}}(V)$ maximizes $\sum_{k=1}^5\beta_k x^k+\sum_{k=1}^5\eta_k^{\mathrm{T}}Vx^k=\sum_{k=1}^5(\beta_k+\eta_k^{\mathrm{T}}V)x^k$. Thus, our estimate $\widehat{x}_{\mathrm{opt}}(V)$ of $x_{\mathrm{opt}}(V)$, which for ease of reference below we call the RCT estimator of $x_{\mathrm{opt}}(V)$, is given by the (assumed unique) value $x(V)$ that maximizes $\sum_{k=1}^5(\widehat{\beta}_k+\widehat{\eta}_k^{\mathrm{T}}V)x^k$ for the given V.

3.3.2. An observational study and dynamic MSMs. To understand how to mimic the above randomized trial, we require the following observations. Suppose data on when, if ever, a subject began HAART and data on successive CD4 counts are available. Suppose no subject has two CD4 measurements in the same week. Consider a subject who started HAART at a CD4 cell count of 250 in week t whose lowest prior CD4 count was 300. Then this subject's observed data are consistent with having followed regime x for $x=251, 252, \ldots, 300$. In fact, the subject followed all of these regimes. Next consider a subject who never started therapy and whose lowest CD4 count was 225. Then this subject followed regimes x for $x=201, 202, \ldots, 225$. Finally, consider a subject who started antiretroviral therapy at a CD4 cell count of 250 in week t whose lowest previous CD4 count was 250 or less. Then this subject failed to follow any regime in the set $\{500, 499, \ldots, 201\}$.

We are now ready to describe our procedure. Time will be in weeks since enrollment in the cohort. As we did for the RCT, we assume the model $E[Y_x|V]=\beta_0+\sum_{k=1}^5\beta_k x^k+\eta_0^{\mathrm{T}}V+\sum_{k=1}^5\eta_k^{\mathrm{T}}Vx^k$. Let Γ_i be the number of regimes in the set $\{500, 499, \ldots, 201\}$ followed by subject i. We create an artificial data set of size $\Gamma=\sum_{i=1}^n\Gamma_i$, with each subject i, for $i=1,\ldots,n$, contributing Γ_i observations $(Y_i, V_i, X_{i1}), (Y_i, V_i, X_{i2}), \ldots, (Y_i, V_i, X_{i\Gamma_i})$, where the $X_{ik}, k=1,\ldots,\Gamma_i$, denote the regimes followed by subject i.

Our estimator of (β,η) is computed as the weighted least-squares estimator $(\widehat{\beta},\widehat{\eta})$ from the fit of the regression model $Y=\beta_0+\sum_{k=1}^5\beta_k X^k+\eta_0^{\mathrm{T}}V+\sum_{k=1}^5\eta_k^{\mathrm{T}}VX^k+\varepsilon$ to the artificial data set of

DOI: 10.1002/sim

size Γ, pretending all observations in the artificial data set are independent and weighting each artificial observation (Y_i, V_i, X_{ik}) with the weights $\widehat{W}_i(K)$, i.e. the earlier weight $\widehat{W}_i(u)$ with u set to the last time K that one could start therapy before administrative end of follow-up at $K+1$. As in our RCT, $\widehat{x}_{\text{opt}}(V)$ is given by the value $x(V)$ that maximizes $\sum_{k=1}^{5}(\widehat{\beta}_k+\widehat{\eta}_k^{\mathrm{T}}V)x^k$. Our observational estimator $\widehat{x}_{\text{opt}}(V)$ of $x_{\text{opt}}(V)$ will be consistent if (i) our assumption of no unmeasured confounders holds for all $x\in\{500, 499, \ldots, 201\}$, i.e. equation (3) with Y_x replacing T_x, (ii) the model $E[Y_x|V]-E[Y_0|V]=\sum_{k=1}^{5}(\beta_k+\eta_k^{\mathrm{T}}V)x^k$ is correct for all $x\in\{500, 499, \ldots, 201\}$, (iii) the model for treatment initiation used to obtain $\widehat{W}(K)$ is correct, and (iv) V is the vector of all baseline covariates Past(0) needed to ensure no unmeasured confounding.

We stress that condition (ii) implies that our estimate $\widehat{x}_{\text{opt}}(V)$ is consistent even if the model $E[Y_0|V]=\beta_0+\eta_0^{\mathrm{T}}V$ for the main effect of V is wrong. The remark in Section 4.3 provides an explanation.

For logistical reasons, we may only be interested in $x_{\text{opt}}(V)$ when V is a strict subset of the set Past(0) of baseline potential confounders. Fortunately, if we redefine

$$W(u)=1\Big/\prod_{k=0}^{u} f[A(k)|\bar{A}(k-1),\text{Past}(k)]$$

$$\widehat{W}(u)=1\Big/\prod_{k=0}^{u} f[A(k)|\bar{A}(k-1),\text{Past}(k);\widehat{\alpha}]$$

to include time 0 in the denominator, the above estimate of $\widehat{x}_{\text{opt}}(V)$ remains consistent when (iv) is replaced by 'V is a subset of Past(0)', as the redefined estimated weight $\widehat{W}(u)$ serves to control confounding by variables included in Past(0) but not in V.

Finally, as discussed further in Section 4.3, we can use dynamic MSMs to optimize a more complex class of candidate regimes in both randomized and observational studies. For example, we might consider the class of regimes:

- if current HIV RNA is greater than z, start HAART if the current CD4 count is less than x,
- if current HIV RNA is not greater than z, start HAART if the current CD4 count is less than q,

and use methods similar to those just described to jointly estimate the three numbers $(z_{\text{opt}}, x_{\text{opt}}, q_{\text{opt}})$.

4. A FORMALIZATION

4.1. The data

We consider a study of the effect of a time-dependent dichotomous exposure $A(t)$ on a utility function Y. The terms exposure and treatment will be used synonymously and interchangeably. We assume a fixed study population, i.e. a closed cohort with a well-defined, known start of follow-up date for each subject. Time t will refer to time in weeks since start of follow-up, which we also refer to as time since baseline. We only consider the estimation of the effect of exposures occurring at or after the start of follow-up because the estimation of the effects of pre-baseline exposures is not possible without making strong untestable assumptions. Subjects change exposure only at

DOI: 10.1002/sim

the beginning of a week so $A(t)$ is recorded at fixed times $t=0,1,\ldots,K$ weeks from baseline. Covariates $L(t)$ are recorded at times $t=0,1,\ldots,K,K+1$. Data on $L(t)$ become available prior to the determination of $A(t)$ but after $A(t-1)$. Baseline covariates $L(0)$ refer to all covariates, including pre-baseline exposure, that occur prior to the baseline exposure $A(0)$. We use overbars to denote history, i.e. the exposure history through time (i.e. week) t is $\bar{A}(t)=\{A(0),A(1),\ldots,A(t)\}$. We denote a subject's total exposure and covariate history by $\bar{A}=\bar{A}(K)$ and $\overline{L}=\overline{L}(K+1)$. We let $\overline{O}=\overline{O}(K+1)=(\bar{A},\overline{L})$. Utility Y is formally defined as $Y=y(\bar{A},\overline{L})$ with $y(\cdot,\cdot)$ being a known function. We use capital letters to denote random variables, scriptic letters to denote their sample space (i.e. the set of possible realizations), and lower case letters to denote elements of the sample space.

4.2. Treatment regimes and counterfactuals

Let $\bar{a}=\bar{a}(K)=\{a(0),a(1),\ldots,a(K)\}$ be an element of the sample space $\overline{\mathscr{A}}$ of $\overline{A}$. For simplicity, we take $\overline{\mathscr{A}}$ to be a finite set. Let $\overline{L}_{\bar{a}}(m)=\{L(0),L_{a(0)}(1),\ldots,L_{\bar{a}(m-1)}(m)\}$ be a subject's counterfactual L-history through m under the regime $\overline{a}$. Note that $\overline{L}_{\bar{a}}(m)$ only depends on $\bar{a}$ through $\overline{a}(m-1)$, as the future treatments cannot determine past responses. Define $\overline{L}_{\bar{a}}=\overline{L}_{\bar{a}}(K+1)$ to be a subject's complete counterfactual L-history under regime $\overline{a}$. Let $\overline{L}_{\overline{\mathscr{A}}}=\{\overline{L}_{\bar{a}};\overline{a}\in\overline{\mathscr{A}}\}$. A realization $\overline{l}_{\overline{\mathscr{A}}}=\{\bar{l}_{\bar{a}};\overline{a}\in\overline{\mathscr{A}}\}$ of $\overline{L}_{\overline{\mathscr{A}}}$ is a set of covariate histories indexed by the elements $\overline{a}$ of $\overline{\mathscr{A}}$. This notation implicitly includes the assumption that a subject's counterfactual responses do not depend on the treatments given to other subjects.

We make the following assumption linking the counterfactual data and the observed data:

Consistency assumption (C): $\overline{L}(m)=\overline{L}_{\overline{A}(m-1)}(m)$.

This assumption implies that $\overline{O}=(\overline{A},\overline{L})$ equals $(\overline{A},\overline{L}_{\overline{A}})$. That is, if $\overline{A}=\overline{a}$, a subject's observed $\overline{L}$ is obtained by selecting from $\overline{L}_{\overline{\mathscr{A}}}$ the element $\overline{L}_{\overline{a}}$.

So far we have only considered static regimes $\overline{a}$. To characterize the optimal treatment strategy, it is usually necessary to consider dynamic regimes as well.

We use g to denote a general regime. A nonrandom regime g is a treatment strategy or rule in which the treatment prescribed by g at time t depends in a deterministic manner on the evolution of a subject's measured time-dependent covariates $\bar{L}(t)$ and, possibly, treatments $\overline{A}(t-1)$ up to t. An example would be the dynamic regime 'take the treatment methotrexate at week t if and only if the neutrophil count has been greater than 1000 for three consecutive weeks and the patient was not on treatment at week $t-1$'. Mathematically, a nonrandom regime g is a collection of functions $\overline{d}=\{d_k[\bar{a}(k-1),\bar{l}(k)];k=0,\ldots,K\}$ with the range of d_k contained in the sample space $\mathscr{A}(k)$ of $A(k)$ and such that $d_k[\bar{a}(k-1),\bar{l}(k)]$ specifies the treatment to be taken at k for a subject with past history $[\bar{a}(k-1),\bar{l}(k)]$. In our methotrexate example, $d_k[\bar{a}(k-1),\bar{l}(k)]$ is 1 if a subject's $a(k-1)$is zero and his $\bar{l}(k)$ implies that his neutrophil count has been greater than 1000 at weeks $k,k-1,k-2$ (so k must be at least 2); otherwise $d_k[\bar{a}(k-1),\bar{l}(k)]$ is 0. We write either $g_{\overline{d}}$ or $g=\overline{d}$ to indicate the dependence on the functions $\overline{d}$. Earlier we considered the nonrandom regime $\overline{d}_x$: begin antiretroviral therapy the first time t the measured CD4 count falls below x. This regime has $d_{x,k}[\bar{a}(k-1),\bar{l}(k)]=1$ if minimum CD4 count through k is less than x and is 0 otherwise. If, for all k's, $d_k[\bar{a}(k-1),\bar{l}(k)]$ is a constant $a(k)$ that does not depend on $(\bar{a}(k-1),\bar{l}(k))$, the regime $\overline{d}$ is said to be nondynamic or static and is written as $g=\overline{a}$.

The regime $\overline{d}$ is naturally associated with counterfactual random variables $(\overline{A}_{\overline{d}},\overline{L}_{\overline{d}})$ that represent a subject's treatment and covariate history when following the regime. Specifically we regard

DOI: 10.1002/sim

$\overline{d}$ as defining a mapping $\overline{d}:\overline{L}_{\overline{\mathscr{A}}}\mapsto d(\overline{L}_{\overline{\mathscr{A}}})\equiv(\overline{A}_{\overline{d}},\overline{L}_{\overline{d}})$ defined recursively as follows: $L_{\overline{d}}(0)=L(0)$, $A_{\overline{d}}(0)=d_0[\overline{L}_{\overline{d}}(0)]$. For $m=1,\ldots,K+1$, $L_{\overline{d}}(m)=L_{\overline{A}_{\overline{d}}(m-1)}(m)$, $A_{\overline{d}}(m)=d_m[\overline{A}_{\overline{d}}(m-1),L_{\overline{A}_{\overline{d}}(m-1)}(m)]$ except $A_{\overline{d}}(K+1)$ is left undefined. Furthermore, $Y_{\overline{d}}=y(\overline{A}_{\overline{d}},\overline{L}_{\overline{d}})$.

In addition to the preceding consistency assumption, we make the following SR assumption and positivity assumption.

Sequential randomization (SR) assumption: For each k, $A(k)$ is independent of $\overline{L}_{\overline{\mathscr{A}}}$ given $(\overline{A}(k-1),\overline{L}(k))$. Equivalently,

$$f(\overline{A}|\overline{L}_{\overline{\mathscr{A}}})=\prod_{k=0}^{K}f[A(k)|\overline{A}(k-1),\overline{L}_{\overline{A}}(k)]=\prod_{k=0}^{K}f[A(k)|\overline{A}(k-1),\overline{L}(k)]$$

where the last equality follows from assumption C.

Positivity assumption (PO): $f\{a(k)|\overline{A}(k-1),\overline{L}(k)\}>0$ with probability one for all $a(k)\in\mathscr{A}(k)$.

4.3. Dynamic and general MSMs

To help understand the general formulation of dynamic and general MSMs it will be helpful to introduce them in connection with the example discussed in earlier sections on determining the optimal CD4 level at which to start treating HIV-positive subjects with HAART. Henceforth, consider an observational study of subjects with *CD4* counts exceeding 500 at time of diagnosis of HIV infection. Subjects return to the clinic weekly to have various clinical and laboratory measurements made. Let $L(t)$ be the vector of measurements made at week t including CD4 cell count. We let $A(t)$ denote the indicator of whether HAART has been initiated at week t or before. Let Y be a utility known at the end of follow-up at $K+1$, higher values of which are preferable. Let $\overline{d}_x$ denote the dynamic regime 'begin antiretroviral therapy the first time t the measured CD4 count falls below x,' where $x\in\mathscr{X}=\{201,\ldots,500\}$. The model

$$E[Y_{\overline{d}_x}|V]=h(x,V,\psi^*) \tag{4}$$

where

$$h(x,V,\psi)=h_1(x,V,\psi_1)+h_2(V,\psi_0)\quad\text{with } h_1(x,V,0)=0 \tag{5}$$

is an example of a dynamic regime MSM for the conditional counterfactual mean of Y, given a subset V of the baseline covariates $L(0)$. Note that the value $\psi_1^*=0$ is equivalent to the null hypothesis that all regimes in $\{\overline{d}_x;x\in\mathscr{X}\}$ have the same mean, given V.

In Section 3.3, we chose $\psi_1=\{\beta_1,\ldots,\beta_5,\eta_1^{\mathrm{T}},\ldots,\eta_5^{\mathrm{T}}\}^{\mathrm{T}}$, $\psi_0=(\beta_0,\eta_0^{\mathrm{T}})^{\mathrm{T}}$, $h_1(x,V,\psi_1)=\sum_{k=1}^{5}(\beta_k+\eta_k^{\mathrm{T}}V)x^k$, and $h_2(V,\psi_0)=\beta_0+\eta_0^{\mathrm{T}}V$. More generally, suppose $h(x,V,\psi)=(r(x,V)^{\mathrm{T}},r^*(V)^{\mathrm{T}})\psi$ for some known functions $r(x,V)$ and $r^*(V)$. Let Γ_i be the number of regimes followed by subject i and create an artificial data set of size $\Gamma=\sum_{i=1}^{N}\Gamma_i$, with each subject i, for $i=1,\ldots,N$, contributing Γ_i artificial observations $(Y_i,V_i,X_{i1}),(Y_i,V_i,X_{i2}),\ldots,(Y_i,V_i,X_{i\Gamma_i})$, where $X_{ik},k=1,\ldots,\Gamma_i$, denote the regimes followed by subject i. Formally, to follow a regime x means that one's observed treatment $\overline{A}$ is equal to the treatment $\overline{A}_{\overline{d}_x}$ the subject would have had under the regime $\overline{d}_x$. To estimate ψ we fit by weighted least squares with an independence working covariance matrix, the regression model

$$E[Y|X,V]=h(X,V,\psi)=(r(X,V)^{\mathrm{T}},r^*(V)^{\mathrm{T}})\psi$$

DOI: 10.1002/sim

to the artificial data set weighting each artificial observation (Y_i, V_i, X_{ik}) by $\widehat{SW}_{ik} = \widehat{W}_i q^*(X_{ik}, V_i)$, where $\widehat{W} = \prod_{k=0}^{K} 1/f[A(k)|\bar{A}(k-1), \bar{L}(k); \widehat{\alpha}]$ and $q^*(\cdot,\cdot)$ is a user-supplied function. Different choices of q^* impact the efficiency with which ψ (and hence the optimal x) is estimated; however, all choices yield consistent estimators of ψ. [In our earlier discussion, we implicitly chose $q^*(\cdot,\cdot)$ to be the constant 1 so $\widehat{SW}_{ik}$ equalled $\widehat{W}_i$ as previously defined.] The optimal treatment regime in the class $\{\overline{d}_x; x \in \mathscr{X}\}$ for a subject with $V = v$ is estimated as the value of x that maximizes $h(x, v, \widehat{\psi})$ or, equivalently, $h_1(x, v, \widehat{\psi}_1)$ over $x \in \mathscr{X}$ for v fixed.

Orellana *et al.* [2] showed that $\widehat{\psi}$ based on the artificial data set converges to the parameter ψ^* of our dynamic MSM under the assumptions SR, C, P, and correct specification of the model $f[A(k)|\bar{A}(k-1), \bar{L}(k); \alpha]$.

The key step in the proof is to note that when $\widehat{W}$ is converging to $W = W(K)$, the estimate $\widehat{\psi}$ is converging to the (assumed unique) ψ satisfying

$$0 = E\left[\sum_{x \in \mathscr{X}} \Delta_x W q^*(x, V)\{Y - h(x, V, \psi)\}(r(x, V)^{\mathrm{T}}, r^*(V)^{\mathrm{T}})^{\mathrm{T}}\right] \tag{6}$$

where $\Delta_x = \prod_{k=0}^{K} I(A(k) = d_{x,k}[\overline{A}(k-1), \overline{L}(k)])$ is the indicator that the subject followed regime $\overline{d}_x$. However, under our assumptions this expectation is

$$E\left[\sum_{x \in \mathscr{X}} \{Y_{\overline{d}_x} - h(x, V, \psi)\}(r(x, V)^{\mathrm{T}}, r^*(V)^{\mathrm{T}})^{\mathrm{T}} q^*(x, V)\right]$$

which is zero under our MSM. Orellana *et al.* [2] also discuss how to construct locally efficient estimators and doubly robust estimators.

A general MSM is defined analogously. Specifically, given any set of regimes $\{\overline{d}_x; x \in \mathscr{X}\}$ (whether static, dynamic, or both) indexed by x taking values in a (possibly infinite) set $\mathscr{X}$ (not necessarily a subset of the real numbers), a general MSM is defined as model (4) satisfying (5). To estimate the parameter ψ^*, we solve the estimating equations (6) where, of course, Δ_x is now the indicator of following regime $\overline{d}_x$. For example, consider another HIV study, where now subjects may repeatedly start and stop HAART therapy and we consider the regimes 'take therapy at t if and only if the current white blood count exceeds w and a certain liver function test has value less than b'. Here b and w are nonnegative integers in the range of 0–10 000. Then $x = (w, b)$. An example of a choice for $h_1(x, V, \psi_1)$ is $\psi_{1,1}(b-100) + \psi_{1,2}(b-100)^2 + \psi_{1,4}(w-1000) + \psi_{1,5}(w-1000)^2 + \psi_{1,6}(b-100)V + \psi_{1,7}(w-1000)V + \psi_{1,8}(w-1000)(b-100)$. Although in our examples the number of regimes that can be possibly followed by any subject is finite, dynamic MSMs need not assume that this is the case. Orellana *et al.* [2] extend our dynamic MSM methods to the case where $\mathscr{X}$ is uncountable rather than finite.

MSMs and the positivity condition: Specifying a general MSM can also allow us to weaken positivity requirements. For each regime $\overline{d}_{x^*}$ for which the PO assumption might fail or nearly fail (as evidenced by some subject's having particularly small estimated probabilities of following the regime), we simply remove any observation (Y, V, x^*) from the artificial data. We can then either interpret our MSM model as a model for $E[Y_{\overline{d}_x}|V],\ x \in \mathscr{X}_{\mathrm{pos}} = \mathscr{X} \backslash \mathscr{X}_{\mathrm{nonpos}} \subset \mathscr{X}$, where $\mathscr{X}_{\mathrm{nonpos}}$ is the set of x^* removed, or as a model for $E[Y_{\overline{d}_x}|V],\ x \in \mathscr{X}$. In the latter case, one is identifying $E[Y_{\overline{d}_x}|V]$ for the regimes $\overline{d}_x$ in $\mathscr{X}_{\mathrm{nonpos}}$ by model-based extrapolation.

DOI: 10.1002/sim

Semilinear MSMs: MSM (4) satisfying (5) is not guaranteed to be correctly specified under the null that $\psi_1^*=0$. If $\psi_1^*=0, h_1(x,V,\psi_1^*)=0$, and thus the MSM reduces to $E[Y_{\overline{d}_x}|V]=h_2(V,\psi_0^*)$. But the assumed functional form $h_2(V,\psi_0)$ may be incorrect. Furthermore, the IPTW estimates of ψ_0 and ψ_1 are generally correlated. Thus, misspecification of the functional form $h_2(V,\psi_0)$ can result in invalid inferences, even under the null $\psi_1^*=0$. To overcome this difficulty, following Robins [9] we can consider the semilinear general MSM

$$E[Y_{\overline{d}_x}|V]=h_1(x,V,\psi_1^*)+h_2^*(V)$$

with $h_2^*(V)$ allowed to be an arbitrary unknown function, so as to prevent bias in the estimation of $h_1(x,V,\psi_1^*)$ from misspecification of a parametric model for $h_2^*(V)$. Given a user-supplied conditional density $f^*[x|V]$ with $\sum_{x\in\mathscr{X}} f^*[x|V]=1$ and a vector function $b(x,V)$, the estimator $\widetilde{\psi}_1$ that sets to zero the weighted sample average of $f^*(X|V)\{Y-h_1(X,V,\psi_1^*)\}\{b(X,V)-\int b(x,V)\,\mathrm{d}F^*[x|V]\}$ with weight $\widehat{W}$ over the artificial data set Γ can be shown to be a consistent asymptotically normal estimator of ψ_1^* when the model $h_1(x,V,\psi_1^*)$ and the model $f[A(k)|\bar{A}(k-1),\bar{L}(k);\alpha]$ for treatment are correct. Robins [9] proved this result for nondynamic MSMs and Orellana *et al.* [2] showed it for general MSMs. Orellana *et al.* [2] also construct locally efficient, doubly robust estimators of ψ_1^* in semilinear general MSMs. As an example, the density $f^*[x|V]$ could be chosen to be an estimate of the density of X, given V based on the artificial data set.

In fact, when $h_1(x,V,\psi_1)=\psi_1^{\mathrm{T}}r(x,V)$ is linear in ψ_1, it is simple to trick standard weighted least-squares software into computing a consistent and asymptotically linear estimator of ψ_1^*. Specifically, we consider the model $h(x,V;\psi)=\psi_1^{\mathrm{T}}r(x,V)+h_2(V;\psi_0)$ with $h_2(V;\psi_0)=\psi_0^{\mathrm{T}}r^*(V)$, where $r^*(V)=\sum_x r(x,V)f^*(x|V)$. Then, the first component $\widehat{\psi}_1$ of the weighted least-squares estimator $\widehat{\psi}=(\widehat{\psi}_1,\widehat{\psi}_0)$ with weights $\widehat{SW}=f^*[X|V]\widehat{W}$ applied to the artificial data Γ is a consistent and asymptotically normal estimator of ψ_1^* when the model for $f[A(k)|\bar{A}(k-1),\bar{L}(k);\alpha]$ is correct even when the model $h_2(V;\psi_0)=\psi_0^{\mathrm{T}}r^*(V)$ for $h_2^*(V)$ is incorrect.

Remark 1

In Section 3.3, we implicitly chose $f^*[x|V]$ to be the constant $1/300$ (i.e. the uniform distribution on $\mathscr{X}$) for all V's. We thus did not need to mention $f^*[x|V]$ at all because multiplying all weights by the same constant has no effect on a weighted least-squares estimate.

4.4. G-estimation of SNMMs

Before discussing optimal regime SNMMs it will help to review standard additive SNMMs. In general, an additive SNMM assumes that for each treatment time $m=0,\ldots,K$,

$$\begin{aligned}&E[Y_{\{\overline{a}(m-1),a(m),\underline{0}(m+1)\}}|\overline{L}(m)=\overline{l}(m),\overline{A}(m-1)=\overline{a}(m-1)]\\&\quad=E[Y_{\{\overline{a}(m-1),\underline{0}(m)\}}|\overline{L}(m)=\overline{l}(m),\overline{A}(m-1)=\overline{a}(m-1)]+\gamma_m[\overline{a}(m),\overline{l}(m),\beta^*]\end{aligned}\tag{7}$$

where (i) $\{\overline{a}(m-1),a(m),\underline{0}(m+1)\}$ and $\{\overline{a}(m-1),\underline{0}(m)\}$ are nondynamic regimes that differ only in that the former has treatment $a(m)$ at m while the latter has treatment 0 at time m, and both have treatment $\overline{a}(m-1)$ through $m-1$ and treatment 0 from $m+1$ to the end of follow-up K, (ii) β^* is an unknown parameter vector, and (iii) $\gamma_m[\cdot,\cdot,\cdot]$ is a known function of $\overline{a}(m)$, $\overline{l}(m)$, and β satisfying $\gamma_m[\overline{a}(m),\overline{l}(m),\beta]=0$ if $\beta=0$ or $a(m)=0$.

DOI: 10.1002/sim

The functions $\gamma_m[\overline{a}(m), \overline{l}(m), \beta^*]$ model the effect on the mean of Y of a last blip of treatment of magnitude $a(m)$ at m, as a function of (i.e. as modified by) past treatment and covariate history $[\overline{a}(m-1), \overline{l}(m)]$.

Examples of choices of $\gamma_m[\overline{a}(m), \overline{l}(m), \beta]$ include (i) $\beta a(m)$, (ii) $(\beta_0+\beta_1 m)a(m)$, and (iii) $\{\beta_0+\beta_1 m+\beta_2 a(m-1)+\beta_3^{\mathrm{T}} l(m)+\beta_4^{\mathrm{T}} l(m)a(m-1)\}a(m)$. Choice (i) is tantamount to assuming that the effect of a last blip of treatment $a(m)$ is the same for all m's. Under choice (ii) the effect varies linearly with the time m of treatment. Under choice (iii), the effect of a last blip of treatment at m is modified by past treatment and covariate history.

We next describe the g-estimation algorithm for estimating the unknown parameter β^* in an observational study under the assumptions C and SR. To do so, it is convenient to define

$$Y_m(\beta) = Y - \sum_{j=m}^{K} \gamma_j[\overline{A}(j), \overline{L}(j), \beta]$$

Note that, for each β, $Y_m(\beta)$ can be computed from the observed data. To carry out the g-estimation algorithm we first need to postulate a logistic regression model (pooled over persons and time)

$$\text{logit Pr}[A(m)=1|\overline{L}(m), \overline{A}(m-1)] = \alpha^{\mathrm{T}} B(m) \tag{8}$$

for the probability of treatment at time (i.e. week) m for $m=0, \ldots, K$. In this model, $B(m) = b_m[\overline{L}(m), \overline{A}(m-1)]$ is a vector of covariates calculated from a subject's covariate and treatment data $[\overline{L}(m), \overline{A}(m-1)]$, α^{T} is a conformable row vector of unknown parameters. An example of $B(m)$ would be the transpose of the row vector

$$(m, A(m-1), L^{\mathrm{T}}(m), A(m-1)L^{\mathrm{T}}(m), A(m-2), L^{\mathrm{T}}(m-1), L^{\mathrm{T}}(m)A(m-1)A(m-2))$$

where $L(m)$ is the vector of covariates measured at time m.

Having postulated model (8) we estimate α with its MLE $\widehat{\alpha}$. To compute $\widehat{\alpha}$ we simply fit the logistic regression model (8) where each person-week is treated as an independent observation, so that each person contributes $K+1$ observations.

Consider first the case in which β is a scalar parameter as in model (i) above. Suppose that β_{low} and β_{up} are numbers much smaller and larger, respectively, than any substantively plausible value of β^*. Then, to carry out the g-estimation algorithm, separately, for each β on a grid from β_{low} to β_{up}, say $\beta_{\text{low}}, \beta_{\text{low}}+0.1, \beta_{\text{low}}+0.2, \ldots, \beta_{\text{up}}$, we perform the score test of the null hypothesis $\theta=0$ in the logistic model

$$\text{logit Pr}[A(m)=1|\overline{L}(m), \overline{A}(m-1), Y_m(\beta)] = \alpha^{\mathrm{T}} B(m) + \theta Y_m(\beta), \quad m=0, \ldots, K \tag{9}$$

that adds the covariate $Y_m(\beta)$ at each time m to the above pooled logistic model (8). A 95 per cent confidence interval for β^* is the set of β for which the two-sided score test of the hypothesis $\theta=0$ at level 0.05 does not reject. The g-estimate of β^* is the value of β for which the score test takes the value zero (i.e. the p-value is one).

A heuristic justification of the g-estimation procedure is as follows. $Y_m(\beta^*)$ is an estimate of $Y_{\{\overline{a}(m-1), \underline{0}(m)\}}$ since the effect $\sum_{j=m}^{K} \gamma_j[\overline{A}(j), \overline{L}(j), \beta^*]$ of all treatments from m onwards have been subtracted from Y. Thus, one would expect that under the SR assumption, $\theta=0$ when $Y_m(\beta^*)$ is added to (9). In fact, it can be shown that this is true. Now, we do not know β^*. Therefore, any value β for which the data are consistent with the parameter θ of the term $\theta Y_m(\beta)$ being zero might

DOI: 10.1002/sim

be the true β^*, and thus belongs in our confidence interval. If consistency with the data is defined at the 0.05 level, then our confidence interval will have coverage of 95 per cent. Furthermore, the g-estimate $\widehat{\beta}$ of β^* is that β for which adding the term $\theta Y_m(\beta)$ does not help to predict $A(m)$ whatsoever, which is the β for which the score test of $\theta=0$, is precisely zero. The g-estimate $\widehat{\beta}$ is also the value of β for which the pooled logistic regression estimator of θ is precisely zero.

Suppose now that the parameter β is a vector. To be concrete suppose we consider the model with $\gamma_m[\overline{a}(m), \overline{l}(m), \beta] = a(m)\{\beta_0+\beta_1 m+\beta_2 a(m-1)+\beta_3 l(m)+\beta_4 l(m)a(m-1)\}$. In this model, β is five dimensional and we suppose, $l(m)$ is one dimensional. To estimate β we use a five-dimensional grid, one dimension for each component of β. Hence, if we had 20 grid points for each component we would have 20^5 different values of β on our five-dimensional grid. Now, to estimate the five components of β we consider $Q_m = q_m[\bar{L}(m), \bar{A}(m-1)]$, a five-dimensional vector of functions of $\overline{L}(m), \overline{A}(m-1)$, such as $q_m^{\mathrm{T}}[\bar{L}(m), \bar{A}(m-1)] = [1, m, A(m-1), L(m), L(m)A(m-1)]$. We postulate a model that extends model (8) with the addition of the five linear functions $Q_{m,j}Y_m(\beta)$ of $Y_m(\beta)$, $j=1,\ldots,5$, as covariates, i.e.

$$\text{logit Pr}[A(m)=1|\overline{L}(m), \overline{A}(m-1), Y_m(\beta)] = \alpha^{\mathrm{T}} B(m) + \theta^{\mathrm{T}} Q_m Y_m(\beta)$$

Our g-estimate $\widehat{\beta}$ is the value of the vector β for which the 5 degrees of freedom score test statistic for testing the null hypothesis $\theta=0$ is precisely zero. The particular choice of the functions $Q_m = q_m[\bar{L}(m), \bar{A}(m-1)]$ does not affect the consistency of the estimator $\widehat{\beta}$, but it determines the width of its associated confidence interval. See Robins [10] for the optimal choice of Q_m.

When the dimension of β is greater than 2, finding $\widehat{\beta}$ by search over a grid is generally computationally prohibitive. However, when, as in all the examples we have discussed, $\gamma_m[\overline{a}(m), \overline{l}(m), \beta] = a(m)\beta^{\mathrm{T}} R_m$ is linear in β with $R_m = r_m(\bar{L}(m), \bar{A}(m-1))$ being a vector of known functions, then, there is an explicit closed-form expression for $\widehat{\beta}$. Specifically, if $\widehat{\alpha}$ denotes the MLE of α under model (8), then $\widehat{\beta}$ takes the explicit form

$$\widehat{\beta} = \left\{ \sum_{i=1,m=0}^{i=n,m=K} X_{im}(\widehat{\alpha}) Q_{im} S_{im}^{\mathrm{T}} \right\}^{-1} \left\{ \sum_{i=1,m=0}^{i=n,m=K} Y_i X_{im}(\widehat{\alpha}) Q_{im} \right\}$$

where $X_{im}(\widehat{\alpha}) = [A_i(m) - \text{expit}\{\widehat{\alpha}^{\mathrm{T}} B_i(m)\}]$, $S_{im} = \sum_{j=m}^{j=K} A_i(j) R_{ij}$. [Note that to compute $\widehat{\alpha}$ we simply fit the logistic regression model (8) where each person-week is treated as an independent observation, so that each person contributes up to $K+1$ observations.] In fact, for $\gamma_m[\overline{a}(m), \overline{l}(m), \beta]$ of the form $a(m)\beta^{\mathrm{T}} R_m$, we can also obtain a closed-form estimator $\widetilde{\beta}$ of β^*, which is consistent and asymptotically normal if either, but not necessarily both, a model $\varsigma^{\mathrm{T}} T_m = \varsigma^{\mathrm{T}} t_m[\bar{L}(m), \bar{A}(m-1)]$ for $E[Y_m(\beta^*)|\bar{L}(m), \bar{A}(m-1)]$ (and hence for $E[Y_{\{\overline{A}(m-1), \underline{0}(m)\}}|\bar{L}(m), \overline{A}(m-1)]$ because by C and SR the two conditional expectations are the same) is correct or model (8) is correct. The estimator $\widetilde{\beta}$ of β^* is commonly referred to as being doubly robust. The estimator $\widetilde{\beta}$, together with the estimator $\widetilde{\varsigma}$ of ς, is jointly defined as

$$\begin{pmatrix} \widetilde{\beta} \\ \widetilde{\varsigma} \end{pmatrix} = \left\{ \sum_{i=1,m=0}^{i=n,m=K} \begin{pmatrix} X_{im}(\widehat{\alpha}) Q_{im} \\ T_{im} \end{pmatrix} (S_{im}^{\mathrm{T}}, T_{im}^{\mathrm{T}}) \right\}^{-1} \left\{ \sum_{i=1,m=0}^{i=n,m=K} Y_i \begin{pmatrix} X_{im}(\widehat{\alpha}) Q_{im} \\ T_{im} \end{pmatrix} \right\}$$

4.4.1. General SNMMs and optimal regime SNMMs. Suppose we are interested in a particular $\overline{d}^*$ with associated functions $d_m^*[\overline{a}(m-1), \overline{l}(m)]$. Then we define a $\overline{d}^*$-SNMM to be a model for the effect of treatment $a(m)$ *versus* treatment 0 at each time m (as a function of treatment and

DOI: 10.1002/sim

covariate history up to m) when regime $\overline{d}^*$ is followed beginning at time $m+1$. Formally, let $\overline{d}=\{\overline{a}(m-1), a(m), \underline{d}^*(m+1)\}$ be the regime that follows the nondynamic regime $(\overline{a}(m-1), a(m))$ through time (e.g. week) m and then the regime $\underline{d}^*(m+1)=(d^*_{m+1},\ldots,d^*_K)$ from $m+1$. Then a $\overline{d}^*$-SNMM is defined exactly like the SNMM (7), except that $Y_{\{\overline{a}(m-1),a(m),\underline{d}^*(m+1)\}}$ replaces $Y_{\{\overline{a}(m-1),a(m),\underline{0}(m+1)\}}$ and $Y_{\{\overline{a}(m-1),0,\underline{d}^*(m+1)\}}$ replaces $Y_{\{\overline{a}(m-1),\underline{0}(m+1)\}}$. Also we write the known function $\gamma_m[\overline{a}(m),\overline{l}(m),\beta]$ as $\gamma_m^{\overline{d}^*}[\overline{a}(m),\overline{l}(m),\beta]$ to remind us we are now estimating a $\overline{d}^*$-SNMM for a given regime $\overline{d}^*$. Note that a $\overline{d}^*$-SNMM with $\overline{d}^*$ the regime where treatment is always withheld is precisely the SNMM (7).

In the following $\gamma_m^{\overline{d}^*}[c,\overline{a}(m-1),\overline{l}(m),\beta^*]$ is used to denote $\gamma_m^{\overline{d}^*}[\overline{a}(m),\overline{l}(m),\beta^*]$ with the value c substituted for $a(m)$. To estimate the parameter β^* of $\gamma_m^{\overline{d}^*}[\overline{a}(m),\overline{l}(m),\beta^*]$, we use g-estimation as described previously except we redefine $Y_m(\beta)$ to be

$$Y_m(\beta)=Y+\sum_{j=m}^{K}\gamma_j^{\overline{d}^*}[d_j^*\{\overline{A}(j-1),\overline{L}(j)\},\overline{A}(j-1),\overline{L}(j),\beta]-\gamma_j^{\overline{d}^*}[\overline{A}(j),\overline{L}(j),\beta] \tag{10}$$

Intuitively at each time $j\geqslant m$, if $\beta=\beta^*$, $Y_m(\beta)$ subtracts from the subject's observed Y an estimate of the effect $\gamma_j^{\overline{d}^*}[\overline{A}(j),\overline{L}(j),\beta]$ of the subject's observed treatment $A(j)$ and replaces it with an estimate of the effect $\gamma_j^{\overline{d}^*}[d_j^*\{\overline{A}(j-1),\overline{L}(j)\},\overline{A}(j-1),\overline{L}(j),\beta]$ of the treatment $d_j^*\{\overline{A}(j-1),\overline{L}(j)\}$ that the subject would have had at j had, possibly contrary to fact, she began to follow regime $\overline{d}^*$ at time j.

Robins [4] proved that, in the absence of model misspecification, under assumptions SR and C, (i) the g-estimate $\widehat{\beta}$, now based on expression (10), is consistent for the parameter β^* of $\gamma_m^{\overline{d}^*}[\overline{a}(m),\overline{l}(m),\beta^*]$ and (ii) the sample average $n^{-1}\sum_{i=1}^n Y_{0,i}(\widehat{\beta})$ of $Y_0(\widehat{\beta})$ is consistent for $E[Y_{\overline{d}^*}]$. When $\gamma_j^{\overline{d}^*}[\overline{A}(j),\overline{L}(j),\beta]=A(j)\beta^{\mathrm{T}}R_j$, we have the closed-form expression

$$\widehat{\beta}=\left\{\sum_{i=1,m=0}^{i=n,m=K}X_{im}(\widehat{\alpha})Q_{im}S_{im}^{\mathrm{T}}\right\}^{-1}\left\{\sum_{i=1,m=0}^{i=n,m=K}Y_iX_{im}(\widehat{\alpha})Q_{im}\right\}$$

with S_{im}^{T} redefined as $\sum_{j=m}^{j=K}\{A_i(j)-d_j^*\{\overline{A}(j-1),\overline{L}(j)\}\}R_{ij}^{\mathrm{T}}$

Optimal regime SNMMs: A primary use of $\overline{d}^*$-SNMMs is in attempting to estimate the optimal treatment strategy $\overline{d}^{\mathrm{opt}}$ that maximizes $E[Y_{\overline{d}}]$ over all treatment regimes $\overline{d}$. To do so we specify an optimal SNMM, $\overline{d}^{\mathrm{opt}}$-SNMM, based on a function $\gamma_m^{\overline{d}^{\mathrm{opt}}}[\overline{a}(m),\overline{l}(m),\beta]=\gamma_m^{\overline{d}^{\mathrm{opt}}}[a(m),\overline{a}(m-1),\overline{l}(m),\beta]$. As an example we might specify that

$$\begin{aligned}\gamma_m^{\overline{d}^{\mathrm{opt}}}[\overline{a}(m),\overline{l}(m),\beta]&=a(m)\{\beta_0+\beta_1m+\beta_2a(m-1)+\beta_3^{\mathrm{T}}l(m)\\&\quad+\beta_4^{\mathrm{T}}l(m)a(m-1)+\beta_5^{\mathrm{T}}l(m-1)+\beta_6^{\mathrm{T}}l(m-1)a(m-1)\}\end{aligned} \tag{11}$$

If the $\overline{d}^{\mathrm{opt}}$-SNMM were correctly specified and we knew the true β^*, then, under assumptions SR and C, we would know the optimal treatment regime. Specifically, the optimal treatment $d_m^{\mathrm{opt}}[\overline{a}(m-1),\overline{l}(m)]$ at time m given past treatment and covariate history $[\overline{a}(m-1),\overline{l}(m)]$ is $a(m)\in\mathscr{A}(m)$ that maximizes $\gamma_m^{\overline{d}^{\mathrm{opt}}}[a(m),\overline{a}(m-1),\overline{l}(m),\beta^*]$. In particular, if $A(m)$ is a binary indicator $d_m^{\mathrm{opt}}[\overline{a}(m-1),\overline{l}(m)]=I[\gamma_m^{\overline{d}^{\mathrm{opt}}}[1,\overline{a}(m-1),\overline{l}(m),\beta^*]>0]$.

DOI: 10.1002/sim

To understand heuristically why this is the case, suppose that $A(m)$ is binary and that at the very last treatment time K, a subject has past history $\overline{a}(K-1), \overline{l}(K)$. If the subject does not take treatment at K, her outcome will be $Y_{\{\overline{a}(K-1),0\}}$ while if she takes treatment it will be $Y_{\{\overline{a}(K-1),1\}}$. Now, since high values of Y are desirable the optimal treatment choice is to take treatment if and only if $\gamma_K^{\overline{d}^{\text{opt}}}[\overline{a}(K-1), \overline{l}(K), \beta^*]$ exceeds zero. (If $\gamma_K^{\overline{d}^{\text{opt}}}[\overline{a}(K-1), \overline{l}(K), \beta^*]$ is precisely zero, it does not matter whether treatment is taken; in such cases, we choose not to treat simply to break the 'tie'.) Now we continue by backward induction. Specifically, suppose we know the optimal regime from $m+1$ onwards. Consider a subject at time m with past history $\overline{a}(m-1), \overline{l}(m)$. Such a subject will follow the known optimal regime from $m+1$ onwards. But she must decide what treatment to take at m. If she does not take treatment at m, her outcome will be $Y_{g=\{\overline{a}(m-1),0,\underline{d}^{\text{opt}}(m+1)\}}$ while if she takes treatment at m, her outcome will be $Y_{\{\overline{a}(m-1),1,\underline{d}^{\text{opt}}(m+1)\}}$. Thus, according to model (11), she should take treatment if and only if $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{a}(m-1), \overline{l}(m), \beta^*]$ exceeds zero.

Now if we knew $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{a}(m), \overline{l}(m), \beta^*]$ and thus we knew the optimal regime, we would simply have each subject in the population follow the optimal regime beginning at time 0, where at each time m the covariates $L(m)$ must be measured and recorded, so the evolving covariate data necessary to follow the optimal regime will be available.

Since the optimal treatment to give at time m is $d_m^{\text{opt}}[\overline{a}(m-1), \overline{l}(m), \beta^*]$, we see that to estimate the optimal regime we need an estimate $\widehat{\beta}$ of β^*, which we can obtain by g-estimation based on the generalization

$$Y_m(\beta) = Y + \sum_{j=m}^{K} \gamma_j^{\overline{d}^{\text{opt}}}[d_j^{\text{opt}}\{\overline{A}(j-1), \overline{L}(j), \beta\}, \overline{A}(j-1), \overline{L}(j), \beta] - \gamma_j^{\overline{d}^{\text{opt}}}[\overline{A}(j), \overline{L}(j), \beta] \tag{12}$$

of expression (10). Note that (12) differs from (10) in that the regime $d_j^{\text{opt}}\{\overline{A}(j-1), \overline{L}(j), \beta\}$ itself is a function of the parameter β. Nonetheless, one can use g-estimation based on $Y_m(\beta)$ of (12) to estimate $\beta^*$and set confidence intervals by searching over a grid of β values. Furthermore, we can also obtain an estimate $N^{-1}\sum_{i=1}^{N} Y_{0,i}(\widehat{\beta})$ of the mean $E[Y_{\overline{d}^{\text{opt}}}]$ of Y when the population is treated optimally. However, there is no longer an explicit closed-form expression for the g-estimate $\widehat{\beta}$ based on (12), even when $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{A}(m), \overline{L}(m), \beta] = A(m)R_m^{\mathrm{T}}\beta$ is linear in β, with $R_m = r_m\{\overline{A}(m-1), \overline{L}(m)\}$ a known vector function. For example, when $A(m)$ is binary, $d_m^{\text{opt}}[\overline{A}(m-1), \overline{L}(m), \beta] = I[R_m^{\mathrm{T}}\beta > 0]$ and so $Y_m(\beta)$ is no longer linear in β because β now occurs within an indicator function. In fact, when the dimension of β is moderate, the g-estimate $\widehat{\beta}$ is exceedingly difficult to compute by search. However, the following alternative, computationally tractable approach can be used when $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{A}(m-1), \overline{L}(m), \beta]$ is linear in β, i.e. it is of the form $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{A}(m), \overline{L}(m), \beta] = A(m)R_m^{\mathrm{T}}\beta$.

A closed-form estimator of the optimal regime: We only consider the case of binary $A(m)$. Suppose for the moment that $\gamma_m^{\overline{d}^{\text{opt}}}[\overline{A}(m), \overline{L}(m), \beta]$ is of the form $A(m)R_m^{\mathrm{T}}\beta_m$, that is, it depends linearly on a separate parameter vector β_m at each time m and we assume that the components of $\beta^{\mathrm{T}} = (\beta_0^{\mathrm{T}}, \ldots, \beta_K^{\mathrm{T}})$ are variation independent. We will describe estimators $\widetilde{\beta}_m$ of β_m^* that are doubly robust in the sense of being consistent and asymptotically normal for β_m^* if either model (8) is correct for all m's or if model

$$E[Y_m(\beta^*)|\bar{L}(m), \bar{A}(m-1)] = \zeta_m^{\mathrm{T}} z_m[\bar{L}(m), \bar{A}(m-1)]$$

DOI: 10.1002/sim

where $Z_m=[\bar{L}(m),\bar{A}(m-1)]$ is a known function of $(\bar{L}(m),\bar{A}(m-1))$ and ς_m is unknown, is correct for all m's. To compute the estimators $\widetilde{\beta}_m$ we first compute the partial MLE $\widehat{\alpha}$ of α under model (8). Next, beginning with $m=K$, we recursively compute $(\widetilde{\beta}_m,\widetilde{\eta}_m,\widetilde{\varsigma}_m)$, the OLS estimator of $(\beta_m,\eta_m,\varsigma_m)$ in the regression model

$$Y_{m+1}(\underline{\widetilde{\beta}}_{m+1})=A(m)R_m^{\mathrm{T}}\beta_m+\mathrm{expit}\{\widehat{\alpha}^{\mathrm{T}}B(k)\}R_m^{\mathrm{T}}\eta_m+Z_m^{\mathrm{T}}\varsigma_m+\varepsilon_{m+1}$$

where $Y_{K+1}(\underline{\widetilde{\beta}}_{K+1})=Y$, $Y_{m+1}(\underline{\widetilde{\beta}}_{m+1})=Y+\sum_{j=m+1}^{K}[I\{R_j^{\mathrm{T}}\widetilde{\beta}_j>0\}-A(j)]R_j^{\mathrm{T}}\widetilde{\beta}_j$ for $0\leqslant m\leqslant K$, and ε_{m+1} is the unobserved error term.

Remark

The estimators $\widetilde{\beta}_m$ are possibly inefficient members of a general class of estimators $\widetilde{\beta}_m(\mathbf{s},\mathbf{q})$, indexed by vectors of functions s_m and $q_m, m=0,\ldots,K$, defined recursively as follows:

$$\widetilde{\beta}_m(\mathbf{s},\mathbf{q})=\left[\sum_{i=1}^{i=n}A_i(m)X_{im}(\widehat{\alpha})Q_{im}R_{im}^{\mathrm{T}}\right]^{-1}\left[\sum_{i=1}^{i=n}\{Y_{m+1}(\underline{\widetilde{\beta}}_{m+1}(\mathbf{s},\mathbf{q}))_j-S_{im}\}X_{im}(\widehat{\alpha})Q_{im}\right]$$

where $X_m(\widehat{\alpha})=[A(m)-\mathrm{expit}\{\widehat{\alpha}^{\mathrm{T}}B(k)\}]B_m$ and $S_m=s_m[\bar{L}(m),\bar{A}(m-1)]$ and $Q_m=q_m[\bar{L}(m),\bar{A}(m-1)]$ are user-specified functions, their choice affecting efficiency but not consistency when the treatment model (8) is correct.

Now suppose in our d^{opt}-SNMM model, the same parameter vector β applies to each time m. To be concrete, consider the d^{opt}-SNMM (11). In that case, we first estimate a bigger model that has a separate variation-independent parameter vector β_m at each time m; model (11) is then the submodel that imposes $\beta_m=\beta$ for all m's. Let $\widetilde{\Omega}^{-1}$ be a nonparametric bootstrap estimate of the covariance matrix of $(\widetilde{\beta}_0,\ldots,\widetilde{\beta}_K)$. We then estimate β by an inverse covariance-weighted average $\widehat{\beta}=1_{K+1}^{\mathrm{T}}\widetilde{\Omega}^{-1}(\widetilde{\beta}_0,\ldots,\widetilde{\beta}_K)^{\mathrm{T}}/(1_{K+1}^{\mathrm{T}}\widetilde{\Omega}^{-1}1_{K+1})$ of the $\widetilde{\beta}_m$, where 1_{K+1} is a $K+1$ vector with all components equal to 1.

Note that the d^{opt}-SNMM model (11) is a nonsaturated model. For example, it assumes that the optimal regime does not depend on covariate values two weeks in the past or treatment values three weeks in the past, which may be incorrect. If the d^{opt}-SNM model is badly misspecified, then the estimated optimal regime $\widehat{d}_m^{\mathrm{opt}}[\bar{a}(m-1),\bar{l}(m)]$ may be a poor estimate of the actual optimal regime. Because in realistic studies highly nonsaturated d^{opt}-SNMM models must be employed, misspecification can be a serious problem.

We note that in using a d^{opt}-SNMM model to find the optimal regime, it was necessary for us to estimate the treatment strategy $d^{\mathrm{opt}}=\{d_0^{\mathrm{opt}}[\bar{l}(0)],d_1^{\mathrm{opt}}[a(0),\bar{l}(1)],\ldots,d_K^{\mathrm{opt}}[\bar{a}(K-1),\bar{l}(K)]\}$ that maximized $E_d[Y]$ over all regimes d, including regimes in which treatment depends on past treatment as well as covariate history. However, one can always construct a regime $d^{\dagger\mathrm{opt}}=\{d_0^{\dagger\mathrm{opt}}[\bar{l}(0)],d_1^{\dagger\mathrm{opt}}[\bar{l}(1)],\ldots,d_K^{\dagger\mathrm{opt}}[\bar{l}(K)]\}$ in which treatment depends only on past covariate history such that following regime $d^{\dagger\mathrm{opt}}$ from time 0 onwards is precisely equivalent to following d^{opt} from time 0. Nonetheless, it can be important to know d^{opt} rather than only $d^{\dagger\mathrm{opt}}$ as the following example shows. Suppose that a (random) member of the source population has observed history $(\overline{A}(m-1),\overline{L}(m))=(\bar{a}(m-1),\bar{l}(m))$ (under standard care) that is not consistent with following the optimal regime d^{opt} and comes to our attention only at time m. We wish to intervene beginning at m and give the subject the optimal treatment strategy from time m onwards. Under the C and SR assumptions,

DOI: 10.1002/sim

the optimal treatment strategy for such a subject is $\{d_m^{\text{opt}}[a(m-1), \bar{l}(m)], \ldots, d_K^{\text{opt}}[\bar{a}(K-1), \bar{l}(K)]\}$. This strategy can be implemented only if we know (or have a good estimate of) d^{opt}; knowledge of $d^{\dagger\text{opt}}$ does not suffice.

5. STRENGTHS AND WEAKNESSES OF MSMs AND SNMMs

MSMs have the advantage that they are easy to understand and easy to fit with standard off-the-shelf software that allows for weights. These two points explain their rapid adoption compared with SNMMs. The usefulness of MSMs has been extended by the introduction of dynamic MSMs to estimate the optimal regime in a class.

However, IPTW estimation of MSMs has four drawbacks not shared by g-estimation of SNMMs. First, if the number of time periods is large the product in the denominator of the weights can become very small for some subjects who then receive inordinately large weights, leading both to bias when the weights must be estimated [and to so-called pseudo-bias even when they are known, see Scharfstein *et al.* [11]] and to imprecision. Problems with large or even truly infinite weights (when positivity does not hold) can be somewhat ameliorated but not cured, by using bounded doubly robust estimators [12], adjusting for baseline covariates and then using the covariates in the numerator of the weights, downweighting or eliminating from consideration regimes $g = d_x$ associated with very small weights, using locally semiparametric efficient estimators or bounded influence function estimators for nonsaturated MSMs (as these estimators downweight regimes $g = d_x$ that result in excessively large weights in a near optimal fashion), and using diagnostics for the undue influence of large weights and for the consequences of truncating large weights [13]. Second, MSMs cannot be used to estimate causal effects when treatment is confounded but an instrumental variable is available. Third, sensitivity analysis models to assess the impact of departures from the assumption of SR are much more restrictive for MSMs and less useful than those for SNMMs. Fourth, SNMMs, in contrast to MSMs, allow one to directly model interactions between treatment and evolving time-dependent covariates in order to look for qualitative effect modification.

In terms of estimation of optimal regimes, both MSMs and SNMMs have their distinct place. General MSMs are excellent for estimating the optimal regime in a prespecified parametrized class of regimes (such as the optimal CD4 cell count at which to start therapy) that still may include all logistically feasible regimes, particularly in settings with resource constraints that preclude implementing complex regimes.

In contrast, the method of backward induction on which g-estimation of optimal SNMMs is based requires that the set of potential regimes from which the optimal is to be selected include all functions of an increasing (in time) amount of information (i.e. of an increasing sigma field). Thus, optimal regime SNMMs are useful for estimating the optimal regime in the huge class of dynamic regimes in which treatment at each m can depend on any function of the entire measured past $\bar{l}(m), \bar{a}(m-1)$ (the case considered above) or, as described in Section 7 of Robins [4] in the smaller, but still large, class in which treatment at each m can depend on any function of $\overline{w}(m), \bar{a}(m-1)$, where $W(m)$ is a subvector of the covariates in $L(m)$. Even if $W(m)$ is just CD4 cell count at m, it is possible that the optimal treatment decision at time m may be a complex function of CD4 cell counts at all past times. Such a regime, though optimal in its class, may be logistically impossible to implement, in which case it may be necessary to choose among a smaller class of logistically feasible regimes by fitting a general MSM.

DOI: 10.1002/sim

6. BETWEEN POPULATION EXTRAPOLATION AND OPTIMIZATION OF PROGNOSTIC TESTING STRATEGIES

Let $\overline{d}_x$ again denote the nonrandom dynamic regime 'begin HAART the first time t the measured CD4 count falls below x', where x is measured in whole numbers. Let $\mathscr{X}=\{201,\ldots,500\}$ and let $\{\overline{d}_x; x\in\mathscr{X}\}$ be the set of candidate regimes. Let x_{opt} be the $x\in\mathscr{X}$ for which the expected utility $E_{\overline{d}_x}[Y]$ is a maximum. It is clear that x_{opt} will depend on the frequency with which subjects have their *CD*4 count measured. For example, suppose that in a particular patient population, *CD*4 counts are obtained very frequently, say weekly, and x_{opt} is 340. If in the same population, the CD4 counts were measured much less frequently (say every 6 months), then x_{opt} would presumably exceed 340, as many subjects who were well above 340 at the time of their last CD4 blood test but were not started on HAART may be well below 340 at the time of their next test 6 months later.

These considerations raise at least two issues. First, suppose the empirical frequency of blood tests to assess CD4 counts differs between two biologically similar populations but good longitudinal data are available only on the first population. How can we use data from the first population to estimate x_{opt} in the second? This extrapolation question presumes that, for logistical or financial reasons, we do not have the ability to change the frequency of *CD*4 blood tests in the second population. Thus, although, for purposes of continuity, we use a CD4 blood test as a prototype, the methods we develop may be more relevant to tests that are either more expensive or present a greater logistical challenge than a CD4 count.

Second, suppose that in a particular population on which we have collected longitudinal data we do have the ability to increase at some financial cost, or decrease with some financial savings, the frequency of CD4 tests and we desire to use the data to determine jointly the optimal CD4 testing schedule and the optimal time to start HAART. This second issue is the well-known issue of assessing the 'value of information' in sequential decision making; a CD4 test has no direct biological effect on a patient whatsoever. Its net value, if any, is that the information supplied by the test can be used to fine tune the time to start HAART, leading to a net increase in expected utility, even when the costs of the test are included in the utility function. Of course, to include the costs of a CD4 test in the utility implies that our utility function is in dollars and that we must place a monetary value, usually adjusted for quality of life, on each additional year of life. It is not our purpose to discuss this highly contentious issue. Our goal here is limited to the development of statistical methodology.

6.1. Formalization

We formalize the problem as follows. The potentially observed data at time m are now $L(m)=(Z(m)^{\mathrm{T}}, I(m))^{\mathrm{T}}$, $A(m)$, and $T(m)$, where $T(m)$ is the indicator of whether a test was performed at time m, $I(m)$ is the biological function measured by a test (e.g. CD4 count) at $m-1$, $Z(m)$ are the other components of the covariate vector $L(m)$, and $A(m)$ is the treatment at m. Let $O(m)=(Z(m), I(m), A(m), T(m))$, which we assume is temporally ordered in the sense that a decision whether to test at m can depend on $Z(m), I(m), A(m)$ and the decision on treatment can potentially depend on $(Z(m), I(m))$. However, we observe $I(m)$ only if $T(m-1)=1$ while $Z(m)$ and $A(m)$ are always observed. [The fact that the test at $m-1$ gives us knowledge of I at m rather than $m-1$ is inconsequential and can be taken as a convention that simplifies the notation.] Thus, the actually observed data at time m are $O^*(m)=(Z(m), I^*(m), A(m), T(m))$,

DOI: 10.1002/sim

where $I^*(m)=T(m-1)I(m)$, i.e. $I^*(m)$ is equal to the biological function at time $m-1$ if the test was actually performed, and it is equal to 0 otherwise. The vector $L^*(m)\equiv(Z(m),I^*(m))$ encodes all new information about the patient's health status that the physician has available at time m. In contrast, the vector $J(m)\equiv(A(m),T(m))$ encodes the pair of intervention variables whose values depend on the doctor's decision at time m (i.e. which treatment to assign at time m and whether or not to take a test at time m). Throughout, we let $\overline{\mathscr{J}}$ denote the sample space of $\overline{J}(K)$, where we continue to use our convention for overbars, i.e. $\overline{J}(m)=(J(0),\ldots,J(m))$.

For later use, it will be convenient to define the vector $\overline{O}^*(k)=(O^*(0),\ldots,O^*(k))$ of all the data observed up to time k, and the vectors $\overline{O}^{*t}(k)=(\overline{O}^*(k-1),Z(k),I^*(k),A(k))$, $\overline{O}^{*,a}(k)=(\overline{O}^*(k-1),Z(k),I^*(k))$, and $\overline{O}^{*,i}(k)=(\overline{O}^*(k-1),Z(k))$, which are equal to $\overline{O}^*(k)$ but with $T(k)$, $(A(k),T(k))$, and $(I^*(k),A(k),T(k))$ removed.

The counterfactual L-history $\overline{L}_{\overline{j}}(m)=\{L(0),L_{a(0),t(0)}(1),\ldots,L_{\overline{a}(m-1),\overline{t}(m-1)}(m)\}$ through m under the joint treatment and testing regime $\overline{j}=(\overline{a},\overline{t})$ is assumed to satisfy the following assumption. To state this assumption, let $\overline{L}_{\overline{a}}(m)=\{L(0),L_{a(0)}(1),\ldots,L_{\overline{a}(m-1)}(m)\}$.

No direct effect (*NDE*) *assumption*: $\overline{L}_{\overline{a},\overline{t}}(m)=\overline{L}_{\overline{a}}(m)$ w.p.1. for all m's.

The NDE assumption encodes the fact that testing *per se* has no direct effect on $L(k)$ when treatment is set to $\overline{a}$. The assumption implies that $\overline{L}_{\overline{a},\overline{t}}=\overline{L}_{\overline{a}}\equiv L_{\bar{a}(K)}(K+1)$ and consequently the set $\overline{L}_{\overline{\mathscr{J}}}\equiv\{\overline{L}_{\overline{j}};\overline{j}\in\overline{\mathscr{J}}\}$, denoting all vectors of counterfactual health status variables in the hypothetical scenarios in which the intervention variables are set to each of their possible configurations $\overline{j}=(\overline{a},\overline{t})$, is equal to the set $\overline{L}_{\overline{\mathscr{A}}}\equiv\{\overline{L}_{\overline{a}};\overline{a}\in\overline{\mathscr{A}}\}$ of all counterfactuals obtained when setting only the treatment variables.

We also define the counterfactual vector $\overline{L}^*_{\overline{a},\overline{t}}(m)\equiv\overline{L}^*_{\overline{j}}(m)=\{L(0),L^*_{j(0)}(1),\ldots,L^*_{\overline{j}(m-1)}(m)\}$ with $L^*_{\overline{j}(m-1)}(m)=(Z_{\overline{j}(m-1)}(m),I^*_{\overline{j}(m-1)}(m))$ and $I^*_{\overline{j}(m-1)}(m)=I_{\overline{j}(m-1)}(m)t(m-1)$. The vector $\overline{L}^*_{\overline{j}(m-1)}(m)$ encodes all health outcomes that would actually be recorded up to time m if treatment $\overline{A}(m-1)=\overline{a}(m-1)$ and test decisions $\overline{T}(m-1)=\overline{t}(m-1)$ were implemented. Under the NDE assumption $I^*_{\overline{j}(m-1)}(m)=I_{\overline{a}(m-1)}(m)t(m-1)$ but $I_{\overline{j}(m-1)}(m)=I_{\overline{a}(m-1)}(m)$; hence, in contrast to $\overline{I}_{\overline{j}}(m)=\overline{I}_{\overline{j}(m-1)}(m)$, $\overline{I}^*_{\overline{j}}(m)=\overline{I}^*_{\overline{j}(m-1)}(m)$ depends on $\overline{t}$ even under the NDE assumption.

For notational convenience, we also define $\overline{I}^*_{\overline{t}(m-1)}(m)=\overline{I}^*_{\overline{A}(m-1),\overline{t}(m-1)}(m)$ and $\overline{L}^*_{\overline{t}(m-1)}(m)=\overline{L}^*_{\overline{A}(m-1),\overline{t}(m-1)}(m)$ so $\overline{L}^*_{\overline{t}(m-1)}(m)$ is a subject's counterfactual L^*-history through m with treatment history equal to the observed history $\overline{A}(m-1)$ but testing history equal to $\overline{t}(m-1)$.

As in the preceding sections, we again make a consistency assumption (C), which links a specific vector of the counterfactual set $\overline{L}_{\overline{\mathscr{J}}}$ to the observed data. Specifically, we assume

Consistency assumption (*C*): $\overline{I}^*(m)=\overline{I}^*_{\overline{J}(m-1)}(m)\equiv\overline{I}^*_{\overline{T}(m-1)}(m)$ and $\overline{L}(m)=\overline{L}_{\overline{A}(m-1)}(m)$ for $m=0,1,\ldots,K$.

A random dynamic j-regime is a conditional law G for $\overline{J}$ given $\overline{L}_{\overline{\mathscr{A}}}$ with density denoted as $g(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})$. We denote the G generating the data in our study as $G^\intercal$ and its density as $g^\intercal(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})$. We let P denote a (generic) marginal distribution of $\overline{L}_{\overline{\mathscr{A}}}$ and $p(\overline{l}_{\overline{\mathscr{A}}})$ its density evaluated at $\overline{l}_{\overline{\mathscr{A}}}$. We let $P^\intercal$ denote the actual P that generates the (counterfactual) data $\overline{L}_{\overline{\mathscr{A}}}$ in the population under study.

Let $F_{(G,P)}$ and $f_{(G,P)}(\overline{l}_{\overline{\mathscr{A}}},\overline{j})$ denote the distribution and density of $(\overline{L}_{\overline{\mathscr{A}}},\overline{J})$ under the joint law (G,P). Thus, $f_{(G,P)}(\overline{l}_{\overline{\mathscr{A}}},\overline{j})=p(\overline{l}_{\overline{\mathscr{A}}})g(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})$. Throughout, in a slight abuse of notation, we use the same symbol, say g, to denote different conditional densities associated with the distribution G,

DOI: 10.1002/sim

the actual random variables connected to the conditional densities being those that take values that appear in the expression for g. Thus, for example, $g[j(k)|\overline{o}^{*,a}(k)]$ stands for the conditional density of $J(k)$ (evaluated at $j(k)$) given $\overline{O}^{*,a}(k)=\overline{o}^{*,a}(k)$ while $g(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})$ stands for the conditional density of $\overline{J}$ (evaluated at $\overline{j}$) given $\overline{L}_{\overline{\mathscr{A}}}=\overline{l}_{\overline{\mathscr{A}}}$.

We say G is sequentially randomized given data $\overline{O}^{*,a}$ ($\overline{O}^{*,a}$-SR) if for all $\overline{j}$, $\overline{l}_{\overline{\mathscr{A}}}$

$$g(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})=\prod_{k=0}^{K} g[j(k)|\overline{o}^{*,a}(k)]$$

The density $f_{(G^\tau,P^\tau)}(\overline{l}_{\mathscr{A}},\overline{j})$ is said to satisfy coarsening at random (CAR) with respect to the underlying full data $\overline{L}_{\overline{\mathscr{A}}}$ and observed data $\overline{O}^*=(\overline{L}^*,\overline{J})$ if $g^\tau(\overline{j}|\overline{l}_{\overline{\mathscr{A}}})$ depends on $\overline{o}^*$only. It follows that if G^τ is $\overline{O}^{*,a}$-SR, then $f_{(G^\tau,P^\tau)}(\overline{l}_{\mathscr{A}},\overline{j})$ satisfies CAR.

As with dynamic treatment regimes, we can also define nonrandom j-regimes that are defined by functions $\overline{b}=(b_0,\ldots,b_K)$, so that at each time k, a subject with observed past $\overline{O}^{*,a}(k)=\overline{o}^{*,a}(k)$ is assigned to treatment and testing status $J(k)=(A(k),T(k))$ equal to $b_k[\overline{o}^{*,a}(k)]\in\mathscr{J}(k)$. Any nonrandom j-regime $\overline{b}=(b_0,\ldots,b_K)$ is indeed a special case of a $\overline{O}^{*,a}$-SR random regime in which, for each k, $g_k[j(k)|\overline{o}^{*,a}(k)]$ is zero for all $j(k)\in\mathscr{J}(k)$, except for $j(k)=b_k[\overline{o}^{*,a}(k)]$ for which it is one. In the next theorem, we shall also need the following generalization of the earlier assumption PO.

Joint positivity assumption (PO^*) *for a regime* G^τ *with respect to another regime* G: $g(\overline{j}|\overline{L}_{\overline{\mathscr{J}}})$ is dominated by (i.e. absolutely continuous w.r.t.) $g^\top(\overline{j}|\overline{L}_{\overline{\mathscr{J}}})$ with $P^\top$-probability 1, i.e.

$$g(\overline{j}|\overline{L}_{\overline{\mathscr{J}}})>0 \text{ implies } g^\top(\overline{j}|\overline{L}_{\overline{\mathscr{J}}})>0 \text{ w.p.1. under } P^\top \tag{13}$$

For $\overline{O}^{*,a}$-SR regimes G^τ and G, the PO* assumption for G^τ w.r.t. G is equivalent to the assumption that, for all $k\leqslant K$, $\Pr_{(G,P^\tau)}[g^\tau\{J(k)|\overline{O}^{*,a}(k)\}>0]=1$. That is, if under a j-regime G, a treatment assignment and test decision were feasible at time k, i.e. had positive probability of occurring at k, for a given observed past $\overline{O}^{*,a}(k)$, then they would also have positive probability of occurring in the actual population from where the observed data were sampled.

The following theorem establishes a key result connecting the mean of any function $h(\overline{O})$ of $\overline{O}$ in the hypothetical world in which the j-regime G is implemented to the mean of $h(\overline{O})$ in the actual study under the joint positivity assumption PO*. In the following subsection, we exploit the results of this theorem to develop between-population extrapolation and estimation of joint optimal treatment assignment and testing decision strategies.

Theorem 6.1
Suppose assumption C holds and the $\overline{O}^{*,a}$-SR holds for G^τ and G. Then,

$$E_{(G,P^\tau)}[h(\overline{O})]=E_{(G^\tau,P^\tau)}\left[\frac{\prod_{k=0}^{K} g[J(k)|\overline{O}^{*,a}(k)]}{\prod_{k=0}^{K} g^\tau[J(k)|\overline{O}^{*,a}(k)]}h(\overline{O})\right]$$

for all $h(\overline{O})$ for which $E_{(G,P^\tau)}[|h(\overline{O})|]<\infty$ if and only if assumption PO* for G^τ with respect to G holds.

Theorem 6.1 is just the Radon-Nikodym theorem applied to the laws (G,P^τ) and (G^τ,P^τ). Note that under assumptions PO*,C, and $\overline{O}^{*,a}$-SR for G^τ and G, $\prod_{k=0}^{K} g[J(k)|\overline{O}^{*,a}(k)]/\prod_{k=0}^{K} g^\tau[J(k)$

DOI: 10.1002/sim

$|\overline{O}^{*,a}(k)]$ is indeed the likelihood ratio (Radon-Nikodym derivative) of $(G, P^{\intercal})$ with respect to $(G^{\intercal}, P^{\intercal})$.

6.2. *Between-population extrapolation and optimal strategies*

In this subsection, we assume that assumption PO* for $G^{\intercal}$ with respect to G holds. Further, as shorthand we refer to the study population on whom we have obtained data as population $G^{\intercal}$.

6.2.1. Between-population extrapolation. In this section, we will apply Theorem 6.1 to derive estimators of the optimal treatment strategy in a population, throughout referred to as population G, which follows the testing regime with conditional probabilities $g[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)]$. The estimators are based on data from the study population $G^{\intercal}$, which follows a possibly different testing regime with conditional probabilities $g^{\intercal}[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)]$. In this section, we only wish to optimize the treatment regime, taking the testing regime g as given. In the following section, we shall consider jointly optimizing treatment and testing regimes. We will assume that the $\overline{O}^{*,a}$-SR assumption holds in both populations $G^{\intercal}$ and G. Furthermore, we assume that for each k, the conditional probability $g[t(k)=1|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)]$ of performing a test at k (with result $I(k+1)$) in population G given the observed data $(\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k))$ is known. In practice, the testing schedule will not be known for population G and our specification of g will reflect a best guess, which we nevertheless treat as known in the following analysis. However, in practice, the selected g could be varied in a sensitivity analysis or be replaced by an empirical estimate if appropriate data were available on population G. We assume that populations G and $G^{\intercal}$ have the same marginal distribution $P^{\intercal}$ of $\overline{L}_{\mathscr{A}}$. In certain settings, this latter assumption can be understood as a formalization of the assumption that the two populations are biologically similar.

Suppose that we are given a set of nonrandom treatment regimes $\{\overline{d}_x, x \in \mathscr{X}\}$ to be implemented in a setting in which testing for $I(k)$ is not necessarily done at each time k. In this setting, $\overline{d}_x$ is equal to a collection of functions $\{d_{x,k}; k=0, \ldots, K\}$ where each $d_{x,k}$ maps the data $(\overline{A}(k-1), \overline{T}(k-1), \overline{L}^*(k))$ to a treatment $a(k)$, that is, $d_{x,k}[\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*(k)] \in \mathscr{A}(k)$.

The collection of conditional densities $\{g[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)], k=0, \ldots, K\}$, together with a given treatment regime $\overline{d}_x$, determines an $\overline{O}^{*,a}$-SR random dynamic j-regime G_x in population G. Specifically, define the conditional densities as

$$g_x[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*(k)] = \begin{cases} 1 & \text{if } a(k) = d_{x,k}[\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*(k)] \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$$g_x[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)] = g[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)] \tag{15}$$

By definition, $g_x[j(k)|\overline{o}^{*,a}(k)] = g_x[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*(k)] g_x[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)]$ and, by the $\overline{O}^{*,a}$-SR assumption, $g_x(\overline{j}|\overline{l}_{\mathscr{A}}) = \prod_{k=0}^{K} g_x[t(k)|a(k), \overline{o}^{*,a}(k)] g_x[a(k)|\overline{o}^{*,a}(k)]$. Note, since $d_{x,k}[\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*(k)] = d_{x,k}[\overline{o}^{*,a}(k)]$ is a deterministic treatment regime, $g_x[a(k)|\overline{o}^{*,a}(k)] = g_x[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}_{\mathscr{A}}]$ automatically holds and thus is not an assumption. Thus, for regime g_x, the $\overline{O}^{*,a}$-SR assumption reduces to the assumption that $g_x[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}_{\mathscr{A}}] = g_x[t(k)|\overline{a}(k), \overline{t}(k-1), \overline{l}^*(k)]$. To be concrete we will consider the following running example.

Running example: Let $\overline{d}_x$ again denote the nonrandom dynamic regime 'begin HAART the first time k the measured CD4 count falls below x,' with $x \in \mathscr{X} = \{201, \ldots, 500\}$. In order to

DOI: 10.1002/sim

avoid additional complications, it will be useful to assume that, under both $g^{\top}$ and g_x, a subject who initiates HAART remains on the same HAART regimen thereafter and takes no other HIV medications, in which case there is no reason to obtain a CD4 test once HAART has been started. We formalize this assumption by assuming

$$\begin{aligned} \mathrm{pr}_{G^*}[A(k)=1|\overline{a}(k-1),\overline{t}(k-1),\overline{l}^*(k)]=1 \quad &\text{if } a(k-1)s(k)=1 \\ \mathrm{pr}_{G^*}[T(k)=1|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]=0 \quad &\text{if } a(k)=1 \text{ or } s(k)=0 \end{aligned} \tag{16}$$

hold for G^* equal to $G^{\top}$ or G_x where $s(k)$ is a realization of $S(k)\in L^*(k)$, the indicator variable of survival at week k. [Without this assumption and with $A(t)$ being the indicator variable that takes the value 1 if and only if HAART has been initiated by time t, the NDE assumption would be unlikely to hold, because, the result of a CD4 test might then have a direct effect on survival not through $A(t)$ if the test result is used to decide when to switch an initial HAART regimen to a new HAART regimen. In studies in which subjects do switch treatment regimens based on CD4 tests obtained after HAART initiation, the NDE assumption would only hold if we redefined $A(t)$ to be the vector of all treatments that physicians (or patients) might modify at time t based on the result of a CD4 test.]

Let $Y=y(\overline{A},\overline{L})$ be a prespecified utility function. The mean $E_{(G_x,P^{\intercal})}(Y)$ computed under the joint law $(G_x,P^{\intercal})$ is interpreted as the expected utility in population G if the treatment strategy $\overline{d}_x$ is followed. Similarly, if $G_x^{\intercal}$ is defined like G_x but with $g^{\intercal}$ instead of g, the resulting mean $E_{(G_x^{\intercal},P^{\intercal})}(Y)$ is the expected utility in the study population $G^{\intercal}$ if the treatment regime $\overline{d}_x$ is followed. Note that this expected utility was denoted as $E(Y_{\overline{d}_x})$ in Section 4.

Our goal is to find $x_{\mathrm{opt}}(V)$ in $\mathscr{X}$ that maximizes $E_{(G_x,P^{\intercal})}[Y|V]$ over $x\in\mathscr{X}$, where V is a (possibly improper) subset of the baseline covariates. The key observation that will allow us to estimate $x_{\mathrm{opt}}(V)$ is that under the conditions of Theorem 6.1 (which, in this subsection, we assume to hold) and the PO* assumption we have

$$E_{(G_x,P^{\intercal})}[Y|V]=E_{(G^{\intercal},P^{\intercal})}[U_xY|V] \tag{17}$$

where

$$U_x\equiv\frac{\prod_{k=0}^{K}g_x[J(k)|\overline{O}^{*,a}(k)]}{\prod_{k=0}^{K}g^{\intercal}[J(k)|\overline{O}^{*,a}(k)]}=\frac{\Delta_x\prod_{k=0}^{K}g_x[T(k)|\overline{O}^{*,t}(k)]}{\prod_{k=0}^{K}g^{\intercal}[J(k)|\overline{O}^{*,a}(k)]} \tag{18}$$

and

$$\Delta_x=\prod_{k=0}^{K}I\left(A(k)=d_{x,k}[\overline{A}(k-1),\overline{T}(k-1),\overline{L}^*(k)]\right)$$

is the indicator that the subject followed treatment regime $\overline{d}_x$.

Suppose, as in Section 4.3, we specify an MSM model

$$E_{(G_x,P^{\intercal})}[Y|V]=h(x,V,\psi^*) \tag{19}$$

with $h(x,V,\psi)=(r(x,V)^{\mathrm{T}},r^*(V)^{\mathrm{T}})\psi$. Then, it follows from (17) that, for any user-specified function $q^*(x,V)$,

$$E_{(G^{\intercal},P^{\intercal})}\left[\sum_{x\in\mathscr{X}}U_x\{Y-h(x,V,\psi^*)\}(r(x,V)^{\mathrm{T}},r^*(V)^{\mathrm{T}})^{\mathrm{T}}q^*(x,V)\right]=0 \tag{20}$$

DOI: 10.1002/sim

Consequently, letting $\mathbb{P}_n(A)$ stand for $n^{-1}\sum_{i=1}^n A_i$, we have that under standard regularity conditions, a CAN estimator of ψ^* is the $\widehat{\psi}$ solving

$$\mathbb{P}_n\left[\sum_{x\in\mathscr{X}} \widehat{U}_x\{Y-h(x,V,\psi)\}(r(x,V)^{\mathrm{T}}, r^*(V)^{\mathrm{T}})^{\mathrm{T}} q^*(x,V)\right]=0 \tag{21}$$

where

$$\widehat{U}_x \equiv \frac{\Delta_x \prod_{k=0}^K g[T(k)|\overline{O}^{*,t}(k)]}{\prod_{k=0}^K g^{\tau}[J(k)|\overline{O}^{*,a}(k);\widehat{\alpha}]} \tag{22}$$

provided (i) the model $g^{\mathrm{T}}[J(k)|\overline{O}^{*,a}(k);\alpha]$ used to estimate $g^{\tau}[J(k)|\overline{O}^{*,a}(k)]$ is correct and $\widehat{\alpha}$ maximizes the partial log-likelihood $\mathbb{P}_n[\sum_{k=0}^K \log g^{\tau}[J(k)|\overline{O}^{*,a}(k);\alpha]]$, (ii) $\widehat{U}_x Y$ is a function of the observed data $\overline{O}^*$, and (iii) the PO* assumption holds for all $x\in\mathscr{X}$. In practice, as discussed in Section 4.3, to ensure (iii), we should also eliminate from $\mathscr{X}$ any x for which there exists empirical evidence that the PO*assumption almost fails. Note a sufficient condition for (ii) to hold is that the utility function $y(\overline{A},\overline{L})$ does not depend on a subject's possibly unmeasured covariate history $\overline{I}$ (e.g. CD4 count history in our running example) and thus is only a function of the always observed data $(\overline{A},\overline{Z})$.

Once we have obtained a CAN estimator of ψ^*, we can estimate $x_{\mathrm{opt}}(V)=x_{\mathrm{opt}}(V;\psi^*)$ as in Section 4.3. The estimator $\widehat{\psi}$ will be far from the most efficient possible estimator of ψ^* under the above assumptions. However, the theory of efficient estimation in CAR models could be used to obtain much more efficient estimators. See Robins and Rotnitzky [14].

6.2.2. Estimation of joint optimal testing and treatment regimes. Suppose we are given a set of nonrandom testing and treatment regimes $\{G_x\equiv\overline{\mathbf{d}}_x\equiv(\overline{d}_{x_a},\overline{d}_{x_t});x\in\mathscr{X}\}$ where $x=(x_a,x_t)\in\mathscr{X}=\mathscr{X}_a\times\mathscr{X}_t$. Here $\{\overline{d}_{x_a};x_a\in\mathscr{X}_a\}$ is a set of treatment regimes with $\overline{d}_{x_a}=\{d_{x_a,k};k=0,\ldots,K\}$, each $d_{x_a,k}[\overline{a}(k-1),\overline{t}(k-1),\overline{l}^*(k)]$ being a function taking values in $\mathscr{A}(k)$. Similarly, $\{\overline{d}_{x_t};x_t\in\mathscr{X}_t\}$ is a set of testing regimes with $\overline{d}_{x_t}=\{d_{x_t,k};k=0,\ldots,K\}$, each $d_{x_t,k}[\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]$ being a function taking values in $\{0,1\}$. Note that $G_x$as defined here is the G_x defined in the preceding section except now with

$$g[t(k)|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]=I(t(k)=d_{x_t,k}[\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)])$$

Also, note that in this section $\overline{\mathbf{d}}_{\mathbf{x}}$ (in bold) denotes a joint nonrandom treatment and testing regime, and $\overline{d}_{x_a}$ denotes the treatment part of the joint regime. In contrast, in earlier sections, a nonrandom treatment regime was denoted simply by $\overline{d}_x$ because we did not need to distinguish between treatment and testing regimes.

Suppose there is a cost c to performing a test. Then if $Y=y(\overline{A},\overline{L})$ is our previously defined public health utility converted somehow into monetary units, our new utility is $Y^*=Y-c\sum_{m=0}^{K-1}T(m)$, where $\sum_{m=0}^{K-1}T(m)$ denotes the number of tests received. Our goal is to find $x_{\mathrm{opt}}(V)=(x_{a,\mathrm{opt}}(V),x_{t,\mathrm{opt}}(V))$ that maximizes the expected utility $E_{(G_x,P^{\tau})}[Y^*|V]$ over all joint regimes $x=(x_a,x_t)$. More generally, we can let the cost at m depend on the entire observed past. That is, $C(m)\equiv c_m(\overline{O}^{*,t}(m))\equiv c_m(\overline{A}(m),\overline{T}(m-1),\overline{L}^*(m))$ for some known function $c_m(\overline{O}^{*,t}(m))\equiv c_m(\overline{A}(m),\overline{T}(m-1),\overline{L}^*(m))$. The utility in such a case becomes $Y^*=Y-\sum_{m=0}^{K-1}T(m)C(m)$.

DOI: 10.1002/sim

Under the conditions of Theorem 6.1 and assumption PO* (which we assume to hold in this subsection), (17), and consequently (20), still holds if Y is replaced by Y^* and U_x is defined as in (18) except that Δ_x is replaced by Δ_{x_a} and $g[T(k)|\overline{O}^{*,t}(k)]$ is replaced by $I(T(k)=d_{x_t,k}[\overline{O}^{*,t}(k)])$. Suppose we specify an MSM

$$E_{(G_x,P^\tau)}[Y^*|V]=h(x,V,\psi^*) \tag{23}$$

with $h(x,V,\psi)=(r(x,V)^{\mathrm{T}},r^*(V)^{\mathrm{T}})\psi$. Then, reasoning as in the preceding section, with Y^* replacing Y, a CAN estimator of ψ^* is the $\widehat{\psi}$ solving equation (21) provided conditions (i)–(iii) of the preceding subsection hold and $\widehat{U}_x$ is defined as in (22) but with Δ_x replaced by Δ_{x_a} and $g[T(k)|\overline{O}^{*,t}(k)]$ replaced by $I(T(k)=d_{x_t,k}[\overline{O}^{*,t}(k)])$. As in the previous subsection, once we have obtained a CAN estimator of ψ^*, we can estimate the optimal joint strategy in our candidate class by the $\widehat{x}_{\mathrm{opt}}(V)$ that maximizes $h(x,V,\widehat{\psi})$.

Running example (*continuation 1*): Let $\overline{d}_{x_a}=\{d_{x_a,k};k=0,\ldots,K\}$ denote the nonrandom dynamic treatment regime 'begin HAART the first time t the measured CD4 count falls below x_a' and $\overline{d}_{x_t}$ the testing regime 'if a subject is still at risk to initiate HAART, perform a CD4 test every x_t weeks'. Our goal is to find the $x_{\mathrm{opt}}(V)=(x_{a,\mathrm{opt}}(V),x_{t,\mathrm{opt}}(V))$ that minimizes $E_{(G_x,P^\tau)}[Y^*|V]=E_{(G_x,P^\tau)}[Y-\sum_{m=0}^{K-1}T(m)C(m)|V]$ as a function of V. A possible choice for $h(x,V,\psi)$ in model (23) that allows the effect of the treatment regime $\overline{d}_{x_a}$ to depend (as it should) on the testing regime $\overline{d}_{x_t}$ is

$$\begin{aligned}h(x,V,\psi)=&\sum_{k=1}^{5}\psi_{1,1,k}x_a^k+\sum_{k=1}^{5}\psi_{1,2,k}^{\mathrm{T}}Vx_a^k+\sum_{k=1}^{5}\psi_{2,1,k}x_t^k+\sum_{k=1}^{5}\psi_{2,2,k}^{\mathrm{T}}Vx_t^k\\&+\sum_{\{(k,l);k+l\leqslant 4,k,l\geqslant 0\}}\psi_{3,1,k,l}x_a^kx_t^l+\sum_{\{(k,l);k+l\leqslant 4,k,l\geqslant 0\}}\psi_{3,2,k,l}^{\mathrm{T}}Vx_a^kx_t^l+\psi_4^{\mathrm{T}}V\end{aligned}$$

Elsewhere, we show that additional substantive knowledge concerning the dependence of $E_{(G_x,P^\tau)}[Y|V]$ on $x=(x_a,x_t)$ can be incorporated into the functional form of our models if we specify separate models for $E_{(G_x,P^\tau)}[Y|V]$ and $E_{(G_x,P^\tau)}[\sum_{m=0}^{K-1}T(m)C(m)|V]$ and then combine the fits to estimate $x_{\mathrm{opt}}(V)$.

6.3. Usefulness of the NDE assumption

The results obtained in Section 6.2 were a straightforward generalization of the results of Section 4.3. In particular, we did not use the fact that the NDE assumption holds. However, the consistency of our estimators required that (a) the PO* assumption hold and (b) that $\widehat{U}_xY$ is a function of the observed data $\overline{O}^*$. In this section, we shall show that when (a) and/or (b) are/is false, between-population extrapolation and estimation of joint optimal treatment and testing regimes often remain possible by exploiting the NDE assumption to construct alternative CAN estimators. Because the methodology developed in this section is new, we treat several special cases based on our running example, beginning with the simplest. The most general case is reserved for Appendix A. It is pedagogically useful to begin with the problem of estimation of a joint optimal testing and treatment regime.

6.3.1. Special case 1. Consider running example (continuation 1) with $x=(x_a,x_t)$ fixed and $Y=y(\overline{A},\overline{L})$ being a prespecified function, say $y_1(\overline{A},\overline{Z})$, of $(\overline{A},\overline{Z})$ only. Suppose in population G^τ

DOI: 10.1002/sim

all subjects at risk to initiate HAART receive a CD4 blood test every week and thus $\text{pr}_{G^\top}[T(k)=1|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]=1$ if $a(k)=1-s(k)=0$. In contrast, suppose under G_x, CD4 tests are performed only every x_t weeks. Thus, $\text{pr}_{G_x}[T(k)=1|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]$ is equal to 1 if $a(k)=1-s(k)=0$ and k is divisible by x_t, but it is equal to 0 otherwise. Then the assumption PO* for $G^\intercal$ with respect to G_x is false for $x_t>1$, since, given $a(k)=1-s(k)=0$ for k not divisible by x_t, the test decision $T(k)=0$ has positive probability under G_x but has probability 0 under $G^\intercal$. Thus, the approach of Section 6.2.2 cannot be used to estimate $E_{(G_x,P^\intercal)}[Y^*]$ or $E_{(G_x,P^\intercal)}[Y^*|V]$. However, under the NDE assumption, consistent estimators of these quantities can be obtained provided the following positivity assumption holds.

Treatment positivity (TPO*) assumption for a regime $G^\intercal$ with respect to another regime G: The conditional measure $g(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})$ is dominated by (i.e. absolutely continuous w.r.t.) $g^\top(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})$ with $P^\top$-probability 1, i.e.

$$g(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})>0 \text{ implies } g^\top(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})>0 \text{ w.p.1. under } P^\top \tag{24}$$

Note that the PO* assumption implies the TPO* assumption but the converse is false.

The following algorithm returns a consistent estimator of $E_{(G_x,P^\intercal)}[Y^*]$ under the NDE and TPO* assumptions when $G^\intercal$ is sequentially randomized.

Algorithm 1

1. Construct a modification of the data set generated under $(G^\intercal, P^\intercal)$ by both recoding $T(k)=1$ as $T(k)=0$ and discarding the CD4 count data $I(k+1)$ whenever k is not divisible by x_t, where x_t is the frequency of CD4 testing under regime G_x.
2. Use the modified data to determine the set of subjects who followed the treatment regime $\overline{d}_{x_a}$. [Note that this set will differ from the set of subjects who followed the treatment regime $\overline{d}_{x_a}$ in the original unmodified data. For example, consider a subject whose CD4 count in the unmodified data first fell below x_a in week k and who started HAART in week k. If k is not divisible by x_t, then the subject will have followed the regime x_a in the unmodified data but not in the modified data. That is, although the treatment history in the modified data agrees with that in the unmodified data, by modifying a subject's testing history, a subject can have followed the treatment regime $\overline{d}_{x_a}$ in the original data set but not in the modified data or *vice versa*.]
3. Estimate $E_{(G_x,P^\intercal)}[Y^*]$ by a weighted average over the members of the set in step 2 of the utilities Y^*_{mod} calculated from the modified data set with weights equal to an estimate of $1/\prod_{k=0}^K g^\intercal[A(k)|\overline{O}^{*,a}(k)]$ with $\overline{O}^{*,a}(k)$ based on the original unmodified data.

Algorithm 1 is formalized as follows:

Define $t^*_{x_t,0}(a(0),l(0))\equiv d_{x_t,0}[a(0),t(-1),l(0)]$ where, by convention, $t(-1)=1$ for all subjects and for $k=1,\ldots,K$, recursively define $t^*_{x_t,k}(\overline{a}(k),\overline{l}_{\overline{a}(k-1)}(k))\equiv d_{x_t,k}[\overline{a}(k),\overline{t}^{**}(k-1),\overline{l}^*_{\overline{a}(k-1),\overline{t}^{**}(k-1)}(k)]$ where $t^{**}(j)=t^*_{x_t,j}(\overline{a}(j),\overline{l}_{\overline{a}(j-1)}(j))$. Furthermore, define $\overline{T}^*_{x_t}(0)=t^*_{x_t,0}(A(0),L(0))$ and $T^*_{x_t}(k)\equiv t^*_{x_t,k}(\overline{A}(k),\overline{L}(k))$. Thus,

$$\overline{T}^*_{x_t}\equiv\overline{t}^*_{x_t}(\overline{A},\overline{L})\equiv\overline{t}^*_{x_t,K}(\overline{A},\overline{L})$$

DOI: 10.1002/sim

is the test history a subject with observed data $(\overline{A}, \overline{L})$ would have if he followed the testing regime $\overline{d}_{x_t}$ and his observed treatment history. Further define

$$\delta_{x_a}(\overline{a}, \overline{t}, \overline{L}_{\overline{a}}) = \prod_{k=0}^{K} I\left(a(k) = d_{x_a,k}[\overline{a}(k-1), \overline{t}(k-1), \overline{L}^*_{\overline{a}(k-1),\overline{t}(k-1)}(k)]\right) \tag{25}$$

$$\Delta_{x_a}(\overline{t}) = \delta_{x_a}(\overline{A}, \overline{t}, \overline{L}_{\overline{A}}) \tag{26}$$

so $\Delta_{x_a}(\overline{t})$ is the indicator that a subject's observed treatment is equal to the treatment he would have received had he followed the treatment regime $\overline{d}_{x_a}$ and the nondynamic testing regime $\overline{t}$. Hence, $\Delta_{x_a}(\overline{T}^*_{x_t})$ is the indicator that a subject's observed treatment is equal to the treatment he would have received had he followed the treatment and testing regime $G_x = \overline{\mathbf{d}}_{x=(x_a,x_t)}$. The set of subjects with $\Delta_{x_a}(\overline{T}^*_{x_t}) = 1$ is precisely the set constructed in step 2 of the above algorithm. The estimator described in step 3 of the algorithm is then

$$\mathbb{P}_n\left[\Delta_{x_a}(\overline{T}^*_{x_t}) Y^*_{\text{mod}} \Big/ \left\{\prod_{k=0}^{K} g^{\intercal}\{A(k)|\overline{O}^{*,a}(k); \widehat{\alpha}\}\right\}\right] \tag{27}$$

divided by the same expression with Y^*_{mod} absent, where, by the NDE assumption that $T(k)$ has no causal effect on $\overline{L}$,

$$Y^*_{\text{mod}} = Y - \sum_{m=0}^{K-1} T^*_{x_t}(m) c_m(\overline{A}(m), \overline{T}^*_{x_t}(m-1), \overline{L}^*_{\overline{T}^*_{x_t}(m-1)}(m)) \tag{28}$$

for any subject with $\Delta_{x_a}(\overline{T}^*_{x_t}) = 1$, since the right-hand side (RHS) of the last display is the counterfactual value of Y^* that would have been observed, if possibly, contrary to fact, the subject had followed the regime G_x by modifying his testing history while leaving his treatment history unchanged. [Note this last statement would not be true were the NDE assumption false, since then simply recoding $T(k) = 1$ as $T(k) = 0$ could not undo the actual causal effect of $T(k) = 1$ on Y]. The consistency of (27) as an estimator of $E_{(G_x, P^{\intercal})}[Y^*]$ is a consequence of Corollary A.1 in Appendix A. Here we give a heuristic argument for its consistency.

Suppose for the moment the set constructed in step 2 consisted of all n study subject's because treatment regime $\overline{d}_x$ is followed with probability 1 when $\overline{T} = \overline{T}^*_{x_t}$. In that case, under the NDE assumption, we could calculate from the modified data set the utility of each of the n subjects under the regime G_x using equation (28). Their sample average would be an unbiased estimate of $E_{(G_x, P^{\intercal})}[Y^*]$. In practice, some subjects will have $\Delta_{x_a}(\overline{T}^*_{x_t}) = 0$ and thus will not be included in the set in step 2. Therefore, to preserve (asymptotic) unbiasedness we must, as in step 3 of the algorithm, weight those with $\Delta_{x_a}(\overline{T}^*_{x_t}) = 1$ by the estimated inverse $1/\prod_{k=0}^{K} g^{\intercal}[A(k)|\overline{O}^{*,a}(k); \widehat{\alpha}]$ of their conditional probability (based on the unmodified data) of having $\Delta_{x_a}(\overline{T}^*_{x_t}) = 1$ so as to represent those with $\Delta_{x_a}(\overline{T}^*_{x_t}) = 0$ that were removed from the data set in step 2. This inverse probability weighting is formally justified by the fact that, as shown in the proof of Theorem A.1 in Appendix A, under the TPO* assumption,

(i) $\text{pr}_{G^{\intercal}}[\overline{A} = \overline{a}|\overline{L}_{\overline{\mathscr{A}}}]$ is nonzero whenever $\overline{a} \in \overline{\mathscr{A}}_g$, where $\overline{\mathscr{A}}_g$ is the support of $\overline{A}$ under $(G, P^{\intercal})$ and

(ii) the expression in square brackets in (27) (with the true α in place of $\widehat{\alpha}$) equals $\sum_{\overline{a} \in \overline{\mathscr{A}}_g} I(\overline{A} = \overline{a}) m(x, \overline{a}, \overline{L}_{\overline{a}}) / \text{pr}_{G^{\intercal}}[\overline{A} = \overline{a}|\overline{L}_{\overline{\mathscr{A}}}]$ for a function $m(x, \overline{a}, \overline{L}_{\overline{a}})$ satisfying

DOI: 10.1002/sim

$\sum_{\overline{a}\in\mathscr{A}_g} E_{P^\intercal}[m(x,\overline{a},\overline{L}_{\overline{a}})] = E_{(G_x,P^\intercal)}[Y^*]$. Specifically, $m(x,\overline{a},\overline{L}_{\overline{a}})$ is $\Delta_{x_a}(\overline{T}^*_{x_t})Y^*_{\text{mod}}$ except with $(\overline{A},\overline{L}_{\overline{A}})$ replaced by $(\overline{a},\overline{L}_{\overline{a}})$. That is, $m(x,\overline{a},\overline{L}_{\overline{a}}) = \delta_{x_a}(\overline{a},\overline{t}^*_{x_t}(\overline{a},\overline{L}_{\overline{a}}),\overline{L}_{\overline{a}}) \times \{y_1(\overline{a},\overline{Z}_{\overline{a}}) - \sum_{m=0}^{K-1} t^*_{x_t,m}(\overline{a},\overline{L}_{\overline{a}}) c_m(\overline{a}(m),\overline{t}^*_{x_t,m-1}(\overline{a},\overline{L}_{\overline{a}}),\overline{L}^*_{\overline{a}(m-1),\overline{t}^*_{x_t,m-1}(\overline{a},\overline{L}_{\overline{a}})}(m))\}$.

6.3.2. Special case 2. Special case 2 differs from special case 1 only in that we no longer assume that all subjects in population $G^\intercal$ at risk to initiate HAART receive a CD4 blood test every week. That is, we no longer place *a priori* qualitative restrictions on $g^\top[t(k)|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]$. In this setting, it will be useful to define

$$\Delta(\overline{t}) = \prod_{k=0}^{K} I(T(k)\geqslant t(k)) \tag{29}$$

Note $\Delta(\overline{t})$ is 1 if and only if at each time k that a subject would have been tested under the given testing history $\overline{t}$, they were, in fact, tested under their observed testing history $\overline{T}$. Furthermore, in special case 1 but not in special case 2, $\Delta(\overline{T}^*_{x_t}) = 1$ with probability one since, in case 1, subjects were in fact tested every week while at risk to initiate HAART. When, as in case 2, $\Delta(\overline{T}^*_{x_t})$ may be 0, $\Delta_{x_a}(\overline{T}^*_{x_t}) = \prod_{k=0}^{K} I(A(k) = d_{x_a,k}[\overline{A}(k-1),\overline{T}^*_{x_t}(k-1),\overline{L}^*_{\overline{T}^*_{x_t}(k-1)}(k)])$ may not be a function of the observed data $\overline{O}^*$ that was generated under $(G^\intercal, P^\intercal)$ because (i) $\overline{L}^*_{\overline{T}^*_{x_t}(k-1)}(k)$ may then be unknown (i.e. missing) for certain values of k and (ii) the function $d_{x_a,k}[\overline{A}(k-1),\overline{T}^*_{x_t}(k-1),\overline{L}^*_{\overline{T}^*_{x_t}(k-1)}(k)]$ associated with the regime 'start HAART the first time CD4 count drops below x_a' depends nontrivially on $\overline{L}^*_{\overline{T}^*_{x_t}(k-1)}(k)$. Thus, estimator (27) of $E_{(G_x,P^\intercal)}[Y^*]$ cannot be used as it may fail to be a function of the observed data. To overcome this difficulty, we instead estimate $E_{(G_x,P^\intercal)}[Y^*]$ for a given $x = (x_a, x_t)$ by

$$\mathbb{P}_n[\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})Y^*_{\text{mod}}\widehat{W}_{x_t}] \tag{30}$$

where

$$\widehat{W}_{x_t} \equiv \left\{ \prod_{k=0}^{K} g^\intercal[A(k)|\overline{O}^{*,a}(k);\widehat{\alpha}] \prod_{\{k;T^*_{x_t}(k)=1\}} g^\intercal[T(k)|\overline{O}^{*,t}(k);\widehat{\alpha}] \right\}^{-1}$$

By replacing $\Delta_{x_a}(\overline{T}^*_{x_t})$ by $\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})$ we guarantee that equation (30) is a function of the observed data. Furthermore, since not all subjects have $\Delta_{x_t}(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t}) = 1$, to ensure unbiasedness (asymptotically) it is necessary to weight by $\widehat{W}_{x_t}$. Note that the product $\prod_{\{k;T^*_{x_t}(k)=1\}} g^\intercal[T(k)|\overline{O}^{*,t}(k);\widehat{\alpha}]$ is computed only over the times k at which $T^*_{x_t}(k)$ equals 1. Furthermore, note that in this product we can replace $g^\intercal[T(k)|\overline{O}^{*,t}(k);\widehat{\alpha}]$ with $\text{pr}_{G^\top}[T(k) = 1|\overline{O}^{*,t}(k);\widehat{\alpha}]$ because this product is premultiplied by the indicator $\Delta(\overline{T}^*_{x_t})$ and in fact, $\Delta(\overline{T}^*_{x_t}) = 1$ implies $T(k) = 1$ whenever $T^*_{x_t}(k) = 1$. An informal understanding as to why $\widehat{W}_{x_t}$ is the appropriate inverse probability weight can be obtained by considering a hybrid intervention in which every x_t weeks we intervene and force subjects to be tested. However, for any week k with k not divisible by x_t, we do not prevent all testing; rather we apply the

DOI: 10.1002/sim

random testing regime $g^{\intercal}[t(k)|\overline{o}^{*,t}(k)]$. However, in deciding treatment, we force each subject to take the treatment history they would have taken if they followed the regime $\overline{\mathbf{d}}_{x=(x_a,x_t)}$, thereby effectively ignoring the results of tests taken in weeks k not divisible by x_t. Under assumption I, the law of the hybrid regime is absolutely continuous with respect to $G^{\top}$ and the likelihood ratio (Radon-Nikodym derivative) of the law under the hybrid regime to that under the regime $G^{\top}$ generating the observed data is precisely $\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})W_{x_t}$ with $W_{x_t}=1/\prod_{k=0}^{K} g^{\intercal}[A(k)|\overline{O}^{*,a}(k)]\prod_{\{k;T^*_{x_t}(k)=1\}} g^{\intercal}[T(k)|\overline{O}^{*,t}(k)]$. Furthermore, if in calculating a subject's utility, only the costs of the tests obtained for weeks k divisible by x_t are included, the utility Y^* calculated under the hybrid regime equals that under the regime $\overline{\mathbf{d}}_{x=(x_a,x_t)}$. Thus our parameter of interest, the expected utility $E_{(G_x,P^{\intercal})}[Y^*]$ under $G_x=\overline{\mathbf{d}}_{x=(x_a,x_t)}$, will equal the expected utility under the hybrid regime. Thus, under assumption I, the expected utility under the hybrid regime will equal $E_{(G^{\intercal},P^{\intercal})}[\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})Y^*_{\mathrm{mod}}W_{x_t}]$ by the Radon-Nikodym theorem. [Recall that for subjects with $\Delta_{x_a}(\overline{T}^*_{x_t})=1$, Y^*_{mod} equals the utility Y^* that would be obtained under $G_x=\overline{\mathbf{d}}_{x=(x_a,x_t)}$.]

In Appendix A, we prove in Corollary A.1 that (30) is a consistent estimator of $E_{(G_x,P^{\intercal})}[Y^*]$ when $G^{\top}$ is $\overline{O}^{*,a}$-SR, the treatment and testing model parametrized by α is correct, the consistency assumption C holds and the following assumption also holds.

Assumption identify (I): The set

$$\left\{\overline{t};\overline{t}\geqslant\overline{T}_{G_x} \text{ and } \prod_{k=0}^{K} g^{\top}[A_{G_x}(k),t(k)|\overline{A}_{G_x}(k-1),\overline{t}(k-1),\overline{L}^*_{\overline{A}_{G_x}(k-1),\overline{t}(k-1)}(k)]>0\right\}$$

is nonempty with w.p.1. under $P^{\intercal}$ where in g^{T} we view the conditioning event as a realization of the observed past and with the inequality interpreted componentwise.

Remark

We now provide a heuristic explanation of assumption I. The counterfactual random variables $\overline{A}_{G_x}$ and $\overline{T}_{G_x}$ in the statement of assumption I are those naturally associated with the deterministic treatment and testing regime $G_x=\overline{\mathbf{d}}_{x=(x_a,x_t)}$. They are recursively constructed in a manner similar to that of Section 4.2. They can alternatively be defined as follows. $\overline{A}_{G_x}$ is the unique solution to $\delta_{x_a}(\overline{a},\overline{t}^*_{x_t}(\overline{a},\overline{L}_{\overline{a}}),\overline{L}_{\overline{a}})=1$. Furthermore, $\overline{T}_{G_x}=\overline{t}^*_{x_t}(\overline{A}_{G_x},\overline{L}_{\overline{A}_{G_x}})$. In addition, in what follows we will also use the counterfactual $\overline{L}_{G_x}$ which is defined as $\overline{L}_{G_x}\equiv\overline{L}_{\overline{A}_{G_x}}$. Note that with these definitions, $\Delta_{x_a}(\overline{T}^*_{x_t})=1$ if and only if $\overline{A}_{G_x}=\overline{A}$. Also, if $\overline{A}_{G_x}=\overline{A}$, then $\overline{L}_{G_x}=\overline{L}$ and $\overline{T}_{G_x}=\overline{T}^*_{x_t}$. Informally assumption I states that for every possible *treatment* and testing history $(\overline{A}_{G_x},\overline{T}_{G_x})$ that can occur under $(G_x,P^{\intercal})$, there must exist a testing history $\overline{t}$ that results in a test each time $\overline{T}_{G_x}$ does (and possibly at other times as well) such that $(\overline{A}_{G_x},\overline{t})$ has a nonzero probability of occurring under $(G^{\intercal},P^{\intercal})$. Assumption I is identifiable in the sense that if we knew the marginal distribution of O^* in population $G^{\intercal}$,we could determine whether assumption I hold. Thus, we can determine from the data whether it 'almost fails' as discussed in Section 4.3.

Although not explicitly stated as one of the conditions for consistency of (30), the TPO* assumption is indeed implicitly assumed since it is implied by the assumption I, as proved in Theorem A.1 and Corollary A.1. Furthermore, it can be shown that the assumption I implies that the hybrid regime is absolutely continuous with respect to $G^{\top}$.

DOI: 10.1002/sim

Estimation for arbitrary treatment and testing regimes when the utility function does not depend on $\overline{I}$: In both special cases 1 and 2, we assumed a particular treatment and testing regime $\overline{\mathbf{d}}_{x=(x_a,x_t)}$. In fact, Corollary A.1 in Appendix A implies that the estimator of equation (30) is consistent for $E_{(G_x,P^\tau)}[Y^*]$ for any deterministic treatment and testing regime $\overline{\mathbf{d}}_{x=(x_a,x_t)}$ when $G^\top$ is $\overline{O}^{*,a}$-SR, assumptions C and I hold, and the model for treatment and testing is correct. Furthermore, if the same assumptions required for the consistency of (30) hold for all $x=(x_a,x_t)$ in a set $\mathscr{X}$, a CAN estimator of the parameter ψ^* of MSM (23) is the $\widehat{\psi}$ solving equation (21) with $\widehat{U}_x$ redefined as $\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})\widehat{W}_{x_t}$ and Y replaced by Y^*_{mod}. Thus, we can view the data as consisting of $n\times\text{card}(\mathscr{X}_a)\times\text{card}(\mathscr{X}_t)$ subjects indexed by the ordered triple (i,x_1,x_2), each of whom is entered into a weighted least-squares regression of Y^*_{mod} on $(r(x,V)^{\text{T}},r^*(V)^{\text{T}})$ with weight $q^*(x,V)\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})\widehat{W}_{x_t}$ if and only if $\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})=1$. The resulting weighted least-squares estimator is precisely the estimator $\widehat{\psi}$. As above, once we have obtained a CAN estimator of ψ^*, we can estimate the optimal joint strategy in our candidate class by the $\widehat{x}_{\text{opt}}(V)$ that maximizes $h(x,V,\widehat{\psi})$.

6.3.3. Special case 3. In special cases 1 and 2 we assumed that Y was a function $y_1(\overline{A},\overline{Z})$ of the always observed variables $(\overline{A},\overline{Z})$ and in particular did not depend on any component of CD4 count history $\overline{I}$. Case 3 differs from case 2 only in that we now suppose instead that $Y=y(\overline{A},\overline{L})$ is a nontrivial function of every component of $\overline{I}$. This would be the case, for example, if the study followed HIV-infected subjects with baseline CD4 count equal to 500 for one year and the utility Y were the integrated CD4 count $\sum_{m=0}^K I(m)$.

Consider the estimator (30) of $E_{(G_x,P^\tau)}[Y^*]$ with $G_x=\overline{\mathbf{d}}_{x=(x_a,x_t)}$ and with Y now a nontrivial function of $\overline{I}$, for example, $Y=\sum_{m=0}^K I(m)$. The quantity displayed in (30) is no longer a function of the observed data $\overline{O}^*$. However, this quantity again becomes a function of the observed data if we substitute $I(\overline{T}=\overline{1})$ for $\Delta(\overline{T}^*_{x_t})$ where $\overline{1}$ is the vector of all 1's.

However, this substitution requires us to replace the product $\prod_{\{k;T^*_{x_t}(k)=1\}}$ by the product $\prod_{k=0}^K$ in the definition of $\widehat{W}_{x_t}$ in order to obtain inverse probability weights that preserve consistency. Corollary A.1 of Appendix A implies that with these changes, the modified quantity in (30) is consistent for $E_{(G_x,P^\tau)}[Y^*]$ provided $G^\top$ is $\overline{O}^{*,a}$-SR, the testing and treatment model is correct, and assumption C and the following assumption holds.

Assumption
For $\overline{t}=\overline{1}$,

$$\prod_{k=0}^{K} g^\top[\overline{A}_{G_x}(k),t(k)|\overline{A}_{G_x}(k-1),\overline{t}(k-1),\overline{L}^*_{\overline{A}_{G_x}(k-1),\overline{t}(k-1)}(k)]>0 \quad \text{with w.p.1. under } P^\tau$$

Note that because $G^\top$ is $\overline{O}^{*,a}$-SR, this latter assumption is equivalent to the assumption that $\text{pr}_{G^\tau}[\overline{A}=\overline{A}_{G_x},\overline{T}=\overline{1}|\overline{L}_{\overline{\mathscr{A}}}]\equiv\text{pr}_{G^\tau}[\Delta_{x_a}(\overline{T}^*_{x_t})=1,\overline{T}=\overline{1}|\overline{L}_{\overline{\mathscr{A}}}]>0$ w.p.1. under P^τ.

6.3.4. Special case 4. We return to the problem of between-population extrapolation. We are given data $\overline{O}^*$ generated under $(G^\top,P^\tau)$ and wish to estimate $E_{(G_x,P^\tau)}[Y]$ for G_x as defined in equations (14) and (15), with $g[t(k)|\overline{a}(k),\overline{t}(k-1),\overline{l}^*(k)]$ the known conditional testing density in a population G and $\overline{d}_x$ a deterministic treatment regime. However, we no longer assume that the

DOI: 10.1002/sim

PO* assumption holds so the approach of Section 6.2.1 is unavailable. For simplicity, we assume that Y is only a function of $(\overline{A}, \overline{Z})$. Consider the estimator

$$\mathbb{P}_n \left[\sum_{\overline{t}^* \in \overline{\mathscr{T}}} \Delta(\overline{t}^*) \Delta_x(\overline{t}^*) Y \widehat{W}(\overline{t}^*) \prod_{k=0}^{K} g[t^*(k) | \overline{A}(k), \overline{t}^*(k-1), \overline{L}^*_{\overline{t}^*(k-1)}(k)] \right] \tag{31}$$

where

$$\widehat{W}(\overline{t}^*) = \left\{ \prod_{k=0}^{K} g^{\intercal}[A(k) | \overline{O}^{*,a}(k); \widehat{\alpha}] \prod_{\{k; t^*(k)=1\}} g^{\intercal}[T(k) | \overline{O}^{*,t}(k); \widehat{\alpha}] \right\}^{-1}$$

with $\Delta_x(\overline{t}^*)$ as in equation (26) with $x_a = x$ and $\overline{\mathscr{T}}$ the set of $K+1$ vectors of 0 and 1's. Note that estimator (31) is a sum of 2^{K+1} terms indexed by the vectors in $\overline{\mathscr{T}}$ with each term identical to those in estimator (30) except that Y^*_{mod} is redefined to be Y and $\overline{T}^*_{x_t}$ is replaced by a given vector $\overline{t}^* \in \overline{\mathscr{T}}$. This estimator, while feasible in theory, will be computationally intractable when K is large because the set $\overline{\mathscr{T}}$ contains 2^{K+1} elements.

To decrease the computational burden, we construct a simulation-based estimator, motivated by the observation that the subject-specific contribution to estimator (31) is the expectation of $\Delta(\overline{t}^*)\Delta_x(\overline{t}^*) Y \widehat{W}(\overline{t}^*)$ with respect to the conditional density $\widetilde{g}(\overline{t}^* | \overline{A}, \overline{L}, \overline{T}) \equiv \prod_{k=0}^{K} g[t^*(k) | \overline{A}(k), \overline{t}^*(k-1), \overline{L}^*_{\overline{t}^*(k-1)}(k)]$. Therefore, we consider the simulation estimator

$$\mathbb{P}_n \left[S^{-1} \sum_{s=1}^{S} \Delta(\overline{T}^*_s) \Delta_x(\overline{T}^*_s) Y \widehat{W}(\overline{T}^*_s) \right]$$

where, for each subject, the $\overline{T}^*_s$ are sampled independently from $\widetilde{g}(\overline{t}^* | \overline{A}, \overline{L}, \overline{T})$ by recursively drawing from the Bernoulli densities $g[t^*(k) | \overline{A}(k), \overline{t}^*(k-1), \overline{L}^*_{\overline{t}^*(k-1)}(k)], k = 0, \ldots, K$. Corollary A.2 of Appendix A implies that both estimator (31) and the associated simulation estimator are consistent for $E_{(G_x, P^{\intercal})}[Y^*]$ provided $G^{\top}$ and G_x are $\overline{O}^{*,a}$-SR, the model for treatment and testing is correct, assumption C and the NDE assumption hold, and the identifying assumption in display (A1) given in Appendix A holds. Assumption (A1) is the generalization of Assumption I above to the setting of a random testing regime. Like Assumption I it implies the TPO* assumption.

Furthermore, under these same conditions, for a given set of treatment regimes $\{\overline{d}_x; x \in \mathscr{X}\}$ for which assumption (A1) holds, a CAN estimator of the parameter ψ^* of the MSM (19) can be obtained by weighted least squares as follows. View the data set as consisting of $n \times S \times \text{card}(\mathscr{X})$ subjects indexed by the ordered triple (i, s, x), each of whom is entered into the weighted least-squares regression of Y on $(r(x, V)^{\mathrm{T}}, r^*(V)^{\mathrm{T}})$ with weight $q^*(x, V)\widehat{W}(\overline{T}^*_s)$ if and only if $\Delta(\overline{T}^*_s)\Delta_x(\overline{T}^*_s) = 1$. Once we have obtained a CAN estimator of ψ^*, we can estimate the optimal strategy in our candidate class by the $\widehat{x}_{\text{opt}}(V)$ that maximizes $h(x, V, \widehat{\psi})$.

7. EXTRAPOLATION UNDER A NONIGNORABLE VISIT PROCESS

Suppose we have data from a comprehensive HMO database that records all patient visits. HIV-infected patients come to a clinic to be seen by a physician at a time t either because of acute

DOI: 10.1002/sim

symptoms or because of a scheduled follow-up appointment. Data on the reason for a given visit are often not recorded for data analysis. This random clinic visit process, unlike the regular weekly visit process we have assumed heretofore, results in an association between visits and risk since patients who are sick are generally more likely to return to the clinic at frequent intervals both spontaneously and to keep frequent appointments (although it is possible moderately ill patients who are not getting better come infrequently and miss many scheduled visits).

In this section, we consider the realistic setting of a random visit process. The visit process will result in nonignorable missing data unless we collect data at t on the health status of not only those who return to the clinic at t (which we typically do) but also those who do not come to the clinic at t (which we almost never do). Although we shall assume that the visit process is nonignorable, it is often reasonable to assume that among patients coming to the clinic at t, the decision to treat or not can be viewed as effectively randomized (i.e. ignorable) given the data recorded in the past including data on the health status measured at a visit at time t. Under this assumption, as shown below, we can still estimate the optimal CD4 count at which to start HAART in the study population (provided the utility Y is always observed) but extrapolation to another population of biologically similar individuals with a different visit process becomes problematic.

We shall assume that a subject's treatment only changes at the time of a clinic visit. This assumption may (approximately) hold if we are interested in the causal effect of a treatment being prescribed by a physician rather than the effect of the actual treatment taken by the patient, as would be the case if either an intervention that would increase patient compliance is unavailable or the only treatment data recorded for data analysis are the prescription data.

We shall also assume that the diagnostic and/or prognostic tests discussed earlier are obtained at and only at each clinic visit. This assumption allows us to consider the clinic visit problem without having to also worry about the diagnostic testing problem as well. We can then reuse our previous notation by redefining $T(m)$ to be the indicator of a clinic visit in week $m+1$, deleting $Z(m)$, and redefining $I(m)$ to be all covariates that would be recorded at week m if a visit occurred so that now $L(m)=I(m)$. Therefore, $L^*(m)=I^*(m)=T(m-1)L(m)$ becomes the covariate data actually observed at m. In this context, the NDE assumption that $\overline{L}_{\overline{a}}=\overline{L}_{\overline{j}}$ with probability one for $\overline{j}=(\overline{a},\overline{t})$ becomes the assumption that a clinic visit has no effect on the responses $L(m)$ except through its effect on the treatments prescribed. For the NDE assumption to hold, $A(t)$ would need to encode the vector of all active treatments that a physician might prescribe to a patient at a clinic visit in week t. Thus, if $A(t)$ were simply the indicator variable that takes the value 1 if and only if HAART has been initiated by time t, the NDE assumption would only hold if, as would never actually be the case, physicians make no treatment interventions whatsoever once HAART has been initiated.

As mentioned above, even were our goal only to estimate the optimal treatment regime in the study population, we require that the utility Y be observed for all study subjects. Therefore, we assume that there exist 'end point' variables, such as the minimum of time of death and end of follow-up at $K+1$, that are observed regardless of visit history but only become available after $K+1$, say by consultation of the national death index. Such 'end point' variables can be easily appended to our data by adding a formal 'time' $K+2$ to the data, setting by convention $T(K+1)$ to 1 for all subjects, and letting $L(K+2)$ encode the vector of 'end point' variables. We take our prespecified utility function to be a function $Y=y(\overline{A}, L(K+2))$ of treatment history and such 'end point' variables.

Again we are given data $\overline{O}^*$ generated under $(G^\top, P^\intercal)$ and wish to estimate the mean utility $E_{(G_x,P^\intercal)}[Y]$ for $x\in\mathscr{X}$ in a second population G with G_x as defined in equations (14) and (15).

DOI: 10.1002/sim

The following encodes the assumption that the treatment process is ignorable in the study population $G^\dagger$.

Assumption treatment (Tr)-$\overline{O}^{*,a}$-*SR*: We say a conditional law G of $(\overline{A}, \overline{T})$ given $\overline{L}_{\overline{\mathscr{A}}}$ is treatment sequentially randomized given data $\overline{O}^*$ if for all k

$$g^\dagger[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}_{\overline{\mathscr{A}}}] = g^\dagger[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*_{\overline{a}(k-1), \overline{t}(k-1)}(k)] \tag{32}$$

where we recall that the RHS of (32) can be written as $g^\dagger[a(k)|\overline{o}^{*,a}(k)]$. The assumption

$$g^\dagger[a(k)|\overline{a}(k-1), \overline{t}(k-1), \overline{l}^*_{\overline{a}(k-1), \overline{t}(k-1)}(k)] = I(a(k) = a(k-1)) \quad \text{if } t(k-1) = 0$$

encodes the assumption that if, in the observed study population, a clinic visit did not occur at time k, then a subject's treatment does not change at k.

Suppose we did not wish to extrapolate to population G but rather wished to estimate the expected utility $E_{(G^\dagger_x, P^\dagger)}(Y)$ in the study population $G^\dagger$ when the treatment regime $\overline{d}_x$ is followed, where $G^\dagger_x$ is defined like G_x but with $g^\dagger$ instead of g. Then, without making any assumptions whatsoever concerning the visit process $\text{pr}_{G^\dagger}[T(k) = 1|\overline{A}(k), \overline{T}(k-1), \overline{L}_{\overline{\mathscr{A}}}]$, $E_{(G^\dagger_x, P^\dagger)}(Y)$ is identified under the assumption that $G^\dagger$ is TR-$\overline{O}^{*,a}$-SR and the positivity assumption that $\text{pr}[A(k) = d_{x,k}[\overline{A}(k-1), \overline{T}(k-1), \overline{L}^*(k)]|\overline{A}(k-1), \overline{T}(k-1), \overline{L}^*(k)] > 0$ for each k w.p.1. under $(G^\dagger, P^\dagger)$. Specifically, $E_{(G^\dagger_x, P^\dagger)}(Y)$ is identified and satisfies

$$E_{G^\dagger_x, P^\dagger}(Y) = E_{G^\top, P^\dagger}[\Delta_x(\overline{T}) Y W], \quad \text{where}$$

$$W = \left\{ \prod_{k=0}^{K} g^\dagger[A(k)|\overline{A}(k-1), \overline{T}(k-1), \overline{L}^*(k)] \right\}^{-1}$$

and

$$\Delta_x(\overline{T}) = \prod_{k=0}^{K} I(A(k) = d_{x,k}[\overline{A}(k-1), \overline{T}(k-1), \overline{L}^*(k)])$$

In contrast, if we want to extrapolate to a population G with a different visit process and estimate $E_{(G_x, P^\dagger)}[Y]$, matters become more difficult. One approach would be to assume a nonignorable selection bias model visit process for G^{T} such as the following. For $k = 0, \ldots, K$,

$$\text{pr}_{G^\dagger}[T(k) = 1|\overline{A}(k) = \overline{a}(k), \overline{T}(k-1), \overline{L}^*(k), \overline{Y}_{\overline{a}} = y] = 1/\{1 + \exp[-\{H^\dagger_k(\overline{a}) + Q^\dagger_k(\overline{a}, y)\}]\} \tag{33}$$

where $H^\dagger_k(\overline{a}) = h^\dagger_k(\overline{a}, \overline{T}(k-1), \overline{L}^*(k))$ is an unspecified function and $Q^\dagger_k(\overline{a}, y) = q^\dagger_k(\overline{a}, \overline{T}(k-1), \overline{L}^*(k), y)$ is a selection bias function that captures the association between the counterfactual utility $\overline{Y}_{\overline{a}}$ and $T(k)$ that remains after conditioning on the observed past $(\overline{O}^{*,a}(k), A(k))$ and that is assumed known for purposes of data analysis. In a sensitivity analysis, we would study how our inferences concerning $E_{(G_x, P^\dagger)}[Y]$ change as we vary the selection bias function. If model (33), the assumption that $G^\dagger$ is TR-$\overline{O}^{*,a}$-SR, and an appropriate positivity assumption all hold, then it follows from Section 7 of Robins *et al.* [15] that if we regard $g[T(k)|\overline{A}(k) = \overline{a}(k), \overline{T}(k-1), \overline{L}^*(k), \overline{Y}_{\overline{a}} = y]$ as a known function, $g^\ddagger[\overline{T}(k), \overline{a}, \overline{L}^*(k), y]$, say, then both $H_k(\overline{A})$ and $E_{G_x, P^\dagger}(Y)$ are identified with the identifying formula for $H_k(\overline{A})$ given in Section 7 of

DOI: 10.1002/sim

Robins *et al.* and

$$E_{G_x,P^\tau}(Y) = E_{G^\top,P^\tau}[\Delta_x(\overline{T})YWW^\ddagger], \quad \text{where}$$

$$W^\ddagger = \left\{ \prod_{k=0}^{K} \frac{g^\ddagger[\overline{T}(k), \overline{A}, \overline{L}^*(k), Y]\{1+\exp\{H_k^\tau(\overline{A})+Q_k^\tau(\overline{A},Y)\}\}}{\exp\{H_k^\tau(\overline{A})+Q_k^\tau(\overline{A},Y)\}^{T(k)}} \right\}$$

In fact, in this context, the NDE assumption is often not required.

However, rather than taking this approach, we shall instead choose to make no assumptions whatsoever concerning the conditional visit probabilities $\mathrm{pr}_{G^\tau}[T(k)=1|\overline{A}(k), \overline{T}(k-1), \overline{L}_{\overline{\mathscr{A}}}]$ (which implies we also make no assumptions concerning $\mathrm{pr}_{G^\tau}[T(k)=1|\overline{A}(k)=\overline{a}(k), \overline{T}(k-1), \overline{L}^*(k), \overline{Y}_{\overline{a}}]$). We will then investigate what assumptions, in addition to those needed to identify $E_{G_x^\tau,P^\tau}(Y)$, are required to identify $E_{(G_x,P^\tau)}[Y]$ in this setting. We shall see that identification requires both the NDE assumption and an assumption that implies the visit rate under G_x is less than the rate under G^τ in a very strong sense.

Our identifying assumptions will require that, unlike earlier, we consider the joint distribution of the counterfactual variable $\overline{T}_{(\overline{a},G^\tau)}$ and $\overline{T}_{G_x}$ recording a subject's visit histories under regimes $(\overline{a}, G^\tau)$ and G_x, where the regime $(\overline{a}, G^\tau)$ denotes the regime with deterministic treatment history $\overline{a}$ and random visit history determined by the visit process $\mathrm{pr}_{G^\tau}[T(k)=1|\overline{a}(k), \overline{t}(k-1), \overline{L}_{\overline{\mathscr{A}}}]$ of the observed data. For notational convenience, we shall let $\overline{T}_{\overline{a}}$ denote $\overline{T}_{\overline{a},G^\tau}$ and $\overline{T}^\ddagger$ denote $\overline{T}_{G_x}$. Further, we shall include $T_{\overline{a}}(k)$ as a component of $L_{\overline{a}}(k)$ and henceforth denote the components of $L_{\overline{a}}(k)$ other than $T_{\overline{a}}(k)$ by $H_{\overline{a}}(k)=H_{\overline{a}(k-1)}(k)$. It will also be notationally useful to define $H^{\ddagger,*}_{\overline{a}(k-1),\overline{t}^\ddagger(k-1)}(k)=H_{\overline{a}(k-1)}(k)t^\ddagger(k-1)$ to distinguish it from $H^*_{\overline{a}(k-1),\overline{t}_{\overline{a}(k-1)}}(k) \equiv H_{\overline{a}(k-1)}(k)t_{\overline{a}(k-1)}(k-1)$.

With these notational changes, each candidate treatment regime $\overline{d}_x$ is equal to a collection of functions $\{\overline{d}_{x,k}; k=0,\ldots,K\}$ where each $\overline{d}_{x,k}$ maps the data $(\overline{a}(k-1), \overline{t}^\ddagger(k-1), \overline{h}^{\ddagger,*}_{\overline{a}(k-1),\overline{t}^\ddagger(k-1)}(k))$ that would be available at k under G_x to a treatment $a(k)$. That is, $\overline{d}_{x,k}[\overline{a}(k-1), \overline{t}^\ddagger(k-1), \overline{h}^{\ddagger,*}_{\overline{a}(k-1),\overline{t}^\ddagger(k-1)}(k)] \in \mathscr{A}(k)$. Furthermore, we assume

$$\overline{d}_{x,k}[\overline{a}(k-1), \overline{t}^\ddagger(k-1), \overline{h}^{\ddagger,*}_{\overline{a}(k-1),\overline{t}^\ddagger(k-1)}(k)] = a(k-1) \quad \text{if } t^\ddagger(k-1)=0$$

to encode the assumption that treatment only changes at a visit. Further, by definition of G_x in equations (14) and (15), there exists a known G such that, for all x's, $\mathrm{pr}_{G_x}[T^\ddagger(k)=1|\overline{A}(k-1), \overline{T}^\ddagger(k-1), \overline{L}_{\overline{\mathscr{A}}}] \equiv \mathrm{pr}_G[T^\ddagger(k)=1|\overline{A}(k-1), \overline{T}^\ddagger(k-1), \overline{L}_{\overline{\mathscr{A}}}]$ whenever the density of the event $(\overline{A}(k-1), \overline{T}^\ddagger(k-1), \overline{L}_{\overline{\mathscr{A}}})$ is nonzero under (G_x, P^τ), where now $\overline{L}_{\overline{\mathscr{A}}}=(\overline{H}_{\overline{\mathscr{A}}}, \overline{T}_{\overline{\mathscr{A}}})$. Our key identifying assumption is the following.

Visit process (VP) *assumption*:

(i) For all k,

$$\mathrm{pr}_G[T^\ddagger(k)=1|\overline{A}(k), \overline{T}^\ddagger(k-1), \overline{L}_{\overline{\mathscr{A}}}] = \mathrm{pr}_G[T^\ddagger(k)=1|\overline{A}(k), \overline{T}^\ddagger(k-1), \overline{T}(k), \overline{H}^*(k)] \quad (34)$$

where $\overline{T}(k)=\overline{T}_{\overline{A}(k)}(k)$ and $\overline{H}^*(k)=\overline{H}^*_{\overline{A}(k-1),\overline{T}(k-1)}(k)$.

DOI: 10.1002/sim

(ii) No new visit (NNV) assumption: With probability 1,

$$\mathrm{pr}_G[T^{\ddagger}(k)=1|\overline{A}(k),\overline{T}^{\ddagger}(k-1),\overline{T}(k),\overline{H}^{*}(k)]=0 \quad \text{if } T(k)=0 \tag{35}$$

Assumption (34) states that the g-conditional density of $T^{\ddagger}(k)$ given $\overline{L}_{\overline{\mathscr{A}}}$ and $(\overline{A}(k),\overline{T}^{\ddagger}(k-1))$ depends on $\overline{L}_{\overline{\mathscr{A}}}$ through and only through the components $(\overline{T}(k),\overline{H}^{*}(k))$ of $\overline{L}_{\overline{\mathscr{A}}}$ that would be observed under the regime $(\overline{a},G^{\intercal})$ with $\overline{a}(k-1)=\overline{A}(k-1)$.

Under the VP assumption, the expected utility $E_{G_x,P^{\intercal}}[Y]$ of $Y=y(\overline{A},L(K+2))$ is mathematically well defined. Specifically, the expectation is w.r.t. to the density $f_{G_x,P^{\intercal}}(\overline{T}^{\ddagger},\overline{A},\overline{L}_{\overline{\mathscr{A}}})=g_x(\overline{T}^{\ddagger},\overline{A}|\overline{L}_{\overline{\mathscr{A}}})p^{\intercal}[\overline{L}_{\overline{\mathscr{A}}}]$ with

$$g_x(\overline{T}^{\ddagger},\overline{A}|\overline{L}_{\overline{\mathscr{A}}}) \tag{36}$$

$$=\prod_{k=0}^{K} g[T^{\ddagger}(k)|\overline{A}(k),\overline{T}^{\ddagger}(k-1),\overline{T}(k),\overline{H}^{*}(k)] \tag{37}$$

$$\times\prod_{k=0}^{K} I(A(k)=d_{x,k}[\overline{A}(k-1),\overline{T}^{\ddagger}(k-1),\overline{H}^{\ddagger *}(k)]) \tag{38}$$

where $\overline{H}^{\ddagger *}(k)$has components $H_{\overline{A}(m-1)}(m)T^{\ddagger}(m-1)$ for $m\leqslant k$.

Equation (35) formalizes the very strong sense in which the NNV assumption requires that the visit rate under G is less than the visit rate under $G^{\intercal}$. Specifically, equation (35) states that for any week $k+1$ in which a subject with treatment history $\overline{A}(k)$ would not have had a visit under $G^{\intercal}$, he also would not have had a visit under G. The NNV assumption will essentially never be precisely true, except in the special case in which all subjects made weekly visits under $G^{\intercal}$ so $T(k)$ is always 1 for all k's. However, as discussed below, it cannot be weakened without sacrificing the identifiability of $E_{G_x,P^{\intercal}}[Y]$ from the data $\overline{O}^{*}$sampled from $(G^{\intercal},P^{\intercal})$. Identification will additionally require the following positivity assumption.

Restricted positivity (RP) *assumption*: With $(G_x,P^{\intercal})$ probability 1, for $k\leqslant K, g^{\intercal}[a(k)|\overline{A}(k-1),\overline{T}(k-2),\overline{H}^{*}(k),T(k-1)=1]>0$ for all $a(k)$ in the support of $\overline{A}(k)$ under the conditional law $g_x(a(k)|\overline{A}(k-1),\ \overline{L}_{\overline{\mathscr{A}}},T(k-1)=1)$, where we recall $T(k-1)=T_{\overline{A}(k-1)}(k-1)$ and $T_{\overline{a}(k-1)}(k-1)\in\overline{L}_{\overline{\mathscr{A}}}$.

Note the RP assumption only refers to the case $T_{\overline{A}(k-1)}(k-1)=1$. The case $T_{\overline{A}(k-1)}(k-1)=0$ follows from the NNV assumption. That is, by assumption, $g^{\intercal}[a(k)|\overline{A}(k-1),\overline{T}(k-2),\overline{H}^{*}(k),T(k-1)=0]=I(a(k)=A(k-1))$. But, under the NNV assumption, $g_x(a(k)|\overline{A}(k-1),\overline{L}_{\overline{\mathscr{A}}},T(k-1)=0)=I(a(k)=A(k-1))$ also holds.

Our main result is the following.

Theorem 7.1

Suppose $G^{\intercal}$ is TR-$\overline{O}^{*,a}$-SR, and for known $g[T^{\ddagger}(k)|\overline{A}(k),\overline{T}^{\ddagger}(k-1),\overline{L}_{\overline{\mathscr{A}}}]$ and $\overline{d}_x$, the NDE, the VP, the RP, and the C assumptions hold. Then, $E_{G_x,P^{\intercal}}[Y]$ with $Y=y(\overline{A},L(K+2))$ is identified from data $\overline{O}^{*}$ generated under $(G^{\top},P^{\intercal})$ and equals $E_{G^{\top},P^{\intercal}}[U^{\dagger}(g,g^{\top})Y]$, where $U^{\dagger}(g,g^{\top})$

DOI: 10.1002/sim

with

$$U^{\dagger}(g,g^{\top})=\frac{\sum_{\bar{t}^{\ddagger}\in\overline{\mathscr{T}}^{\ddagger}}\prod_{k=0}^{K}g[t^{\ddagger}(k)|\overline{A}(k),\bar{t}^{\ddagger}(k-1),\overline{T}(k),\overline{H}^{*}(k)]\Delta_{x}(\bar{t}^{\ddagger})}{\prod_{k=0}^{K}g^{\intercal}[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{H}^{*}(k)]}$$

with $\Delta_{x}(\bar{t}^{\ddagger})=\prod_{k=0}^{K}I\,(A(k)=d_{x,k}[\overline{A}(k-1),\bar{t}^{\ddagger}(k-1),\overline{H}^{\ddagger,*}_{\overline{A}(k-1),\bar{t}^{\ddagger}(k-1)}(k)])$.

Proof

Under $(G_{x},P^{\intercal})$ and $(G^{\intercal},P^{\intercal})$, respectively, the likelihood of an observation $(\overline{A},\overline{L}_{\overline{\mathscr{A}}})$ is $\sum_{\bar{t}^{\ddagger}\in\overline{\mathscr{T}}^{\ddagger}}g(\overline{A},\bar{t}^{\ddagger}|\overline{L}_{\overline{\mathscr{A}}})p^{\intercal}(\overline{L}_{\overline{\mathscr{A}}})$ and $g^{\intercal}(\overline{A}|\overline{L}_{\overline{\mathscr{A}}})p^{\intercal}(\overline{L}_{\overline{\mathscr{A}}})$. Under our assumptions, $\sum_{\bar{t}^{\ddagger}\in\overline{\mathscr{R}}^{\ddagger}}g(\overline{A},\bar{t}^{\ddagger}|\overline{L}_{\overline{\mathscr{A}}})p^{\intercal}(\overline{L}_{\overline{\mathscr{A}}})$ is absolutely continuous w.r.t. $g^{\intercal}(\overline{A}|\overline{L}_{\overline{\mathscr{A}}})p^{\intercal}(\overline{L}_{\overline{\mathscr{A}}})$ and the Radon-Nikodym derivative is $U^{\dagger}(g,g^{\top})$. Thus, $E_{G,P^{\intercal}}[Y]=E_{(G^{\intercal},P^{\intercal})}[U^{\dagger}(g,g^{\top})Y]$. Finally, $U^{\dagger}(g,g^{\top})$ and Y are functions of the observed data $\overline{O}^{*}$. □

The above proof of identifiability would have failed if the NNV assumption did not hold. First, absolute continuity would have failed because, the joint event $A(k)\neq A(k-1)$ and $T_{\overline{A}(k-1)}(k)=0$, which has probability zero under $(G^{\intercal},P^{\intercal})$, could then have nonzero probability under $(G_{x},P^{\intercal})$. In addition, if the NNV assumption did not hold, $\Delta_{x}(\bar{t}^{\ddagger})$ need not be a function of $\overline{O}^{*}$.

Recall that our final goal is to optimize $E_{(G_{x},P^{\intercal})}[Y|V]$ over $x\in\mathscr{X}$. Suppose we specify an MSM $E_{(G_{x},P^{\intercal})}[Y|V]=h(x,V,\psi^{*})$ with $h(x,V,\psi)=(r(x,V)^{\mathrm{T}},r^{*}(V)^{\mathrm{T}})\psi$. Then, for a given $q^{*}(x,V)$, if the assumptions of Theorem 7.1 hold for all $x\in\mathscr{X}$, Theorem 7.1 implies the estimator $\widehat{\psi}$ solving (21) remains CAN when we replace the numerator of $\widehat{U}_{x}$ in (21) by the numerator of $U^{\dagger}(g,g^{\top})$ and the denominator by $\prod_{k=0}^{K}g^{\intercal}[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{H}^{*}(k);\widehat{\alpha}]$.

However, the estimator $\widehat{\psi}$ will often be computationally infeasible when K is large because the set $\overline{\mathscr{R}}^{\ddagger}$ contains 2^{K+1} elements. Therefore, we replace the previous expression with its unbiased simulation-based estimator $S^{-1}\sum_{s=1}^{S}\Delta_{x}(\bar{t}^{\ddagger}_{s})$, where, for $s=1,\ldots,S$ and $k=0,\ldots,K+1$, $\bar{t}^{\ddagger}_{s}(-1)$ is defined to be zero and $\bar{t}^{\ddagger}_{s}(k)$ is obtained recursively as a random draw from the (assumed known) Bernoulli density $g[t^{\ddagger}(k)|\overline{A}(k),\bar{t}^{\ddagger}(k-1),\overline{H}^{\ddagger,*}_{\overline{A}(k-1),\bar{t}^{\ddagger}(k-1)}(k)]$ and $\bar{t}^{\ddagger}_{s}=\bar{t}^{\ddagger}_{s}(K)$. With this replacement, the solution $\widetilde{\psi}$ to the resulting estimating equation remains a CAN estimator of ψ^{*}, when the model $g^{\intercal}[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{H}^{*}(k);\alpha]$ is correct. Once we have obtained a CAN estimator of ψ^{*}, we can estimate $x_{\mathrm{opt}}(V)=x_{\mathrm{opt}}(V,\psi^{*})$ by $x_{\mathrm{opt}}(V,\widehat{\psi})$ as above.

In summary, the above development documents the very strong assumptions required to identify $E_{(G_{x},P^{\intercal})}[Y|V]$ from data generated under $(G^{\intercal},P^{\intercal})$ in the presence of an arbitrary unmodelled nonignorable visit process $G^{\intercal}$ that differs from the visit process of G_{x}. Two particular problems that we have yet to address are: How does one use one's limited knowledge of the population G to which we wish to extrapolate to choose the function $g[T^{\ddagger}(k)|\overline{A}(k),\overline{T}^{\ddagger}(k-1),\overline{T}(k),\overline{H}^{*}(k)]$ needed to implement our estimator and how does one quantify the bias that results from imposing the NNV assumption when that assumption is incorrect. These are very difficult questions with no single 'correct' answer and any answer is lengthy. For this latter reason, we defer our answers to a future paper in which we will also consider in greater detail the alternative approach to the nonignorable visit process problem based on a nonignorable model such as equation (33).

DOI: 10.1002/sim

APPENDIX A

In this appendix, we prove a general theorem that subsumes the special cases treated in Section 6.3. Let $\widetilde{c}(\overline{a},\overline{t},\overline{l})$ be a general function of $(\overline{a},\overline{t},\overline{l})$. We think of $\widetilde{c}(\overline{a},\overline{t},\overline{l})$ as the most general possible utility function. That is, the utility Y^* is given by $Y^*=\widetilde{c}(\overline{A},\overline{T},\overline{L})$. Define $u_{\widetilde{c}}(k,\overline{a},\overline{t},\overline{l})=1$ if $\widetilde{c}(\overline{a},\overline{t},\overline{l})$, as a function of $\overline{l}=(\overline{z},\overline{i})$, depends nontrivially on the $k+1$th component $i(k+1)$of $\overline{i}$ and define $u_{\widetilde{c}}(k,\overline{a},\overline{t},\overline{l})=0$, otherwise. Given a function $\widetilde{c}(\overline{a},\overline{t},\overline{l})$ and a conditional law $g^{\top}(\overline{t}|\overline{L},\overline{A})$, define the sets

$$\overline{\mathscr{B}}^*_{\widetilde{c}}(\overline{l},\overline{a},\overline{t}^*)=\{\overline{t};t(k)\geqslant\max\{t^*(k),u_{\widetilde{c}}(k,\overline{a},\overline{t}^*,\overline{l})\}\text{ for }k=0,\ldots,K+1\}$$

$$\overline{\mathscr{B}}_{\widetilde{c},g^{\top}}(\overline{l},\overline{a},\overline{t}^*)=\{\overline{t}\in\overline{\mathscr{B}}^*_{\widetilde{c}}(\overline{l},\overline{a},\overline{t}^*);g^{\top}(\overline{t}|\overline{l},\overline{a})>0\}$$

and

$$\mathscr{K}(\overline{a},\overline{t},\overline{l})=\{k;t(k)+u_{\widetilde{c}}(k,\overline{a},\overline{t},\overline{l})>0,0\leqslant k\leqslant K\}$$

Theorem A.1

Suppose G and $G^{\intercal}$ are $\overline{O}^{*a}$-SR with the same marginal distribution $P^{\intercal}$ for $\overline{L}_{\overline{\mathscr{A}}}$ and that the NDE assumption holds. Further, suppose that

$$\overline{\mathscr{B}}_{\widetilde{c},g^{\top}}(\overline{L}_{\overline{a}},\overline{a},\overline{t}^*)\text{ is nonempty for all }(\overline{a},\overline{t}^*)\in\overline{\mathscr{J}}_g(\overline{L}_{\overline{\mathscr{A}}})\text{ w.p.1. under }P^{\top}\tag{A1}$$

where $\overline{\mathscr{J}}_g(\overline{L}_{\overline{\mathscr{A}}})$ is the support of $\overline{J}=(\overline{A},\overline{T})$ under $g(\overline{a},\overline{t}^*|\overline{L}_{\overline{\mathscr{A}}})$. Then,

(a) $E_{G,P^{\intercal}}[\widetilde{c}(\overline{A},\overline{T},\overline{L})]$ is identified from data $\overline{O}^*$ generated under $(G^{\top},P^{\intercal})$ and it equals $E_{G^{\top},P^{\intercal}}[U(g,g^{\top})]$ where $U(g,g^{\top})\equiv u(\overline{O}^*;g,g^{\intercal})$ is defined as

$$\frac{\sum_{\overline{t}^*\in\overline{\mathscr{T}}_g(\overline{A},\overline{L})}I(\overline{T}\in\overline{\mathscr{B}}_{\widetilde{c}}(\overline{L},\overline{A},\overline{t}^*))\widetilde{c}(\overline{A},\overline{t}^*,\overline{L})\prod_{k=0}^{K}g[A(k),t^*(k)|\overline{A}(k-1),\overline{t}^*(k-1),\overline{L}^*_{\overline{t}^*(k-1)}(k)]}{\prod_{k=0}^{K}g^{\intercal}[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{L}^*(k)]\prod_{k\in\mathscr{K}(\overline{A},\overline{t}^*,\overline{L})}g^{\intercal}[T(k)|\overline{A}(k),\overline{T}(k-1),\overline{L}^*(k)]}\tag{A2}$$

with $\overline{\mathscr{T}}_g(\overline{A},\overline{L})$ the support of $\overline{T}$ under $g(\overline{t}|\overline{A},\overline{L})$, and

(b) the TPO* assumption of equation (24) holds.

Remark A.0

Note that by G being $\overline{O}^{*a}$-SR, it holds that $g(\overline{t}|\overline{A},\overline{L})=g(\overline{t}|\overline{A},\overline{L}_{\overline{\mathscr{A}}})$ and consequently, $\overline{\mathscr{T}}_g(\overline{A},\overline{L})$ is also the support of $\overline{T}$ under $g(\overline{t}|\overline{A},\overline{L}_{\overline{\mathscr{A}}})$, which we denote by $\overline{\mathscr{T}}_g(\overline{A},\overline{L}_{\overline{\mathscr{A}}})$.

Remark A.1

Note $\widetilde{c}(\overline{A},\overline{t}^*,\overline{L})$ and $\overline{L}^*_{\overline{t}^*(k-1)}(k)$ are both functions of $\overline{O}^*=(\overline{A},\overline{T},\overline{L}^*_{\overline{T}})$ whenever $\overline{T}\in\overline{\mathscr{B}}_{\widetilde{c}}(\overline{A},\overline{t}^*,\overline{L})$. It then follows as stated in the theorem that $U(g,g^{\top})$ is a function of $\overline{O}^*$.

Corollary A.1

If the suppositions of Theorem A.1 hold, $G=G_x=\mathbf{d}_{x=(x_a,x_t)}$, $u_{\widetilde{c}}(k,\overline{a},\overline{t},\overline{l})$ is the zero function, then $E_{G,P^{\intercal}}[\widetilde{c}(\overline{A},\overline{T},\overline{L})]$ is equal to

$$E_{G^{\top},P^{\intercal}}\left[\left\{\frac{\Delta(\overline{T}^*_{x_t})\Delta_{x_a}(\overline{T}^*_{x_t})\widetilde{c}(\overline{A},\overline{T}^*_{x_t},\overline{L})}{\prod_{k=0}^{K}g^{\intercal}[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{L}^*(k)]\prod_{k\in\mathscr{K}(\overline{A},\overline{T}^*_{x_t},\overline{L})}g^{\intercal}[T(k)|\overline{A}(k),\overline{T}(k-1),\overline{L}^*(k)]}\right\}\right]$$

DOI: 10.1002/sim

Furthermore, equation (A1) is equivalent to the assumption that the set

$$\left\{\overline{t}; t(k) \geqslant T_{G_x}(k) \text{ and } \prod_{k=0}^{K} g^{\top}[A_{G_x}(k), t(k) | \overline{A}_{G_x}(k-1), \overline{t}(k-1), \overline{L}^{*}_{\overline{A}_{G_x}(k-1), \overline{t}(k-1)}(k)] > 0\right\}$$

is nonempty w.p.1. under $P^{\intercal}$.

Corollary A.2
If the suppositions of Theorem A.1 hold and $G = G_{x_a}$, with G_x as defined in equations (14) and (15), then $E_{G,P^{\intercal}}[\widetilde{c}(\overline{A}, \overline{T}, \overline{L})]$ is equal to

$$E_{G,G^{\top},P^{\intercal}}\left[\frac{1}{S}\sum_{s=1}^{S}\left[\frac{I(\overline{T} \in \overline{\mathscr{B}}_{\widetilde{c},g^{\top}}(\overline{L}, \overline{A}, \overline{T}^{*}_{s},))\widetilde{c}(\overline{A}, \overline{T}^{*}_{s}, \overline{L})\delta_{x_a}(\overline{A}, \overline{T}^{*}_{s}, \overline{L})}{\prod_{k=0}^{K} g^{\intercal}[A(k) | \overline{A}(k-1), \overline{T}(k-1), \overline{L}^{*}(k)] \prod_{k \in \mathscr{K}(\overline{A}, \overline{T}^{*}_{s}, \overline{L})} g^{\intercal}[T(k) | \overline{A}(k), \overline{T}(k-1), \overline{L}^{*}(k)]}\right]\right]$$

with $E_{G,G^{\top},P^{\intercal}}$ the expectation taken under $f_{G,G^{\top},P^{\intercal}}(\overline{t}^{*}_{1}, \ldots \overline{t}^{*}_{S}, \overline{t}, \overline{a}, \overline{l}) \equiv \{\prod_{s=1}^{S} \widetilde{g}(\overline{t}^{*}_{s} | \overline{a}, \overline{l}, \overline{t})\} f_{G^{\top},P^{\intercal}}(\overline{a}, \overline{t}, \overline{l})$, where $\widetilde{g}(\overline{t}^{*}_{s} | \overline{a}, \overline{l}, \overline{t}) \equiv \prod_{k=0}^{K} g[t^{*}_{s}(k) | \overline{a}(k), \overline{t}^{*}_{s}(k-1), \overline{l}^{*}_{\overline{t}^{*}_{s}(k-1)}(k)]$.

Proof of Theorem A.1
Let $\overline{\mathscr{A}}_{g}(\overline{L}_{\overline{\mathscr{A}}})$ denote the support of $\overline{A}$ under $g(\overline{a} | \overline{L}_{\overline{\mathscr{A}}})$. Then

$$\begin{aligned}
&E_{G,P^{\intercal}}[\widetilde{c}(\overline{A}, \overline{T}, \overline{L})] \\
&= E_{P^{\intercal}}\left[\sum_{(\overline{a},\overline{t}^{*}) \in \overline{\mathscr{J}}_{g}(\overline{L}_{\overline{\mathscr{A}}})} \widetilde{c}(\overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) g(\overline{a}, \overline{t}^{*} | \overline{L}_{\overline{\mathscr{A}}})\right] \\
&= E_{P^{\intercal}}\left[\sum_{(\overline{a},\overline{t}^{*}) \in \overline{\mathscr{J}}_{g}(\overline{L}_{\overline{\mathscr{A}}})} \widetilde{c}(\overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) \prod_{k=0}^{K} g[a(k), t^{*}(k) | \overline{a}(k-1), \overline{t}^{*}(k-1), \overline{L}^{*}_{\overline{a}(k-1), \overline{t}^{*}(k-1)}(k)]\right] \\
&= E_{P^{\intercal}}\left[\sum_{(\overline{a},\overline{t}^{*}) \in \overline{\mathscr{J}}_{g}(\overline{L}_{\overline{\mathscr{A}}})} \widetilde{c}(\overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) \prod_{k=0}^{K} g[a(k), t^{*}(k) | \overline{a}(k-1), \overline{t}^{*}(k-1), \overline{L}^{*}_{\overline{a}(k-1), \overline{t}^{*}(k-1)}(k)]\right. \\
&\quad \left. \times \left\{\sum_{\overline{t} \in \overline{\mathscr{B}}_{\widetilde{c},g^{\top}}(\overline{L}_{\overline{a}}, \overline{a}, \overline{t}^{*},)} \prod_{\{k; t^{*}(k) = u(k, \overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) = 0\}} g^{\intercal}[t(k) | \overline{a}(k), \overline{t}(k-1), \overline{L}^{*}_{\overline{a}(k-1), \overline{t}^{*}(k-1)}(k)]\right\}\right] \\
&= E_{G^{\top},P^{\intercal}}\left\{\sum_{(\overline{a},\overline{t}^{*}) \in \overline{\mathscr{J}}_{g}(\overline{L}_{\overline{\mathscr{A}}})} \sum_{\overline{t} \in \overline{\mathscr{B}}_{\widetilde{c},g^{\top}}(\overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}})} \frac{I(\overline{A} = \overline{a}) I(\overline{T} = \overline{t})}{g^{\intercal}[\overline{a}, \overline{t} | \overline{L}_{\overline{\mathscr{A}}}]}\right. \\
&\quad \times \prod_{\{k; t^{*}(k) = u(k, \overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) = 0\}} g^{\intercal}[t(k) | \overline{a}(k), \overline{t}(k-1), \overline{L}^{*}_{\overline{a}(k-1), \overline{t}^{*}(k-1)}(k)] \\
&\quad \left. \times \widetilde{c}(\overline{a}, \overline{t}^{*}, \overline{L}_{\overline{a}}) \prod_{k=0}^{K} g[a(k), t^{*}(k) | \overline{a}(k-1), \overline{t}^{*}(k-1), \overline{L}^{*}_{\overline{a}(k-1), \overline{t}^{*}(k-1)}(k)]\right\}
\end{aligned}$$

DOI: 10.1002/sim

$$
\begin{aligned}
&= E_{G^\top, P^\intercal}\left\{ \sum_{\overline{t}^* \in \overline{\mathscr{T}}_g(\overline{A},\overline{L})} I[\overline{A} \in \overline{\mathscr{A}}_g(\overline{L}_{\overline{\mathscr{A}}})] \frac{I(\overline{T} \in \overline{\mathscr{B}}_{\widetilde{c}}(\overline{L},\overline{A},\overline{t}^*))}{g^\intercal[\overline{A},\overline{T}|\overline{L}_{\overline{\mathscr{A}}}]} \right. \\
&\quad \times \prod_{\{k;t^*(k)=u(k,\overline{A},\overline{t}^*,\overline{L})=0\}} g^\intercal[T(k)|\overline{A}(k),\overline{T}(k-1),\overline{L}^*(k)] \\
&\quad \left. \times \widetilde{c}(\overline{A},\overline{t}^*,\overline{L}) \prod_{k=0}^{K} g[A(k),t^*(k)|\overline{A}(k-1),\overline{t}^*(k-1),\overline{L}^*_{\overline{A}(k-1),\overline{t}^*(k-1)}(k)] \right\} \\
&= E_{G^\top, P^\intercal}[u(\overline{O}^*; g, g^\intercal)]
\end{aligned}
$$

where the second equality follows because G is $\overline{O}^{*a}$-SR, the third follows because, by the definition of $\overline{\mathscr{B}}_{\widetilde{c},g^\top}(\overline{L}_{\overline{a}},\overline{a},\overline{t}^*)$, and assumption (A1) the expression in set braces equals $\prod_{\{k;t^*(k)=u(k,\overline{a},\overline{t}^*,\overline{L}_{\overline{a}})=0\}}[\sum_{t(k)=0}^{1} g^\intercal[t(k)|\overline{a}(k),\overline{t}(k-1),\overline{L}^*_{\overline{a}(k-1),\overline{t}^*(k-1)}(k)]]$, the fourth follows because

$$
E_{G^\top}\left[\frac{I(\overline{A}=\overline{a})I(\overline{T}=\overline{t})}{g^\intercal[\overline{a},\overline{t}|\overline{L}_{\overline{\mathscr{A}}}]} \middle| \overline{L}_{\overline{\mathscr{A}}} \right] = E_{G^\top}\left[\frac{I(\overline{A}=\overline{a})}{g^\intercal[\overline{a}|\overline{L}_{\overline{\mathscr{A}}}]} \frac{I(\overline{T}=\overline{t})}{g^\intercal[\overline{t}|\overline{L}_{\overline{a}},\overline{a}]} \middle| \overline{L}_{\overline{\mathscr{A}}} \right] = 1
$$

for $\overline{t} \in \overline{\mathscr{B}}_{\widetilde{c},g^\top}(\overline{L}_{\overline{a}},\overline{a},\overline{t}^*)$ by $g^\intercal[\overline{a},\overline{t}|\overline{L}_{\overline{\mathscr{A}}}]$ being $\overline{O}^{*a}$-SR, part (b) of this theorem (to be shown next) and equation (A1), the fifth follows because $I(\overline{T} \in \overline{\mathscr{B}}_{\widetilde{c},g^\top}(\overline{L},\overline{A},\overline{t}^*))=1$ is equivalent to $I(\overline{T} \in \overline{\mathscr{B}}_{\widetilde{c}}(\overline{L},\overline{A},\overline{t}^*))=1$ under $G^\intercal$, and the last equality follows because

$$
\begin{aligned}
&\prod_{\{k;t^*(k)=u(k,\overline{A},\overline{t}^*,\overline{L})=0\}} g^\intercal[T(k)|\overline{A}(k),\overline{T}(k-1),\overline{L}^*(k)]/g^\intercal[\overline{A},\overline{T}|\overline{L}_{\overline{\mathscr{A}}}] \\
&= \prod_{k=0}^{K} g^\intercal[A(k)|\overline{A}(k-1),\overline{T}(k-1),\overline{L}^*(k)] \prod_{\{k;t^*(k)+u(k,\overline{A},\overline{t}^*,\overline{L})>0\}} g^\intercal[T(k)|\overline{A}(k),\overline{T}(k-1),\overline{L}^*(k)]
\end{aligned}
$$

by $g^\intercal[\overline{A},\overline{T}|\overline{L}_{\overline{\mathscr{A}}}]$ being $\overline{O}^{*a}$-SR.

Finally, the TPO* assumption of equation (24) holds because (i) by (A1) for all $\overline{a}$ such that $g(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})>0$, there exists a $\overline{t}$ such that $g^\top(\overline{t}|\overline{L}_{\overline{a}},\overline{a})>0$ and (ii) by $G^\intercal$ being $\overline{O}^{*a}$-SR, $g^\top(\overline{t}|\overline{L}_{\overline{a}},\overline{a})>0 \Rightarrow g^\top_{\overline{A},\overline{T}|\overline{L}_{\overline{a}}}(\overline{a},\overline{t}|\overline{L}_{\overline{a}}) = g^\top(\overline{a},\overline{t}|\overline{L}_{\overline{\mathscr{A}}})>0 \Rightarrow g^\top(\overline{a}|\overline{L}_{\overline{\mathscr{A}}})>0$. □

Proof of Corollary A.1

The numerator of equation (A2) equals $\Delta(\overline{T}^*_{x_2})\Delta_{x_1}(\overline{T}^*_{x_2})\widetilde{c}(\overline{A},\overline{T}^*_{x_2},\overline{L})$ when $G=\overline{\mathbf{d}}_{x=(x_1,x_2)}$.

Moreover, under regime $G_x=\overline{\mathbf{d}}_{x=(x_a,x_t)}$ the support of $(\overline{A},\overline{T})$ under $g(\overline{a},\overline{t}^*|\overline{L}_{\overline{\mathscr{A}}})$ is the singleton $(\overline{A}_{G_x},\overline{T}_{G_x})$. However, arguing as in the proof of part (b) of the theorem

$$
\begin{aligned}
g^\top(\overline{t}|\overline{L}_{\overline{a}},\overline{a})_{|\overline{a}=\overline{A}_{G_x}} > 0 \;\Leftrightarrow\; & g^\top_{\overline{A},\overline{T}|\overline{L}_{\overline{a}}}(\overline{a},\overline{t}|\overline{L}_{\overline{a}})_{|\overline{a}=\overline{A}_{G_x}} = g^\top(\overline{a},\overline{t}|\overline{L}_{\overline{\mathscr{A}}})_{|\overline{a}=\overline{A}_{G_x}} \\
&= \prod_{k=0}^{K} g^\top[a(k),t(k)|\overline{a}(k-1),\overline{t}(k-1),\overline{L}^*_{\overline{a}(k-1),\overline{t}(k-1)}(k)]_{|\overline{a}=\overline{A}_{G_x}} > 0 \qquad \square
\end{aligned}
$$

DOI: 10.1002/sim

Proof of Corollary A.2

Corollary A.2 follows at once from the observation that the numerator of equation (A2) equals

$$\sum_{\overline{t}^* \in \overline{\mathscr{T}}_g(\overline{A},\overline{L})} \left\{ I(\overline{T} \in \overline{\mathscr{B}}_{\widetilde{C},g^\top}(\overline{L},\overline{A},\overline{T}^*)) \Delta_{x_a}(\overline{A},\overline{t}^*,\overline{L}) \right.$$

$$\left. \times \widetilde{c}(\overline{A},\overline{t}^*,\overline{L}) \prod_{k=0}^{K} g[t^*(k)|\overline{A}(k),\overline{t}^*(k-1),\overline{L}^*_{\overline{t}^*(k-1)}(k)] \right\}$$

which can be written as $E_{\widetilde{g}}[\delta_{x_a}(\overline{A},\overline{T}^*,\overline{L}) I(\overline{t} \in \overline{\mathscr{B}}_{\widetilde{C},g^\top}(\overline{L},\overline{A},\overline{T}^*)) \widetilde{c}(\overline{A},\overline{T}^*,\overline{L})|\overline{A},\overline{L},\overline{T}]$. □

ACKNOWLEDGEMENTS

We thank Miguel Hernan for useful advice. This work was partially funded by NIH grants R01 GM48704 and R0132475.

REFERENCES

1. Palella FJ, Deloria-Knoll M, Chmiel JS, Moorman AC, Wood KC, Greenberg AE, Holmberg SD, HIV Outpatient Study (HOPS) Investigators. Survival benefit of initiating antiretroviral therapy in HIV-infected persons in different CD4+ cell strata. *Annals of Internal Medicine* 2003; **138**:620–626.
2. Orellana L, Rotnitzky A, Robins JM. Generalized marginal structural models for estimating optimal treatment regimes. *Technical Report*, Department of Biostatistics, Harvard School of Public Health, 2006.
3. van der Laan MJ. Causal effect models for intention to treat and realistic individualized treatment rules. *Working Paper 203*, U.C. Berkeley Division of Biostatistics Working Paper Series, 2006. (Available from: http://www.bepress.com/ucbbiostat/paper203.)
4. Robins JM. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium on Biostatistics*, Lin DY, Heagerty P (eds). Springer: New York, 2004.
5. Murphy SA. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society*, *Series B* 2003; **65**:331–366.
6. Moodie EEM, Richardson TS, Stephens D. Demystifying optimal dynamic treatment regimes. *Biometrics* 2007; **63**(2):447–455.
7. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic and Clinical Pharmacology and Toxicology* 2006; **98**:237–242.
8. Lanoy M, Mary-Krause P, Tattevin R, Dray-Spira C, Duvivier P, Fischer Y, Obadia F, Lert D, Costagliola C, the Clinical Epidemiology Group. Predictors identified for losses to follow-up among HIV-seropositive patients. *Journal of Clinical Epidemiology* 2006; **59**(8):829–835.
9. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Models in Epidemiology*: *The Environment and Clinical Trials*, Halloran ME, Berry D (eds). Springer: New York, 1999; 95–134.
10. Robins JM. Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics* 1994; **23**:2379–2412.
11. Scharfstein D, Rotnitzky A, Robins JM. Adjusting for non-ignorable drop-out using semiparametric non-response models. *Journal of the American Statistical Association* 1999; **94**:1096–1120.
12. Robins JM, Sued M, Lei-Gomez Q, Rotnitzky A. Performance of double-robust estimators when 'inverse probability' weights are highly variable. Discussion of demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data, by Kang and Schaffer. *Statistical Science* 2007; in press.
13. Wang Y, Petersen M, Bangsberg D, van der Laan M. Diagnosing bias in the inverse probability of treatment weighted estimator resulting from violation of experimental treatment assignment. *U.C. Berkeley Division of Biostatistics Working Paper Series*, 2006. (Available from: http://www.bepress.com/ucbbiostat/paper211/.)

DOI: 10.1002/sim

14. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology—Methodological Issues*, Jewell N, Dietz K, Farewell V (eds). Birkhäuser: Boston, MA, 1992; 297–331 (includes errata sheet).
15. Robins JM, Scharfstein D, Rotnitzky A. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In *Statistical Models in Epidemiology*: *The Environment and Clinical Trials*, Halloran ME, Berry D (eds). Springer: New York, 1999; 1–94.

DOI: 10.1002/sim