University of California, Berkeley

From the SelectedWorks of Maya Petersen

April, 2007

Statistical Learning of Origin-Specific Statically Optimal Individualized Treatment Rules

Mark J. van der Laan
Maya L Petersen, University of California, Berkeley

Available at: http://works.bepress.com/maya_petersen/38/
Statistical Learning of Origin-Specific Statically Optimal Individualized Treatment Rules

Mark J. van der Laan∗ Maya L. Petersen†

∗Division of Biostatistics, School of Public Health, University of California, Berkeley, laan@stat.berkeley.edu
†University of California, Berkeley, mayaliv@gmail.com

Copyright ©2007 The Berkeley Electronic Press. All rights reserved.
Statistical Learning of Origin-Specific
Statically Optimal Individualized Treatment Rules

Mark J. van der Laan and Maya L. Petersen

Abstract

Consider a longitudinal observational or controlled study in which one collects chronological data over time on a random sample of subjects. The time-dependent process one observes on each subject contains time-dependent covariates, time-dependent treatment actions, and an outcome process or single final outcome of interest. A statically optimal individualized treatment rule (as introduced in van der Laan et. al. (2005), Petersen et. al. (2007)) is a treatment rule which at any point in time conditions on a user-supplied subset of the past, computes the future static treatment regimen that maximizes a (conditional) mean future outcome of interest, and applies the first treatment action of the latter regimen. In particular, Petersen et. al. (2007) clarified that, in order to be statically optimal, an individualized treatment rule should not depend on the observed treatment mechanism. Petersen et. al. (2007) further developed estimators of statically optimal individualized treatment rules based on a past capturing all confounding of past treatment history on outcome. In practice, however, one typically wishes to find individualized treatment rules responding to a user-supplied subset of the complete observed history, which may not be sufficient to capture all confounding. The current article provides an important advance on Petersen et. al. (2007) by developing locally efficient double robust estimators of statically optimal individualized treatment rules responding to such a user-supplied subset of the past. However, failure to capture all confounding comes at a price; the static optimality of the resulting rules becomes origin-specific. We explain origin-specific static optimality, and discuss the practical importance of the proposed methodology. We further present the results of a data analysis in which we estimate a statically optimal rule for switching antiretroviral therapy among patients infected with resistant HIV virus.

KEYWORDS: counterfactual, causal inference, double robust estimating function, dynamic treatment regime, history-adjusted marginal structural model, inverse probability weighting
1 Introduction.

Dynamic treatment regimes, also called adaptive treatment strategies (Lavori and Dawson (2000)), assign a treatment decision at each time point based on a subject’s observed past up till that time point. Examples of dynamic regimes include strategies in which a treatment course is continued only as long as adverse effects are absent or markers of response are favorable, as well as strategies in which treatment is initiated based on the severity of symptoms, presence of risk factors, and/or degree of response to alternative interventions. The ability to respond to changes in a subject’s observed past differentiates dynamic treatment regimes from static treatment regimens; static regimens allow treatment to change over time, but not in response to changes in subject covariates. The adaptive nature of dynamic regimes makes them particularly relevant to clinical practice, as well as to many other applications, and has motivated a large body of research aimed at their estimation.

Substantial prior work has discussed estimation of the mean (or some other parameter) of an outcome under a candidate dynamic regime (or a user-supplied set of such regimes) (for example, Robins (1986), Robins (1993), Robins (1997), Murphy et al. (2002), Murphy (2003), Hernan et al. (2006), van der Laan and Petersen (2007)). Estimation of the optimal dynamic regime, or the treatment rule that, if followed, would result in the highest mean outcome for the population, has also been addressed by multiple authors spanning several disciplines. Notably, Murphy (2003) and Robins (1994, 2004) proposed related methods to estimate the optimal dynamic treatment regime without assuming a model on the entire data-generating distribution. The methods of both Murphy and Robins rely on the counterfactual framework for causal inference (Robins (1986), Robins (1987)) and the concept of potential or counterfactual outcomes (Neyman (1990), Rubin (1978)). Murphy (2003) and Moodie et al. (2007) provide more extensive reviews of the literature surrounding the estimation of optimal dynamic treatment regimes.

In this manuscript we discuss estimation of statically optimal treatment rules, an alternative type of dynamic regime. A statically optimal individualized treatment rule conditions at each time point on a user-supplied subset of the past, computes the future static treatment regimen that maximizes a (conditional) mean future outcome of interest, and applies the first treatment action of the latter regimen. Repetition of the process at subsequent time points results in a rule that is dynamic (or individualized), in that the treatment plan to which a subject is assigned can change over time in response to the subject’s observed past.

An optimal dynamic regime is, by definition, the treatment rule that results in the best expected outcome. The mean outcome achieved under a statically optimal rule will thus generally be inferior to that achieved under the truly
optimal rule; van der Laan et al. (2005) provide a simulation demonstrating an example of the potential sub-optimality of statically optimal rules. However, estimation of statically optimal rules may offer several practical advantages. The statically optimal rule is a less ambitious parameter than the optimal dynamic regime, suggesting that statically optimal rules can be estimated with either a greater degree of precision or fewer model assumptions than optimal dynamic regimes. In addition, estimation of the statically optimal rule can be accomplished by estimating a series of static treatment effects, allowing a familiar set of statistical tools and software to be used.

Statically optimal individualized treatment rules were first discussed by van der Laan et al. (2005) in an article introducing (observed) history-adjusted marginal structural models (HA-MSM). However, as made explicit in Petersen et al. (2007b), the optimality of rules estimated by HA-MSM can depend on the way in which treatment was assigned in the observed data (the observed treatment mechanism), and as a result, HA-MSM-derived rules can fail to select the optimal future treatment plan (given individual covariate values) if applied to an equivalent population with a different treatment mechanism. Dependence on the observed treatment mechanism implies, in particular, that a rule can fail to select the statically optimal treatment at each time point in a setting where the rule itself has been applied to the population beginning at baseline (as, for example, would occur if the rule were tested in a clinical trial). As a result, Petersen et al. (2007b) refine the definition of static optimality to specify that a statically optimal rule must not depend on the observed treatment mechanism. They further point out that HA-MSM-derived rules are only truly statically optimal if the covariates in the history-adjusted marginal structural model are sufficient to capture all confounding of past treatment on outcome.

The statically optimal rules discussed by Petersen et al. (2007b) have the attractive property that they select the optimal future treatment plan regardless of how past treatment has been assigned. As a result, these rules will retain their static optimality if applied to an exchangeable population that has been following the rule in question (as would occur in the context of a clinical trial) or if applied to an exchangeable population that has been following some unknown treatment mechanism (as could occur in the course of clinical practice). In order to achieve this generality, however, the statically optimal rules must incorporate sufficient covariate history to control for confounding of past treatment history on outcome.

Due to limited resources, practical sensibility, and reliability, clinical trials are generally forced to focus on candidate individualized treatment rules that only respond to a small number of user-supplied bio-markers or other relevant measurements. The current article provides an important advance by presenting models and corresponding locally efficient double robust estima-
tors of individualized treatment rules responding to a \textit{user-supplied} subset of a subject’s past. However, failure to incorporate sufficient covariates to control for confounding has a price: the static optimality of the resulting rules is \textit{origin-specific}. In other words, the individualized treatment rules described in this paper will not necessarily choose the first action of the optimal static treatment at any given time for subjects who have not followed the statically optimal course through that time (since a specified origin). Thus the origin-specific statically optimal rules described in this paper may be evaluated appropriately in clinical trials, where the rule itself is applied beginning at enrollment.

Under assumptions discussed below, statically optimal rules can be estimated using observational cohort data. As an illustration, in Section 6 we provide an example of a statically optimal rule derived from an observational cohort of HIV-infected individuals. Subjects become eligible for the analysis when they virologically fail a combination antiretroviral therapy regimen. The goal is to identify a rule for deciding when to modify a subject’s failing regimen. Waiting to modify therapy can result in disease progression, reflected in CD4 T cell count declines, as well as in the accumulation of viral mutations that confer resistance to potential future treatments. However, early therapy modification runs the risk of prematurely depleting future treatment options (Deeks (2003)).

In this research application, a static regimen would simply assign a fixed modification time to each subject, possibly conditional on the subjects’ covariates at time of virologic failure. It is likely, however, that a more clinically beneficial strategy would base the decision to modify drug regimens on markers of disease progression and viral evolution over the course of treatment with the failing therapy, or in other words, would employ a dynamic regime in deciding when to switch. In the data example presented in this article, we thus aim to estimate a statically optimal rule of the following type: if a subject has not yet modified therapy, assess, based on a subset of subject’s observed past, whether waiting longer to switch is estimated to result in a poorer expected outcome than switching immediately. If so, assign the subject to switch therapy at the next time point; otherwise, wait until the following time point and re-evaluate the expected consequences of waiting longer to switch. Such a rule is \textit{dynamic} in that the decision whether to modify therapy is based on updated values of a patient’s covariates. For example, the rule might indicate that a subject should switch therapy when his CD4 T cell count falls below a given threshold. The rule is \textit{statically optimal}, in that the rule assigns at each time point the first action of the future static regimen that is estimated to result in the best expected outcome.

As we illustrate in the data example, the methods introduced in this paper can be implemented using standard software as a series of inverse probabil-
ity weighted (IPTW) regressions. However, the parameter estimated is novel—we estimate the counterfactual history-adjusted mean outcome at each time point, conditional on a user-supplied subset of the observed past. Petersen et al. (2007b) used the same data example to estimate a statically optimal treatment rule based on observed history-adjusted marginal structural models (HA-MSM) that incorporated the treatment mechanism as a covariate. In contrast, here we illustrate an alternative approach to the estimation of an origin-specific statically optimal treatment rule based on counterfactual history-adjusted marginal structural models (CF-HA-MSM).

1.1 Organization of article.

In Section 2 we define the observed data and a statistical framework/model for the data generating distribution. This framework is used to formally define the origin-specific statically optimal treatment rule and to introduce the counterfactual history-adjusted mean outcome as a novel parameter. We show that the counterfactual history-adjusted mean implies our desired origin-specific statically optimal rule, and thus demonstrate the rule’s identifiability. Our model corresponds with only modelling the counterfactual history-adjusted mean outcome, while leaving all nuisance parameters unspecified. In Section 3 we derive the class of all estimating functions in this model for the data generating distribution. The estimating functions for the parameter of interest are indexed by nuisance parameters: the treatment mechanism and the conditional distributions of the covariates given the past. However, the estimating functions are orthogonal to these nuisance parameters, and as a result remain unbiased if one (not both) of the two nuisance parameters is misspecified. Such estimating functions are referred to as double robust inverse probability of treatment weighted (IPTW) estimating functions (van der Laan and Robins (2003)). By assuming models for the two nuisance parameters, which represent factors in the likelihood of the observed data, and estimating the parameters accordingly using maximum likelihood procedures, we obtain double robust IPTW estimators of the origin-specific statically optimal individualized treatment rules. Section 4 discusses statistical inference. In Section 5 we review conditions under which the origin-specific statically optimal treatment rule estimated using the counterfactual history-adjusted mean, as presented in this paper, is equivalent to the statically optimal rule defined by Petersen et al. (2007b). Section 6 presents the results of a data analysis based on the HIV example described above, in which we estimate origin-specific statically optimal rules for deciding when to switch antiretroviral therapy. Section 7 discusses various generalizations of practical interest.
2 The origin-specific statically optimal individualized treatment rule and counterfactual history-adjusted mean.

In this section we review the counterfactual framework for causal inference (Robins (1986, 1987). This framework is used to formally define the the counterfactual history-adjusted mean outcome and resulting origin-specific statically optimal individualized treatment rule as parameters of the data generating distribution.

2.1 The statistical framework.

The observed data structure on a randomly sampled subject is defined as a missing data structure on the set of treatment-specific time-dependent processes, where each treatment-specific process represents the counterfactual data we would have observed on the subject, if, possibly contrary to the fact, the subject would have followed that particular treatment regimen. In addition, we make the sequential randomization assumption (defined below), allowing us to identify the probability distribution of the counterfactual processes, and thereby allowing us to learn causal parameters of interest from the observed data.

Representation of the observed data as a missing data structure.

The chronological data structure observed on a randomly sampled subject/experimental unit is given by

$$O = (L(0), A(0), L(1), A(1), \ldots, L(K), A(K), L(K + 1)),$$

where $L(t)$ is data collected at time $t$, $A(t)$ is treatment at time $t$ assigned after $L(t)$, and $K + 1$ denotes the maximal follow-up time. It is assumed that for each subject and each possible treatment regimen $\bar{a} = (a(0), \ldots, a(K)) \in \mathcal{A}$ there exists a process $L_{\bar{a}} = L_{\bar{a}}(0), \ldots, L_{\bar{a}}(K+1)$, and that the observed process $L$ on the subject is the treatment-specific process indexed by the treatment regimen the subject actually took: $L = L_{\bar{A}}$. Here $\mathcal{A}$ denotes the support of the random treatment process $\bar{A} = (A(0), \ldots, A(K))$. Thus, we define a collection of treatment-specific time-dependent processes $(L_{\bar{a}}(t) : 0 \leq t \leq K+1)$, indexed by a treatment regimen $\bar{a} = (a(0), \ldots, a(K))$.

We assume that there is an experiment resulting in the observation of all these random processes, and we denote this random variable with $X = (L_{\bar{a}} : \bar{a} \in \mathcal{A})$ (in the censored data literature, $X$ is often referred to as the full data). Let $P_{X_0}$ denote the probability distribution of this collection $X$ of treatment-specific processes. It is assumed that $L_{\bar{a}}(t) = L_{\bar{a}(t-1)}(t)$, where we use the notation $\bar{a}(t) = (a(0), \ldots, a(t))$. The latter assumption is implied by the time-ordering assumption that $A(t)$ occurs after $L(t)$ and before $L(t+1)$. Let $Y_{\bar{a}}(t)$
be a treatment-specific outcome process, and $S_{\bar{a}}(t)$ denote the components of $L_{\bar{a}}(t)$ on which the individualized treatment rule is based. To conclude, the observed data structure $O$ on a randomly sampled subject can be represented as

$$O = (\bar{A}, L_{\bar{A}}),$$

or equivalently,

$$O = (L(0), A(0), L_{A(0)}(1), A(1), \ldots, L_{\bar{A}(K-1)}(K), A(K), L_{\bar{A}(K)}(K+1)).$$

Thus the observed random process $O$ is a missing data structure on the full data structure $X$.

Let $G_0(\cdot | X)$ denote the conditional probability distribution of $\bar{A}$ given $X$, which is called the treatment mechanism since it determines how treatment is assigned. We note that the observed data structure $O$ is a random variable with a probability distribution $P_{X_0,G_0}$ implied by $P_{X_0}$ and $G_0(\cdot | X)$. We assume that we observe $n$ i.i.d. copies $O_1, \ldots, O_n$ of this random variable $O \sim P_{X_0,G_0}$.

**Sequential randomization assumption.** In order to be able to identify parameters of the probability distribution of $X$ from the probability distribution of $O$, we assume the coarsening at random assumption on the missingness/treatment mechanism $G_0(\cdot | X)$. That is,

$$g_0(\bar{a}(K) | X) \equiv \prod_{t=0}^{K} \Pr(A(t) = a(t) | \bar{A}(t-1) = \bar{a}(t-1), X)$$

$$= \prod_{t=0}^{K} \Pr(A(t) = a(t) | \bar{A}(t-1), L_{\bar{A}}(t)).$$

In the context of causal inference this is often referred to as the sequential randomization assumption.

In the following sections we first define the origin-specific statically optimal treatment rule. We then show that this rule is identified by the counterfactual history-adjusted mean, a parameter of $P_{X_0}$.

### 2.2 The origin-specific statically optimal individualized treatment rule.

We begin by defining counterfactuals indexed by individualized treatment rules. For a given rule $d$ with treatment assignment at time $t$ a function of $\bar{L}(t)$, the corresponding counterfactual process $L_d$ is a random variable defined by the deterministic function of the collection of treatment-specific processes $X = (L_{\bar{a}} : \bar{a})$ and the rule $d$ given by $L_d \equiv L_{\bar{a}(d,X)}$, where $\bar{a}(d, X)$ is the treatment vector assigned by rule $d$ for a subject with full data structure $X$. Thus,
van der Laan and Petersen: Origin-Specific Statically Optimal Individualized Treatment Rules

counterfactual processes/random variables \( L_d \) indexed by any individualized treatment rule \( d \) are also well defined, given our definition of static treatment regimen-specific processes \( L_d \) for all \( \bar{a} \).

Using this definition of a counterfactual covariate process indexed by an individualized treatment rule, we provide the following definition of an origin-specific statically optimal treatment rule. \( Y_d(t, m) \) is a future rule-specific (counterfactual) outcome such as, for example, \( Y_d(t+m) \) for some user-supplied integer \( m \geq 0 \). In order to keep some generality, we allow this outcome \( Y_d(t, m) \) to be any function of the future outcome process \( (Y_d(s) : s \geq t) \) starting at time \( t \), indexed by a scalar \( m \). We let \( d_t \) denote the function assigning the treatment decision under rule \( d \) at time \( t \), \( \bar{d}_t = (d_0, d_1, \ldots, d_t) \) and \( a(t) \) denote a future static treatment regimen beginning at time \( t \), \( a(t) = (a(t), a(t+1), \ldots, a(K)) \). We use \( K(m) \) to denote the last time point \( t \) for which \( Y(t, m) \) is defined. For example, if \( Y(t, m) = Y(t + m) \), then \( Y(t, m) \) is only defined for \( t = 0, \ldots, K(m) = K + 1 - m \).

**Definition 1** Below we define an origin-specific statically optimal dynamic treatment rule

\[
\bar{d}_K \equiv \bar{d}_K(\bar{S}_d(K)) = (d_0(S(0)), d_1(\bar{S}_d(1)), \ldots, d_K(\bar{S}_d(K))),
\]

where each function \( S_d(t) \rightarrow d_t(\bar{S}_d(t)) \) describes how \( A(t) \) is assigned in response to \( \bar{S}_d(t) \) for all \( t = 0, \ldots, K \).

This rule \( d \) is defined by the following algorithm:

\[
\begin{align*}
a^*(0 \mid S(0)) &= \arg \max_{a(0)} E(Y_{d_{(0)}a(0)} \mid S(0)) \\
d_0 &= a^*(0 \mid S(0))(1) \\
a^*(1 \mid \bar{S}_d(1)) &= \arg \max_{a(1)} E(Y_{\bar{d}_0a(1)} \mid \bar{S}_d(1)) \\
d_1 &= a^*(1 \mid \bar{S}_d(1))(1) \\
a^*(t \mid \bar{S}_d(t)) &= \arg \max_{a(t)} E(Y_{\bar{d}_{t-1}a(t)} \mid \bar{S}_d(t)) \\
d_t &= a^*(t \mid \bar{S}_d(t))(1), \quad t = 2, \ldots, K(m),
\end{align*}
\]

where \( Y_{\bar{d}_{t-1}a(t)} \equiv Y_{\bar{d}_{t-1}a(t)}(t, m) \) is the counterfactual random variable corresponding to following the treatment rule \( d \) from time \( 0 \) to time \( t - 1 \), and then static treatment regimen \( a(t) \). Similarly, \( \bar{S}_d(t) \equiv \bar{S}_{\bar{d}_{t-1}}(t) \) is the counterfactual random variable corresponding to following the treatment rule \( d \) from time \( 0 \) to time \( t - 1 \). \( a^*(t \mid \bar{S}_d(t))(1) \) is used to denote the first component of the optimal static treatment regimen \( a^*(t \mid \bar{S}_d(t)) \).
For time points \( t = K(m) + 1, \ldots, K \), the rule \( d_t \) assigns the next action of the static regimen

\[
\tilde{a}^*(K(m)) \mid \tilde{S}_d(K(m))) = \arg \max_{\tilde{g}(K(m))} E(Y_{\tilde{d}_{K(m) - 1}, \tilde{g}(K(m))} \mid \tilde{S}_d(K(m)))
\]

Note that the origin-specific statically optimal rule must assign treatment deterministically at every time point. Thus, while in many settings there may be several choices of \( \tilde{a}(t|\tilde{S}_d(t)) \) that optimize the expected outcome (i.e. the \( \arg \max_{\tilde{a}(t)} \) is not unique), in this case the user must specify a deterministic way of choosing between these choices.

The origin-specific statically optimal treatment rule of Definition 1 satisfies the following property: it selects, at each time point \( t \), the initial treatment of the future static regimen \( \tilde{a}(t) \) which optimizes the expected outcome \( Y(t, m) \), given the history \( \tilde{S}(t) \), in the world where treatment up till that time point corresponds to following the statically optimal treatment rule \( d \) (and thus \( \tilde{S}(t) = \tilde{S}_d(t) \) and the counterfactual outcome is also indexed by \( \tilde{d}_{t-1} \)). We specify that the rule is origin-specific to clarify that, if applied to a population with an identical full data generating distribution, the rule will assign the optimal future static regimen at each time point if past treatment (beginning at baseline, or the origin) has been assigned according to the rule itself. In other words, the rule will not necessarily choose the optimal future static treatment at any given time for subjects who have not followed the statically optimal course through that time. Thus the rule is appropriate for evaluation in the context of a clinical trial, where the rule itself is applied beginning at enrollment. The origin-specific statically optimal rule can be distinguished from the statically optimal rule defined by Petersen et al. (2007b), in that the latter selects the optimal future static regimen at each time point regardless of how past treatment has been assigned.

### 2.3 The counterfactual history-adjusted mean outcome and corresponding treatment rule.

In this section, we define a novel parameter of the full data generating distribution: the counterfactual history-adjusted mean outcome. This parameter is then used to define a specific individualized treatment rule (Definition 2). Finally, we show that the treatment rule based on the counterfactual history-adjusted mean (presented in Definition 2) is equivalent to the origin-specific statically optimal treatment rule (presented in Definition 1). Thus the counterfactual history-adjusted mean outcome is demonstrated to identify the origin-specific statically optimal treatment rule of interest.

The counterfactual history-adjusted mean outcome is defined as

\[
E(Y_{\tilde{a}(t-1), \tilde{g}(t)}(t, m) \mid \tilde{S}_{\tilde{a}(t-1)}(t))
\]
van der Laan and Petersen: Origin-Specific Statically Optimal Individualized Treatment Rules

This counterfactual history-adjusted mean outcome provides us with the following individualized treatment rule:

Definition 2 Define

\[ \theta_0(t, a(t) \mid \bar{a}(t-1), \bar{s}(t)) \equiv E(Y_{\bar{a}(t-1),a(t)} \mid \bar{S}_{\bar{a}(t-1)}(t) = \bar{s}(t)). \]

We define the following treatment rule:

\[ \tilde{d}_K(\theta_0)(\bar{S}(K)) = (d_0(S(0)), d_1(\bar{S}(1)), \ldots, d_K(\bar{S}(K))), \]

where each function \( \bar{S}(t) \rightarrow d_t(\bar{S}(t)) \) describes how \( A(t) \) is assigned in response to \( \bar{S}(t) \) for all \( t = 0, \ldots, K \).

This rule \( \tilde{d}_K(\theta_0) \) is defined by the following algorithm applied to \( \bar{S}(K) = (S(0), \ldots, S(K)) \):

\[
\begin{align*}
\bar{a}^*(0 \mid S(0)) &= \arg \max_{\bar{a}(0)} \theta(0, \bar{a}(0) \mid S(0)) \\
\vspace{1em}
d_0(S(0)) &= \bar{a}^*(0 \mid S(0))(1) \\
\vspace{1em}
\bar{a}^*(1 \mid \bar{S}(1)) &= \arg \max_{\bar{a}(1)} \theta(1, \bar{a}(1) \mid d_0(S(0)), \bar{S}(1)) \\
\vspace{1em}
d_1(\bar{S}(1)) &= \bar{a}^*(1 \mid \bar{S}(1))(1) \\
\vspace{1em}
\bar{a}^*(t \mid \bar{S}(t)) &= \arg \max_{\bar{a}(t)} \theta(t, \bar{a}(t) \mid \bar{d}_{t-1}(\bar{S}(t-1)), \bar{S}(t)) \\
\vspace{1em}
d_t(\bar{S}(t)) &= \bar{a}^*(t \mid \bar{S}(t))(1), \ t = 2, \ldots, K(m).
\end{align*}
\]

Here we use the notation \( \bar{d}_{t-1}(\bar{S}(t-1)) \equiv (d_0(S(0)), d_1(\bar{S}(1)), \ldots, d_{t-1}(\bar{S}(t-1))) \) for the first \( t-1 \) components of the dynamic treatment rule \( d \) applied to \( \bar{S}(t-1) \), where we note that \( \bar{d}_t(\bar{S}(t)) \) corresponds to a specific \( \bar{a}(t) \). As above, \( \bar{a}^*(t \mid \bar{S}(t))(1) \) denotes the first component of the optimal static treatment regimen \( \bar{a}^*(t \mid \bar{S}(t)) \). And as above, for time points \( t = K(m) + 1, \ldots, K \), the rule \( \bar{d}_t \) assigns the next action of the static regimen

\[
\begin{align*}
\bar{a}^*(K(m) \mid \bar{S}(K(m))) &= \arg \max_{\bar{a}(K(m))} \theta(t, \bar{a}(K(m)) \mid \bar{d}_{K(m)-1}(\bar{S}(K(m)-1)), \bar{S}(K(m)))
\end{align*}
\]

We now show that, once we condition on the covariates on which the treatment rule depends, \( E(Y_{\bar{a}(t-1),a(t)} \mid \bar{S}_{\bar{a}(t-1)}(t) = \bar{s}(t)) \) is equal to \( E(Y_{\bar{a}(t-1),a(t)} \mid \bar{S}_{\bar{a}(t-1)}(t) = \bar{s}(t)) \) at a particular \( \bar{a}(t - 1) \), and thus the origin-specific statically optimal regimen of interest (Definition 1) is identified by applying Definition 2 to the counterfactual history-adjusted mean (1). This is because, as shown in Definition 1, given \( \bar{s}(t) \), the origin-specific statically optimal treatment rule applied at time \( t \) corresponds to a deterministic choice of \( a(t) \), \( t = 0, \ldots, K \). We present these results formally as a Lemma:
Lemma 1 Define
\[ \theta_0(t, A(t) | \bar{a}(t-1), \bar{s}(t)) = E(Y_{\bar{a}(t-1)A(t)}(t, m) | \bar{S}_{\bar{a}(t-1)}(t) = \bar{s}(t)), \]

Given a dynamic treatment rule \( \bar{S}(K) \rightarrow d(\bar{S}(K)) \), we have
\[ E(Y_{\bar{d}_{t-1}A(t)}(t, m) | \bar{S}_d(t) = \bar{s}(t)) = \theta(t, A(t) | \bar{a}_{\bar{d}_{t-1}}(t-1), \bar{s}(t)), \quad (2) \]
where \( Y_{\bar{d}_{t-1}A(t)}(t, m) \) is the counterfactual random variable corresponding with following rule \( d \) from time 0 till \( t-1 \), and subsequently, following the static treatment regimen \( a(t) \). Note that \( Y_{\bar{d}_{t-1}A(t)}(t, m) \) is a random variable defined as a deterministic function of \( X \), the rule \( d \), and static treatment \( A(t) \). We use \( \bar{a}_{\bar{d}_{t-1}} \) to denote the treatment history (through time \( t \)) corresponding to applying rule \( d \) to \( \bar{s}(t-1) \).

Proof Given \( \bar{S}_d(t) = \bar{S}_{\bar{d}_{t-1}}(t) = \bar{s}(t) \), we have that \( \bar{d}_{t-1} = (a(0), \ldots, a(t-1)) \) for some fixed \( \bar{a}_{\bar{d}_{t-1}}(t-1) \) defined by \( \bar{s}(t) \). Thus, given \( \bar{S}_d(t) = \bar{s}(t) \), with probability equal to 1 we have \( Y_{\bar{d}_{t-1}A(t)}(t, m) = Y_{\bar{a}_{\bar{d}_{t-1}}A(t)}(t, m) \) and \( \bar{S}_d(t) = \bar{S}_{\bar{a}_{\bar{d}_{t-1}}(t-1)}(t) \): that is, counterfactuals indexed by dynamic treatment regimens are identical to counterfactuals indexed by a corresponding static treatment regimen, which proves the result (2). \( \square \)

Origin-specific static optimality of \( d(\theta_0) \). The importance of this identity (2) is established as follows. Suppose that treatment decisions \( A(0), \ldots, A(t-1) \) have been assigned according to the treatment rule \( d = d(\theta_0) \) so that \( A(0) = d_0(S(0)), A(1) = d_1(\bar{S}_d(1)), \ldots, A(t-1) = d_{t-1}(\bar{S}_d(t-1)) \), and that we are now confronted with making a treatment decision at time \( t \): thus, we are given the treatment past \( \bar{A}(t-1) = \bar{d}_{t-1}(\bar{S}_d(t-1)) \) and covariate past \( \bar{S}_d(t) \) in the world in which we have been applying rule \( d = d(\theta_0) \). We want to show that the origin-specific statically optimal treatment decision at time \( t \), given \( \bar{S}_d(t) \), is now precisely given by \( d_t(\bar{S}_d(t)) \) with \( d = d(\theta_0) \). The origin-specific statically optimal treatment decision at time \( t \), given \( \bar{S}_d(t) \), is defined by optimizing the wished expected outcome \( E(Y_{\bar{d}_{t-1}A(t)}(t, m) | \bar{S}_d(t) = \bar{s}(t)) \) over all static future treatment regimens \( \bar{A}(t) \), and carrying out the first component of this latter treatment regimen. By the previous lemma applied to \( d = d(\theta_0) \), it follows that at time \( t \), optimizing the wished expected outcome \( E(Y_{\bar{d}_{t-1}A(t)}(t, m) | \bar{S}_d(t) = \bar{s}(t)) \) over all statically future treatment regimens \( \bar{A}(t) \) is equivalent with optimizing \( \theta_0(t, \bar{A}(t) | \bar{a}_{\bar{d}_{t-1}}(t-1), \bar{s}(t)) \) over \( \bar{A}(t) \). This proves that indeed the origin-specific statically optimal treatment decision at time \( t \), given \( \bar{S}_d(t) \), is now precisely given by \( d_t(\bar{S}_d(t)) \). Thus, if treatment decisions are assigned deterministically according to rule \( d(\theta_0) \), then it follows that at each point in time \( t \), given the observed \( \bar{S}(t) = (\bar{S}_d(t)) = \bar{s}(t) \), our treatment decision at time \( t \) equals the first treatment in the future treatment regimen maximizing
the mean outcome of $Y_{\bar{a}_t,(t,m)}(t,m)$, given $\bar{S}(t) = \bar{s}(t)$, over all future static treatment regimens $\bar{a}(t)$. This proves that indeed the rule $d(\theta_0)$ is an origin-specific statically optimal treatment rule, and thus that estimates of this rule $\hat{d}(\theta_0)$ are potential candidates for treatment regimens to be evaluated in clinical trials.

In this section we have established that the counterfactual history-adjusted mean $\theta_0$ identifies the origin-specific statically optimal dynamic treatment rule, and illustrated why such a rule is of interest. In the next section we discuss estimation of $\theta_0$, and thus estimation of the origin-specific statically optimal dynamic treatment rule $d(\theta_0)$.

2.4 A model for the counterfactual history-adjusted mean.

In order to deal with the curse of dimensionality, we will assume a model for our parameter of interest $\theta(P_X)(t, a(t) | \bar{a}(t-1), \bar{s}(t)) = E_{P_X}(Y_{\bar{a}(t-1),\bar{a}(t)}(t,m) | \bar{S}_{\bar{a}(t-1)}(t) = \bar{s}(t))$ of $P_X$:

$$\theta_{P_X}(t, a(t) | \bar{a}(t-1), \bar{s}(t)) = m_{\beta(P_X)}(t, a(t) | \bar{a}(t-1), \bar{s}(t))$$

for some parametrization $(m_\beta : \beta)$ indexed by a Euclidean parameter $\beta$. Let $\beta_0 = \beta(P_{X0})$ denote the true parameter value of $\beta$. Note that we could also extend the definition of this parameter to the nonparametric model consisting of all full data distributions $P_X$ so that, if this model (3) is wrong, then $\beta_0$ can be interpreted as a summary measure of interest of the true $\theta_0$, in the same manner as we might interpret a linear regression fit as a summary measure of the true underlying regression curve.

The CF-HA-MSM parameter of interest $\theta$ can be contrasted with the parameter of interest estimated by (observed) history-adjusted models (HA-MSM): $E(Y_{\bar{A}(t-1),\bar{a}(t,m)} | \bar{S}_{\bar{A}(t-1)}(t), \bar{A}(t-1))$. The former parameter is indexed by a counterfactual treatment regimen, while the latter is indexed by the observed treatment up till time $t$, and a fixed counterfactual regimen only after time $t$.

2.5 Model for the observed data.

Because of the sequential randomization assumption, the density of the data structure $O$ can be factorized into a $P_{X0}$-part and $G_0$-part as follows:

$$p_{P_{X0},G_0}(O) = Q_{X0}(O)g_0(\bar{A}(K) | X),$$

where the $P_{X0}$-part of the density is defined as

$$Q_0(O) = Q_{X0}(O) \equiv \prod_{t=0}^{K+1} Pr(L(t) | \bar{L}(t-1), \bar{A}(t-1)).$$
We derive the class of estimating functions for $\beta$ in the model for the observed data structure $O$ only assuming (3). This approach is based on the general estimating function methodology of Robins and Rotnitzky (1992) and van der Laan and Robins (2003). The result is a class of double robust inverse probability of treatment weighted estimating functions for $\beta$ indexed by nuisance parameters $g_0$ and $Q_0$, where the estimating functions remain unbiased at $\beta_0$ if one (but not both) of these two nuisance parameters is misspecified. That is, it is not possible to construct consistent estimators of $\beta_0$ without either consistently estimating $Q_0$ or consistently estimating the treatment mechanism $g_0$. As a consequence, beyond the sequential randomization assumption and the model for $\theta_0$, we either need a model $G$ for $g_0$, or a model $Q$ for $Q_{X0}$. (Alternatively, we can assume the union model which states that either $g_0 \in G$ or $Q_{X0} \in Q$.) Given these models, we assume that valid estimators $g_n$ of $g_0$ according to model $G$ and $Q_n$ of $Q_0$ according to model $Q$ are provided. For example, in the case that the models are small enough, $g_n$ and $Q_n$ could be maximum likelihood estimators:

$$g_n = \arg \max_{g \in G} \sum_{i=1}^{n} \log g(\bar{A}_i | X_i)$$

$$Q_n = \arg \max_{Q \in Q} \sum_{i=1}^{n} \log Q(O_i).$$

If the models are large, then it is typically necessary to use a sieve-based maximum likelihood estimator which involves selection of sub-models of $Q$ and/or $G$.

### 2.6 Identifiability of the statically optimal individualized treatment regimen.

In order to identify the statically optimal individualized treatment regimen $d(\theta_0)$ from the observed data probability distribution we need to be able to identify $\theta_0(t, a(t) | \bar{a}(t-1), \bar{s}(t)) \equiv E_0(Y_{\bar{a}(t-1), \bar{s}(t)}(t, m) | S_{\bar{a}(t-1)}(t) = \bar{s}(t))$, for all $\bar{a} \in \mathcal{A}$. Thus, it suffices to identify the joint distribution $(Y_{\bar{a}}, S_{\bar{a}})$ for all treatment regimens $\bar{a} \in \mathcal{A}$. This requires the so-called experimental treatment assignment assumption (ETA) given by: for all $\bar{a} \in \mathcal{A}$

$$g(\bar{a}(K) | X) > 0 \ P_{X_0} \text{-a.e.} \quad (4)$$

Equivalently, at each time $t \leq K$, we need that for all possible observed histories $\bar{A}(t-1) = \bar{a}(t-1), \bar{L}(t) = \bar{l}(t)$

$$P(A(t) = a(t) | \bar{A}(t-1) = \bar{a}(t-1), \bar{L}(t) = \bar{l}(t)) > 0 \text{ for all } a(t)$$
compatible with $\bar{a}(t-1)$ in the sense that $\bar{a}(t)$ is a possible regimen (i.e., $\bar{a}(t)$ is in the support of $\bar{A}(t)$). Under the ETA, we have that the probability distribution of the treatment-specific counterfactual process $L_{\bar{a}}$ is given by

$$P(L_{\bar{a}} = l) = \prod_{t=0}^{K+1} Pr(L(t) = l(t) \mid \bar{L}(t-1) = \bar{l}(t-1), \bar{A}(t-1) = \bar{a}(t-1)).$$ (5)

This formula (5) for the probability distribution of $X_{\bar{a}}$ was named the $G$-computation formula by Robins (Robins (1999)). That is, the $\bar{a}$-specific marginal distribution of $X$ is identified by a simple intervention on the $Q_X$-part of the density of $O$. One can evaluate this probability distribution by simulating many realizations from this $\bar{a}$-specific density of a time-dependent process $(L(0), \ldots, L(K+1))$, which, in particular, provides us with a Monte-Carlo approximation of the probability distribution of $(Y_{\bar{a}}, S_{\bar{a}})$. Given our model (3), a large collection of realizations $(Y_{\bar{a}}, S_{\bar{a}})$ can now also be used to obtain the corresponding approximation for $\beta_0$. Application of this Monte-Carlo approach to a maximum likelihood-based estimate of $Q_X$ results in a likelihood-based estimator of $\beta_0$.

The disadvantage of likelihood-based estimation of $\beta_0$ is that a misspecified model for $Q_0$ immediately implies a biased representation of $\beta_0$ so that, for example, testing a null hypothesis $H_0 : \beta_0 = 0$ based on this likelihood-based estimator will practically fail to control the probability on a false rejection of the null hypothesis. We are concerned with constructing maximally robust estimators of $\beta_0$. In particular, we are interested in estimating $\beta_0$ based on data generated in a clinical trial such as a sequentially randomized trial, in which case the treatment mechanism $g_0$ is known. The knowledge about $g_0$ is not of any help for the likelihood-based approach so that, in particular, the likelihood-based estimator still fails to provide a valid test of the null hypothesis when $g_0$ is known. On the other hand, the inverse probability of treatment weighted (IPTW) and double robust (DR)-IPTW estimators of $\beta_0$, presented in the next section, are known to be consistent and asymptotically linear if $g_0$ is known. In this case, the latter estimators yield an asymptotically valid test of a null hypothesis $H_0 : \beta_0 = 0$, and yield root-n consistent estimators of our origin-specific statically optimal individualized treatment regimen $d(\theta_0)$ accompanied by valid confidence intervals.

3 Double robust inverse probability of treatment weighted estimating functions.

As presented in van der Laan and Robins (2003) (e.g. Chapter 6), given the model $m_3$, the class of all estimating functions can be represented in terms of a class of double robust IPTW estimating functions, derived by orthogonalizing a
class of IPTW estimating functions with respect to the treatment mechanism. This work builds on the early work of Robins on the estimation of marginal structural models, or models on the counterfactual outcome (Robins (1999, 2000); a more extensive review of the literature is provided in van der Laan and Robins (2003).

### 3.1 IPTW estimating functions.

We begin by providing the class of IPTW estimating functions, and thereby the corresponding class of IPTW estimators of $\beta_0$. In addition to contributing to the derivation of the double robust class of estimating functions for the CF-HA-MSM, the IPTW estimating functions are themselves of interest, as they can be solved using standard weighted least squares regression, as illustrated in the data example provided in Section 6. The intuition behind the IPTW estimating functions for CF-HA-MSM presented here is the same as that for IPTW estimating functions generally; namely, non-exchangeability between subjects following different treatment courses in the observed data is corrected for by assigning weights.

**Result 1** Consider the following class of IPTW-estimating functions for $\beta_0$ in the model for $O$ only assuming $\theta_0(t, a(t) \mid \bar{a}(t-1), \bar{s}(t)) = m_{\beta_0}(t, a(t) \mid \bar{a}(t-1), \bar{s}(t))$:

$$D_{h,IPTW}(O \mid \beta, g) \equiv \frac{1}{g(A \mid X)} \sum_{t=0}^{K(m)} h(t, \bar{A}, \bar{S}(t))(Y(t, m) - m_{\beta_0}(t, a(t) \mid \bar{a}(t-1), \bar{s}(t))).$$

If (4) holds, then

$$E_0 D_{h,IPTW}(O \mid \beta_0, g_0) = 0.$$

**Proof.** The conditional expectation of $D_{h,IPTW}(O \mid \beta_0, g_0)$, given $X$, is given by

$$\sum_{\bar{a}} \sum_t h(t, \bar{a}, \bar{S}_\bar{a}(t))(Y_\bar{a}(t, m) - m_{\beta_0}(t, a(t) \mid \bar{a}(t-1), \bar{s}(t))).$$

Now, move the expectation operator within the sums and condition on $\bar{S}_\bar{a}(t)$, giving us the term $E(Y_\bar{a}(t, m) \mid \bar{S}_\bar{a}(t)) - m_{\beta_0}(t, a(t) \mid \bar{a}(t-1), \bar{s}(t))$, which equals zero. This completes the proof. $\square$

As a particular choice for the IPTW-estimating function we propose $D_{h^*,IPTW}$ with

$$h^*(t, \bar{A}, \bar{S}(t)) \equiv g(\bar{A} \mid \bar{S}(t)) \frac{d}{d\beta_0} m_{\beta_0}(t, A(t) \mid \bar{A}(t-1), \bar{S}(t)),$$

where

$$g(\bar{A} \mid \bar{S}(t)) = \prod_{j=0}^{K} g(A(j) \mid \bar{A}(j-1), \bar{s}(\min(j, t))).$$

http://www.bepress.com/ijb/vol3/iss1/6
Such a choice of $h$ is based on the idea of stabilizing weights, as proposed by Robins, et. al. (see for example Robins (1999), Robins et al. (2000)). If the model $m_\beta$ is linear in $\beta$, then $h^*$ does not depend on $\beta$ and is thus known up to the stabilizing factor $g(\bar{A} | \bar{S}(t))$. The advantage of this choice is that the solution $\beta_{n,IPTW}$ of the estimating equation $\sum_{i=1}^n D_{h^*,IPTW}(O_i | \beta, g_n) = 0$ corresponds with a weighted least squares estimator:

$$\beta_{n,IPTW} = \arg \min_{\beta} \sum_{i=1}^n \sum_{t=0}^{K(m)} w_i(t) \left\{ Y_i(t, m) - m_\beta(t, A_i(t) | \bar{A}_i(t-1), \bar{S}_i(t)) \right\}^2$$

with weights given by

$$w_i(t) = \frac{g_n(\bar{A}_i | \bar{S}_i(t))}{g_n(\bar{A}_i | X_i)}.$$

This estimator can be calculated with standard regression software applied to a pooled sample in which each subject contributes $K(m)$ lines of data, using the weight option.

### 3.2 Double robust IPTW estimating functions for $\beta_0$.

In the next result we present the class of double robust IPTW estimating functions

**Result 2** Consider the following class of DR-IPTW-estimating functions for $\beta_0$ in the model for $O$ only assuming $\theta_0(t, a(t) | \bar{a}(t-1), s(t)) = m_{\beta_0}(t, a(t) | \bar{a}(t-1), s(t))$:

$$D_{h,DR}(O | g, Q, \beta_0) \equiv D_{h,IPTW}(O | g, \beta) - D_{h,SRA}(O | g, Q),$$

where

$$D_{h,SRA}(O | g, Q) \equiv \sum_{t=0}^{K(m)} E_{g,Q}(D_{h,IPTW}(O | g, \beta(Q)) | \bar{A}(t), \bar{L}(t))$$

$$- \sum_{t=0}^{K(m)} E_{g,Q}(D_{h,IPTW}(O | g, \beta(Q)) | \bar{A}(t-1), \bar{L}(t)),$$

We have that

$$E_{g_0,Q_0}(D_{h,DR}(O | g, Q, \beta_0) = 0,$$

if $g$ satisfies (4), and either $g = g_0$ or $Q = Q_0$.

Given estimators $g_n$ and $Q_n$, the corresponding estimator of $\beta_0$ based on the estimated likelihood (i.e. the $G$-computation estimator $\beta(Q_n)$), and a possibly
estimated index $h_n$, the double robust IPTW estimator $\beta_{n,DR}$ is defined as the solution in $\beta$ of the estimating equation

$$0 = \sum_{i=1}^{n} D_{h_n,DR}(O_i \mid g_n, Q_n, \beta).$$

If $\beta \rightarrow m_{\beta}$ is linear, then this estimating equation in $\beta$ is linear in $\beta$ so that the solution $\beta_{n,DR}$ exists in closed form.

### 3.3 Special case of counterfactuals indexed by restricted treatment history.

We note that, in the special case that $Y_\bar{A}(t, m) = Y_{\bar{A}(t^*(m))}(t, m)$, so that the counterfactuals of interest are only indexed by treatment up till time $t^*(m)$, then the IPTW and DR estimating equations can be altered so that $g(\bar{A}|X) = g(\bar{A}(t^*(m))|X)$ and $h(t, \bar{A}, \bar{S}(t)) = h(t, \bar{A}(t^*(m)), \bar{S}(t))$. Such a situation occurs, for example, if the outcome of interest is $Y(t + m)$, so that the counterfactuals are indexed only by treatment $\bar{A}(t + m - 1)$. In this case, the IPTW estimating function can be written as:

$$D_{h,IPTW}(O \mid \beta, g) = \sum_{t=0}^{K(m)} \frac{h(t, \bar{A}(t^*(m)), \bar{S}(t))}{g(\bar{A}(t^*(m)) \mid X)} (Y(t, m) - m_{\beta}(t, A(t, t^*(m)) \mid \bar{A}(t-1), \bar{S}(t))),$$

where $\bar{A}(t, t^*(m)) = (A(t), A(t + 1), ..., A(t^*(m)))$ denotes future treatment until the outcome is measured. We modify $h^*$ accordingly to

$$h^*(t, \bar{A}(t^*(m)), \bar{S}(t)) \equiv g(\bar{A}(t^*(m)) \mid \bar{S}(t)) \frac{d}{d\beta_0} m_{\beta_0}(t, A(t, t^*(m)) \mid \bar{A}(t-1), \bar{S}(t)),$$

where

$$g(\bar{A}(t^*(m)) \mid \bar{S}(t)) = \prod_{j=0}^{t^*(m)} g(A(j) \mid \bar{A}(j-1), \bar{S}(\min(j,t))).$$

The corresponding DR-IPTW estimating function is derived simply by subtracting off

$$D_{h,SRA}(O \mid g, Q) = \sum_{t=0}^{K(m)} E_{g,Q}(D_{h,IPTW}(O \mid g, \beta(Q)) \mid \bar{A}(t), \bar{L}(t))$$

$$- \sum_{t=0}^{K(m)} E_{g,Q}(D_{h,IPTW}(O \mid g, \beta(Q)) \mid \bar{A}(t-1), \bar{L}(t)).$$
4 Statistical inference.

Under appropriate conditions, and the assumption that either $g_n$ converges to $g_0$ or $Q_n$ converges to $Q_0$, it can be shown that these estimators of $\beta_0$ are asymptotically linear with specified influence curve (see van der Laan and Robins (2003) chapter 2). For example, if $g_n$ converges to $g_0$, and $Q_n$ converges to a possibly misspecified $Q_1$, then under regularity conditions, we have that $\beta_{n,DR}$ is a consistent and asymptotically linear estimator of $\beta_0$ with influence curve

$$IC(O) \equiv -c(\beta_0)^{-1}D_{h,DR}(O \mid g_0, Q_1, \beta_0) - \Pi(-c(\beta_0)^{-1}D_{h,DR} \mid T_G(P_0)),$$

where

$$c(\beta) = \frac{d}{d\beta}E_0D_{h,DR}(O \mid g_0, Q_1, \beta)$$

is the usual derivative matrix of the estimating equation, $T_G(P_0)$ is the tangent space of the nuisance parameter $g$ at $P_0$ under model $G$, and $\Pi(\cdot \mid T_G(P_0))$ is the projection operator onto this tangent space within the Hilbert space $L_2^0(P_0)$ endowed with covariance inner product $\langle f_1, f_2 \rangle = E_0f_1(O)f_2(O)$. As a consequence, conservative inference can be based upon the following influence curve, which is simple to calculate:

$$IC^*(O) \equiv -c(\beta_0)^{-1}D_{h,DR}(O \mid g_0, Q_1, \beta_0).$$

In particular, in the case that $G$ is correctly specified, a conservative asymptotic 0.95 confidence interval for $\beta_0(j)$ is given by

$$\beta_{n,DR}(j) \pm 1.96\sigma_n/\sqrt{n},$$

where

$$\sigma_n^2 \equiv \frac{1}{n} \sum_{i=1}^{n} \left( IC^*_n(O_i) - \frac{1}{n} \sum_{i=1}^{n} IC^*_n(O_i) \right)^2,$$

and $IC^*_n$ is an estimator of the function $IC^*$ obtained by substituting the estimators $g_n, Q_n$, and estimating the derivative matrix $c(\beta_0)$ with its empirical counterpart.

Since influence curve inference is heavily based on the first-order behavior of the estimator, in the case that $g_n$ and $Q_n$ are highly data-adaptive estimators we suggest the bootstrap method as a more honest method for establishing the true variability of $\beta_{n,DR}$ and obtaining corresponding confidence intervals.

Regarding inference for the individualized treatment rule $d(\theta_0) = d(\beta_0)$, we propose to use an estimate of the sampling distribution of $\beta_{n,DR}$. For example, one could use as estimate of this sampling distribution the distribution $\beta_{n,DR}^\# \sim N(\beta_{n,DR}, \sigma_n^2/n)$ or the bootstrap distribution of $\beta_{n,DR}$ defined by the
distribution of the double robust IPTW estimator when applied to samples of \( n \) i.i.d. observations from the empirical distribution. In this manner, one can obtain the sampling distribution of \( d_i(\hat{\beta}_{nDR}^{\#})(\bar{S}(t)) \) for treatment assignment at time \( t \) for any given history \( \bar{S}(t) \). That is, the estimate \( d(\hat{\beta}_{nDR}^{\#}) \) of the statically optimal individualized treatment rule will be accompanied with a measure of uncertainty when applied at any time \( t \) and history \( \bar{S}(t) \).

5 Comparison with statically optimal treatment rules.

In the preceding sections, we have illustrated how an origin-specific statically optimal treatment rule can be estimated based on the counterfactual history-adjusted mean outcome. The results of Petersen et al. (2007b) demonstrate that this treatment rule is also statically optimal in a more general sense when the following equality holds:

\[
E(Y_{\bar{A}(t-1)|\bar{A}(t)}|\bar{A}(t-1) = \bar{a}(t-1), \bar{S}(t) = \bar{s}(t)) = E(Y_{\bar{a}(t-1)|\bar{S}(\bar{a}(t)) = \bar{s}(t)}) \tag{6}
\]

Specifically, when the counterfactual history-adjusted mean outcome equals the observed history-adjusted mean outcome (as estimated using the history-adjusted marginal structural models of van der Laan et al. (2005)) then the static optimality of the individualized treatment given by Definition 2 is no longer origin-specific. That is, if equality (6) holds, then the resulting rule chooses the future static treatment plan that optimizes expected outcome regardless of how past treatment has been assigned. The rule thus retains its optimality properties not only if applied to a population that has been following the rule of interest, but also if applied to a population that has been following some other treatment mechanism.

There are several practical implications of this finding. If \( \bar{S}(t) \) is chosen so that equality (6) holds, the individualized treatment rules estimated using the counterfactual history-adjusted mean will gain an additional property; they will be generally statically optimal rather than origin-specific statically optimal, and thus will be appropriate for application in contexts where the past treatment mechanism is unknown. This suggests that, if general static optimality is desirable, the researcher may wish to choose the covariates to be included in the rule accordingly.

Petersen et al. (2007b) provide criteria for \( \bar{S} \) sufficient to ensure (general) static optimality. Specifically, equality (6) will hold if the covariates on which the rule depends are sufficient to control for confounding of past treatment history on future outcome. More formally,

\[
\text{If } P(\bar{A}(t-1) = \bar{a}(t-1)|Y_{\bar{a}} = y, \bar{S}_{\bar{a}}(t) = \bar{s}(t)) = \]

http://www.bepress.com/ijb/vol3/iss1/6
\[
s(t) = P(A(t-1) = \bar{a}(t-1) | \bar{S}(t) = \bar{s}(t))
\]
then
\[
E(Y_{\bar{A}(t-1) \bar{a}(t-1)} | \bar{A}(t-1) = \bar{a}(t-1), \bar{S}(t) = \bar{s}(t)) = E(Y_{\bar{a}(t-1) \bar{a}(t)} | \bar{S}(t) = \bar{s}(t)).
\]

Petersen et al. (2007b) point out that if past treatment assignment is only a function of the covariates of interest \(S(t)\), or if the covariates of interest \(\bar{S}(t)\) \(d\)-separate \(A(t-1)\) from \(Y_{\bar{a}(t,m)}\), then this identity will hold, and estimation of either the observed history-adjusted parameter or the counterfactual history-adjusted parameter will estimate the (general) statically optimal treatment rule.

Inclusion of sufficient covariates in the rule to ensure that past treatment assignment is only a function of \(S(t)\) may be undesirable or unpractical. In the case where this latter condition is not met, the question may still arise as to whether the static optimality of a rule based on the counterfactual history-adjusted mean is origin-specific. The \(d\)-separation criteria provides one means to evaluate the claim of general vs. origin-specific static optimality; however, this approach relies on background knowledge sufficient to inform the underlying causal graph. Alternatively, the observed history-adjusted parameter, as described in van der Laan et al. (2005), can be estimated, and the null hypothesis that the counterfactual history-adjusted parameter is equal to the observed history-adjusted parameter (i.e. that equality 6 holds) can be tested, using, for example, a chi-square statistic.

6 Data example: When to switch antiretroviral therapy?

This section describes a data example focused on making treatment decisions for individuals infected with resistant HIV. While antiretroviral regimens are generally able to suppress HIV replication, viral drug resistance frequently emerges. Resistance allows HIV replication to resume, resulting in an increase in the amount of virus detectable in a patient’s blood (plasma HIV RNA level or viral load), and potentially accelerating immunologic decline (reflected in a falling CD4 T cell count) and disease progression. Ideally, a patient infected with resistant virus will be switched to a new regimen to which the virus remains susceptible (DHHS (2004)). However, a limited number of antiretroviral regimens are available, and alternative regimens may be more toxic or difficult to adhere to than a patient’s current regimen. Given evidence that some antiretroviral regimens continue to confer immunologic benefits in the presence of viral resistance, it is unclear how long the clinician should wait before switching a patient who has lost viral suppression to a new antiretroviral regimen (Deeks (2003)). Switching too early risks prematurely depleting future treat-
ment options, while switching too late risks accelerating disease progression, as well as allowing the virus to evolve new resistance mutations. We applied the method described in this paper to estimate an origin-specific statically optimal treatment rule for deciding when to switch therapy among HIV-infected individuals who have lost virologic suppression due to the emergence of resistant virus.

6.1 Data.

The data are drawn from the Study of the Consequences of the Protease Inhibitor Era (SCOPE), an observational clinical cohort of HIV-infected individuals in San Francisco, California. Subjects were followed longitudinally over time, and data were collected on all antiretroviral drug use, AIDS-defining illnesses, use of recreational drugs, adherence to prescribed antiretroviral therapies, homelessness, presence of hepatitis C virus antibody, CD4 and CD8 T cell counts, and plasma HIV RNA levels. In addition, baseline data were collected on demographics (age, sex, income, race), sexual orientation, and treatment history. We denote these covariates $L$.

We identified all episodes of virologic failure among patients followed in SCOPE between 2000 and 2004. Virologic failure ($t=0$) was defined as at least 2 detectable and no undetectable plasma HIV RNA levels in either 1) the first 6 months after starting a new regimen; or 2) over a 4 month period on a stable regimen. The outcome of interest for a given time $t$ was CD4 T cell count $m = 8$ months in the future ($Y(t+8) \subset L(t+8)$). The treatment of interest was time until treatment modification (switch), where treatment modification was defined as change or interruption of at least 1 drug in the failing regimen. At each time point during follow up, treatment was defined using a binary variable ($A$) indicating whether a subject remained on his original non-suppressive therapy ($A = 1$ until a subject switched, after which $A = 0$).

The analysis focused on the 8 months following loss of viral suppression ($t = 0, ..., 8$); $K + 1 = 16$ and $K(m) = 8$. Because a subject could only switch therapy once, counterfactual outcomes of interest were only defined for time points up till the point that a subject switched treatment. We thus note that time points more than 8 months following a subject’s switch time did not contribute.

In the absence of censoring, the observed data thus would have consisted of $n$ i.i.d. copies of

$$O^* = (L(0), A(0), L(1), A(1), ...L(K), A(K), L(K + 1)) = (\bar{A}(K), \bar{L}_A(K + 1))$$

We note that this observed data, in the absence of censoring, can also be considered a time-dependent process:

$$O^*(t) = (\bar{A}(t-1), \bar{L}_A(t)),$$
where $t = 0, \ldots, K + 1$.

Subjects were further subject to two distinct censoring processes. The full data on a subject could be censored 1) when follow-up ended in 2004, or 2) as a result of death or loss to follow-up (here, we consider death a censoring process rather than an outcome of interest). We denote the time at which censoring occurred due to the end of follow-up as $C_1$, and the time at which censoring occurred due to death or loss to follow-up as $C_2$. $C = \min(C_1, C_2)$ denotes a subject’s censoring time, and we define $\bar{T} = \min(K + 1, C)$. We further define a censoring process over time:

$$\bar{C}(t) = (\bar{C}_1(t), \bar{C}_2(t)) = (I(C_1 \leq t), I(C_2 \leq t))$$

The observed data thus consisted of $n$ i.i.d. copies of

$$O = (O^*(\bar{T}), \bar{C}(\bar{T}), \bar{T})$$

In all, 133 subjects (167 episodes of failure) were evaluated. Of these, 66 episodes were censored due to the end of follow-up in 2004, and 18 were censored due to death or loss to follow-up (3 deaths and 15 losses to follow up). In total, 116 episodes (100 subjects) had at least one outcome available (corresponding to $t = 0$). Of these subjects, median time to switch was 6 months (IQR=4,11). The study population was primarily male (86%), and primarily men who have sex with men (49%). Subjects were heavily treatment experienced; 49% were treated with antiretroviral drugs prior to the availability of protease inhibitors in 1996. Petersen et al. (2007a) describe the sample in greater detail.

### 6.2 Parameter of interest.

We aimed to identify the origin-specific statically optimal rule for deciding when to modify treatment, given a specific set of covariates $\bar{S}(t)$. In other words, we estimated for each time point the future switch time expected to maximize CD4 T cell count 8 months later, given covariate values, among individuals who had not yet modified treatment. Following, at each time point, the first action (switch or not) of this optimal treatment plan provided an individualized treatment rule. The static optimality of the rule was origin-specific because it identified, for each time point, the optimal future switch time given that subjects had followed the statically optimal rule itself up till that time point.

Specifically, we considered treatment rules based on current CD4 T cell count and an indicator of viral re-suppression prior to switching regimens. The latter covariate was included because our goal was to identify rules for switching among individuals who were infected with resistant HIV. Individuals
who achieved viral re-suppression without switching regimens almost certainly did not initially lose suppression due to the presence of resistant virus. Thus, 
\[ S(t) = (CD4(t), Sup(t)) \] where \( CD4(t) \) denoted CD4 T-cell count at time \( t \), and \( Sup(t) \) denoted an indicator that re-suppression of the virus had occurred by time \( t \).

As demonstrated in Lemma 1, the origin-specific statically optimal treatment rule for deciding when to switch (among individuals who have not already switched, i.e. \( \bar{a}(t - 1) = 1 \)) is identified by the parameter
\[
\theta(t, a(t)|\bar{a}(t - 1) = 1, Sup(t) = 0, CD4(t)) = E(Y_{\bar{a}(t-1)=1,a(t)}(t+8)|Sup(t) = 0, CD4(t)).
\]

We further note that, as the outcome is measured at time \( t + 8 \), the counterfactuals of interest are in fact indexed only by treatment up till time \( t^* = t + 8 - 1 \) \( (Y_{\bar{a}(t-1)=1,a(t)}(t+8) = Y_{\bar{a}(t-1)=1,a(t,t^*)}) \), where we remind the reader that \( a(t, t^*) = (a(t), a(t + 1), \ldots, a(t^*)) \).

### 6.3 Model for counterfactual history-adjusted mean.

We assumed the following model on the parameter \( \theta(t, a(t, t^*)|\bar{a}(t - 1), \bar{s}(t)) = m_{\beta}(t, a(t, t^*)|\bar{a}(t - 1))\bar{s}(t)) \), where
\[
m_{\beta}(t, a(t, t^*)|\bar{a}(t - 1))\bar{s}(t)) = \beta_0 + \beta_1 \sum_{j=t}^{t^*} a(j) + \beta_2 CD4(t) + \beta_3 t + \beta_4 \sum_{j=t}^{t^*} a(j) \times CD4(t) + \beta_5 \sum_{j=t}^{t^*} a(j) \times CD4(t) \times t + \beta_6 a(t, t^*) \sum_{j=t}^{t^*} a(j) \times CD4(t) \times t,
\]
where \( \sum_{j=t}^{t^*} a(j) \) is the residual amount of time until either treatment is modified or the outcome is measured, under treatment regimen \( a(t, t^*) \).

### 6.4 Model for observed data.

**Treatment mechanism.** As defined in Subsection 2.1, we assumed sequential randomization; in other words, we assumed that the decision whether to switch treatment or not at each time point only depended on covariates measured prior to that time point. In addition, as defined in Subsection 2.6, we assumed experimental treatment assignment; namely that an individual who had not already switched had some positive probability of both switching treatment and not switching, regardless of her observed past.

**Censoring mechanism.** We assumed that the probability of being censored
at every time point, given that censoring had not already occurred, only depended on the observed past (censoring at random):

\[ g(\bar{C}(K) = 0|O^*) \equiv \prod_{t=0}^{K} Pr(C > t|\bar{C}(t-1) = 0, O^*) \]

\[ = \prod_{t=0}^{K} Pr(C_1 > t|\bar{C}(t-1) = 0, \bar{A}(t), \bar{L}(t)) \]

\[ \prod_{t=0}^{K} Pr(C_2 > t|C_1 > t, \bar{C}(t-1) = 0, \bar{A}(t), \bar{L}(t)) \]

We also made two additional identifiability assumptions (counterpart to the experimental treatment assignment assumption). For each type of censoring and every time point, we assumed that, given that censoring had not already occurred, an individual had some positive probability of not being censored regardless of his observed past:

\[ Pr(C_1 > t|\bar{C}(t-1) = 0, \bar{A}(t), \bar{L}(t)) > 0, t = 0, ..., K \]

and

\[ Pr(C_2 > t|C_1 > t, \bar{C}(t-1) = 0, \bar{A}(t), \bar{L}(t)) > 0, t = 0, ..., K \]

We note that censoring due to the end of follow-up in 2004 \((C_1)\) is not necessarily non-informative, as calendar time at baseline \((t = 0)\) in the current data analysis is itself a random variable that deterministically predicts censoring due to end of follow-up in 2004, and could also potentially be related to outcome (due to differences in the characteristics of subjects that lose virologic suppression at different calendar times).

### 6.5 IPTW estimation.

In the absence of censoring, the IPTW estimating function would be

\[ D_{h, IPTW}(O|\beta, g) = \sum_{t=0}^{K(m)} h(t, A(t^*), S(t)) \frac{g(A(t^*)|X)}{g(\bar{A}(t^*)|X)} \]

\[ I(\bar{A}(t-1) = 1)I(Sup(t) = 0)\{Y(t+8) - m_\theta(t, A(t, t^*)|\bar{A}(t-1), S(t))\}, \]

where \(K(m) = 8\). We chose \( h \) as

\[ h^*(t, \bar{A}(t^*)S(t)) \equiv g(\bar{A}(t^*)|S(t)) \frac{d}{d\beta} m_\theta(t, A(t, t^*)|\bar{A}(t-1), S(t)) \]

where \( g(\bar{A}(t^*)|S(t)) = \prod_{j=0}^{t} g(A(j)|\bar{A}(j-1), S(\min(j,t))) \).
However, in the presence of censoring, we use the following estimating function, which incorporates an additional inverse probability of censoring component:

$$D_{h, IPTW}(O|\beta, g) \equiv \sum_{t=0}^{K(m)} \frac{I(C > t^*)g(\bar{C}(t^*) = 0, \bar{A}(t^*), \bar{S}(t)) h(t, \bar{A}(t^*), \bar{S}(t))}{g(C(t^*) = 0|O^*) g(A(t^*)|X)} I(\bar{A}(t-1) = 1) I(\text{Sup}(t) = 0) \{Y(t+8) - m_\beta(t, \bar{A}(t), t^*)|\bar{A}(t-1)\bar{S}(t)\}.$$  

The estimator was implemented using weighted least squares, as described in Section (3). Specifically, each subject contributed one weighted line of data for each time point \(t \leq K(m)\) for which censoring did not occur before the outcome was measured \((t + 8 \leq C)\), and for which the subject had not already switched treatments \((I(A(t-1) = 1) = 0)\), or achieved re-suppression of the virus \((I(\text{Sup}(t-1) = 0) = 0)\). In this pooled dataset, we regressed the observed CD4 T cell count 8 months in the future \((Y(t+8))\) on future time until switching treatment \((g^*(t, t^*))\), elapsed time \(t\) since failure, and current CD4 T cell count \((CD4(t))\), according to the model \(m_\beta\).

For a given time point \(t\), the weight was estimated as the product of a treatment component,

$$\frac{g(\bar{A}(t^*)|\bar{S}(t))}{g(A(t^*)|X)},$$

and a censoring component,

$$\frac{g(\bar{C}(t^*) = 0, \bar{A}(t^*), \bar{S}(t))}{g(C(t^*) = 0|O^*)}.$$  

Note that there is flexibility in choosing a numerator for these weights. Given fits of the treatment mechanism and censoring mechanism, one generally selects a numerator with the purpose of making the weights minimally variable (i.e. of making the weights as close to 1 as possible). Several approaches are available to do this; one general strategy involves simply using the treatment/censoring mechanism selected, but setting all terms not included in \(\bar{S}(t), \bar{A}(t^*)\) equal to zero.

By factorizing the censoring component, it can be further rewritten as a product of a weight for censoring mechanism 1 (end of follow-up in 2004),

$$\prod_{j=0}^{t^*} Pr(C_1 > j|C(j-1) = 0, \bar{A}(j), \bar{S}(j)) \prod_{j=t^*+1}^{t^*} Pr(C_1 > j|C(j-1) = 0, \bar{A}(j), L(j)) \prod_{j=0}^{t^*} Pr(C_1 > j|C(j-1) = 0, \bar{A}(j), \bar{S}(j))$$

and a weight for censoring mechanism 2 (death or loss to follow-up),

$$\prod_{j=0}^{t^*} Pr(C_2 > j|C_1 > j, C(j-1) = 0, \bar{A}(j), \bar{S}(j)) \prod_{j=t^*+1}^{t^*} Pr(C_2 > j|C_1 > j, C(j-1) = 0, \bar{A}(j), \bar{S}(j)) \prod_{j=0}^{t^*} Pr(C_2 > j|C_1 > j, C(j-1) = 0, \bar{A}(j), L(j))$$
Similarly, the treatment component of the weights can be written as:

\[
\prod_{j=0}^t g(A(j)|\tilde{S}(j), \tilde{A}(j - 1) = 1) \prod_{j=t+1}^{t^*} g(A(j)|\tilde{S}(t), \tilde{A}(j - 1) = 1) \\
\prod_{j=0}^{t^*} g(A(j)|\tilde{L}(j), \tilde{A}(j - 1) = 1)
\]

Implementation of the IPTW estimator, then, relied on estimation of the following nuisance parameter models:

1. **Treatment mechanism:**

\[
\prod_{j=0}^{t^*} g(A(j)|\tilde{L}(j), \tilde{A}(j - 1) = 1)
\]

We used the Deletion/Substitution/Addition algorithm (Sinisi and van der Laan (2004)) and 5-fold cross validation to fit a pooled logistic regression model of the probability of switching treatment given the observed past. Note that the model was fit only among those who had not already been censored or switched, as these were the only subjects at risk of switching. Estimation of the treatment mechanism employed inverse probability of censoring weights for each time point \( j \) (\( \frac{I(C > j)}{g(C(j) = 0|O^*)} \)). In modelling the treatment mechanism, we assumed that treatment assignment at time \( j \) was independent of covariate history at time \( j - 1 \) given covariates at time \( j \). In other words,

\[
g(A(j)|\tilde{L}(j), \tilde{A}(j - 1) = 1) = g(A(j)|L(j), \tilde{A}(j - 1) = 1).
\]

2. **Numerator for treatment weight:** In calculating the numerator of the treatment weights, for \( j \leq t \) we made the similar assumption that \( g(A(j)|\tilde{S}(j), \tilde{A}(j - 1) = 1) = g(A(j)|S(j), \tilde{A}(j - 1) = 1) \). For time points \( j > t \), a different model had to be adopted, as the latest available covariates were measured at time \( t \). To avoid the need to fit a separate model for each time point after \( t \) (e.g., \( g(A(j)|S(j - 1), \tilde{A}(j - 1) = 1) \), \( g(A(j)|S(j - 2), \tilde{A}(j - 1)) \), etc.), for \( j > t \) we used the model \( g(A(j)|\tilde{S}(t), \tilde{A}(j - 1) = 1) = g(A(j)|S(0), \tilde{A}(j - 1) = 1) \). Thus, the numerator of the treatment weight consisted of

\[
\prod_{j=0}^t g(A(j)|S(j), \tilde{A}(j - 1) = 1) \prod_{j=t+1}^{t^*} g(A(j)|S(0), \tilde{A}(j - 1) = 1),
\]

Estimates of \( g(A(j)|S(j), \tilde{A}(j - 1) = 1) \) and \( g(A(j)|S(0), \tilde{A}(j - 1) = 1) \) were fit using logistic regression of the probability of switching on most recent or baseline CD4 T cell count, respectively, and suppression history.
3. Censoring mechanisms:

\[
\prod_{j=0}^{t^*} Pr(C_1 > j | \bar{C}(j-1) = 0, \bar{A}(j), \bar{L}(j))
\]

\[
\prod_{j=0}^{t^*} Pr(C_2 > j | C_1 > j, \bar{C}(j-1) = 0, \bar{A}(j), \bar{L}(j))
\]

As with the treatment mechanism, we used the D/S/A algorithm to fit, for each censoring mechanism, a pooled logistic regression model of the probability of being censored given that censoring had not already occurred and the observed past. As in modelling the treatment mechanism, we assumed that censoring probability at time \( j \) was independent of covariate history at time \( j - 1 \) given covariates at time \( j \). In other words,

\[
Pr(C_1 > j | \bar{C}(j-1) = 0, \bar{A}(j), \bar{L}(j)) = Pr(C_1 > j | \bar{C}(j-1) = 0, \bar{A}(j), L(j)),
\]

and similarly for \( C_2 \).

4. Numerators for censoring weight: In estimating the numerator for the censoring weight, we made equivalent assumptions as when estimating the numerator for the treatment weight. The numerators of the censoring weights consisted of:

\[
\prod_{j=0}^{t} Pr(C_1 > j | \bar{C}(j-1) = 0, \bar{A}(j), S(j)) \times
\prod_{j=t+1}^{t^*} Pr(C_1 > j | \bar{C}(j-1) = 0, \bar{A}(j), S(0))
\]

\[
\prod_{j=0}^{t} Pr(C_2 > j | C_1 > j, \bar{C}(j-1) = 0, \bar{A}(j), S(j)) \times
\prod_{j=t+1}^{t^*} Pr(C_2 > j | C_1 > j, \bar{C}(j-1) = 0, \bar{A}(j), S(0))
\]

The resulting estimates \( \beta_n \) of \( \beta \) provided a origin-specific statistically optimal treatment rule according to Definition (2). Standard error estimates and confidence intervals for the parameter \( \beta \), and variability in the resulting decision rule, were calculated by applying the entire estimation algorithm to 100 non-parametric bootstrap samples, respecting the subject (rather than failure episode) as the independent experimental unit.
6.6 Results.

The D/S/A algorithm and cross-validation selected a treatment mechanism with 8 main terms. The corresponding odds ratios are reported in Table (1). The same algorithm applied to the censoring mechanisms selected an intercept-only model for each of the censoring mechanisms; thus, the censoring component of the weights was estimated as 1. Both prior work (Petersen et al. (2007b)) and background knowledge suggest that CD4 T cell count may be the most important potential source of bias due to informative censoring. To address this concern, we performed a sensitivity analysis, in which we fit a model of censoring due to loss to follow-up/death ($C_2$) based on most recent CD4 T cell count and used this model in the estimation of the censoring component of the weights. Changes in the causal coefficients estimated using this censoring model were minimal (relative change of 3% or less), supporting the presence of minimal bias due to informative censoring.

Table 1: Odds ratios for switching treatment based on data-adaptive fit of treatment mechanism (Petersen et al. (2007a))

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current diagnosis with an opportunistic disease</td>
<td>1.21</td>
</tr>
<tr>
<td>Number of protease inhibitor drugs experienced</td>
<td>1.11</td>
</tr>
<tr>
<td>Most recent HIV RNA level undetectable</td>
<td>0.44</td>
</tr>
<tr>
<td>Percent average adherence (per 10%)</td>
<td>0.92</td>
</tr>
<tr>
<td>Most recent CD4 T cell count (per 100 CD4 T cells)</td>
<td>0.92</td>
</tr>
<tr>
<td>Nadir CD4 T cell count (per 100 CD4 T cells)</td>
<td>1.06</td>
</tr>
<tr>
<td>Most recent HIV RNA level more than one month prior</td>
<td>0.90</td>
</tr>
<tr>
<td>Age (per 5 years)</td>
<td>0.80</td>
</tr>
</tbody>
</table>

IPTW estimation relying on these fits yielded the following estimate of the counterfactual history-adjusted parameter of interest:

$$m_{\beta}(t, a(t, t^*)|\bar{a}(t - 1) = 1, Sup(t) = 0, CD4(t)) =$$

$$92.8 - 9.4 \times \sum_{j=t}^{t^*} a(j) + 0.48 \times CD4(t) - 16.12 \times t + 0.05 \times \sum_{j=t}^{t^*} a(j) \times CD4(t)$$

$$+ 1.46 \times \sum_{j=t}^{t^*} a(j) \times t + 0.07 \times CD4(t) \times t - 0.009 \times \sum_{j=t}^{t^*} a(j) \times CD4(t) \times t,$$

This model yields the following origin-specific statically optimal treatment rule:

$$d_t = I\{ -9.4 \times \sum_{j=t}^{t^*} a(j) + 0.05 \times \sum_{j=t}^{t^*} a(j) \times CD4(t),$$

Published by The Berkeley Electronic Press, 2007
\[ +1.46 \times \sum_{j=t}^{t^*} a(j) \times t - 0.009 \times \sum_{j=t}^{t^*} a(j) \times CD4(t) \times t < 0, \]

where, \( d_t \) is the treatment decision at time \( t \) (if \( d_t = 1 \) then switch, if \( d_t = 0 \) then wait). The coefficients which contribute to this rule, together with their 95% confidence intervals, are provided in Table 2.

### Table 2: Coefficients contributing to origin-specific statically optimal rule for when to switch therapy, based on model \( m_\beta(t, g(t, t^*)|a(t - 1) = 1, Sup(t) = 0, CD4(t)) \)

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sum_{j=t}^{t^*} a(j) )</td>
<td>-9.4</td>
<td>-17.8, -0.9</td>
</tr>
<tr>
<td>( \sum_{j=t}^{t^*} a(j) \times CD4(t) )</td>
<td>0.05</td>
<td>0.02, 0.08</td>
</tr>
<tr>
<td>( \sum_{j=t}^{t^*} a(j) \times t )</td>
<td>1.46</td>
<td>-0.5, 3.4</td>
</tr>
<tr>
<td>( \sum_{j=t}^{t^*} a(j) \times t \times CD4(t) )</td>
<td>-0.009</td>
<td>-0.02, -0.002</td>
</tr>
</tbody>
</table>

In addition to providing confidence intervals for estimates of the coefficients that contribute to the individualized treatment rule, bootstrap sampling provides a means to judge the variability of the treatment decision provided by the rule. Figure 1 shows the proportion of bootstrap samples in which the origin-specific statically optimal treatment rule indicated a switch, plotted for each month following loss of viral suppression, and for four different CD4 T cell counts.

The results of these analyses can be summarized as follows:

- Immediately following loss of suppression, individuals with high CD4 T-cell counts can wait to switch, while individuals with low CD4 T cell counts should switch immediately.

- At later time points, an individual’s current CD4 T cell count is less important to the decision whether to wait to switch. Thus at later time points, the decision whether to switch or not is less clear, and may better be made based on additional considerations.

It is interesting to compare these results, based on a CF-HA-MSM, with the results reported in Petersen et al. (2007a). Petersen et al. (2007a) reported estimates of the observed HA-MSM parameter, using the same SCOPE dataset, same \( \bar{S}(t) \), and same model \( m_\beta \), but estimating the observed history-adjusted parameter:

\[ E[Y_{\bar{A}(t-1)|g(t,t^*)}|\bar{S}(t), \bar{A}(t - 1)]. \]

The coefficients contributing to the resulting individualized treatment rule are reported in Table 3. The similarity between the estimate of the counterfactual history-adjusted mean (Table 2) and the estimate of the observed
history-adjusted mean (Table 3) supports the claim of Petersen et al. (2007b) that in this dataset, the choice $\bar{S}(t) = (CD4(t), Sup(t))$ is sufficient to control confounding of the effect of past treatment history (up till time $t - 1$) on the outcome. Such a finding suggests that the static optimality of the treatment rule presented here is not origin-specific; the rule should remain statically optimal if applied to individuals remaining on their original non-suppressive therapy at a given time point, regardless of how the decision whether to switch therapy up till that time point has been made.

Table 3: Estimated effect of each additional month waiting to switch on CD4 T cell count 8 months later, based on observed history-adjusted parameter (Petersen et al. (2007a))

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_{j=t} a(j)$</td>
<td>-9.2</td>
<td>-17.6, -7.6</td>
</tr>
<tr>
<td>$\sum_{j=t} a(j) \times CD4(t)$</td>
<td>0.05</td>
<td>0.02, 0.08</td>
</tr>
<tr>
<td>$\sum_{j=t} a(j) \times t$</td>
<td>1.5</td>
<td>-0.4, 3.4</td>
</tr>
<tr>
<td>$\sum_{j=t} a(j) \times t \times CD4(t)$</td>
<td>-0.009</td>
<td>-0.02, -0.004</td>
</tr>
</tbody>
</table>

This data example illustrates how CF-HA-MSM can be implemented using standard weighted regression, and how the resulting estimates provide an origin-specific statically optimal treatment rule. The scientific question addressed in the example is of real clinical interest. However, the results presented are intended as an illustration of the methodology rather than as a guide to clinical practice, due to several practical decisions made as a result of the small sample size and in order to simplify the data analysis. Specifically, treatment modification was defined in these analyses as any change in a subject’s failing regimen; thus modification did not require that a subject switch to a new regimen aimed at viral re-suppression, but could also include treatment interruption. This approach was taken to avoid the need to define an additional component to the treatment. In addition, the outcome optimized was a relatively short-term biomarker which may not reflect the implications of waiting to switch regimens for long-term mortality and disease progression. Finally, the counterfactual history-adjusted mean was modelled as a linear function of time until switching therapy; such a linear model may be ill-suited to capture any tradeoff between early and delayed switch times.

7 Discussion.

This paper has presented a new parameter of the full data-generating distribution, together with corresponding estimating equations, and demonstrated
Figure 1: Variability in statically optimal decision whether to switch therapy, depending on current CD4 T cell count and elapsed time since failure (among subjects who have not already switched therapy and have not achieved viral re-suppression).

that this parameter directly identifies an origin-specific statically optimal individualized treatment rule. The proposed individualized treatment rule is relatively easy to estimate with standard software. We have further shown that, applied to a data example, the method can provide both practical and interpretable results.

If these methods are applied to data generated in a sequentially randomized trial, in which the treatment mechanism is known, then the DR-IPTW estimator is known to be asymptotically consistent and asymptotically linear under the assumption that the model for the counterfactual history-adjusted mean is correct. In particular, since our model \( \{m_\beta : \beta\} \) for the counterfactual history-adjusted mean always contains the null hypothesis \( H_0 : \beta = 0 \), it follows that the IPTW or DR-IPTW estimator provide a valid test of the null.
As clarified in the discussion of van der Laan et al. (2005), our model for the counterfactual history-adjusted mean of the outcome $Y_\bar{a}(t, m)$ can be replaced by a model for a counterfactual history-adjusted parameter of the conditional distribution $P_{Y_\bar{a}(t,m)|S_\bar{a}(t)}$, such as the conditional median or conditional survival function of $Y_\bar{a}(t, m)$. In this manner, our models yield estimators of origin-specific statically optimal individualized treatment rules which are optimal with respect to any user-supplied parameters of the distribution of the future outcome. For example, in the case that the outcome process of interest is an indicator process jumping from 0 to 1 at a survival time (e.g., time till recurrence of cancer) our methods can be used to estimate individualized treatment rules which at each point in time, conditional on a user-supplied subset of the observed history, select the treatment action (statically) optimizing the survival probability at (e.g.) 5 additional years.

Further, as illustrated in the data example, the general estimating function methodology (van der Laan and Robins (2003)) for censored data can be used to map the estimating functions based on observing $(\bar{A}, L_{\bar{A}})$ presented in this article into estimating functions for the censored longitudinal causal inference data structure $O = (C, \bar{A}(C), L_{\bar{A}}(C))$ for a right-censoring variable $C$ (Chapter 3, van der Laan and Robins (2003)).

As noted in the introduction, the statically optimal treatment rules estimated based on models of the counterfactual history-adjusted mean will generally be inferior to the optimal dynamic treatment regime. However, they appear to provide an interesting alternative, in that they are less ambitious and should thus be able to be estimated with greater precision, as well as being straightforward to implement using standard software. In addition, the idea of using a model of the history-adjusted mean outcome to generate interesting rules is very general. The origin-specific statically optimal rules addressed here choose, at each time point, the best rule from among a user-supplied set of static treatment rules. Alternatively, one might estimate the the best rule at each time point from among a user-supplied set of dynamic rules, where the dynamic rule itself can be updated at subsequent time points in response to changes in patient covariates. In settings where the time scale is not itself meaningful, such an approach should be able to provide an additional increase in precision by pooling across time points.
8 Appendix.

8.1 An alternative derivation of DR-IPTW estimating functions.

It is interesting to note the link between the DR-IPTW estimating functions presented in the previous section and the estimating functions for the observed history-adjusted mean presented in van der Laan et al. (2005). Consider a treatment mechanism \( g^*(A \mid X) \) so that

\[
E_{P_{X_0}}(Y_{\tilde{a}(t)} | \tilde{S}(t)) = E_{P_{X_0},g^*}(Y_{\tilde{a}(t)}, t, m) \mid \tilde{S}(t) = \tilde{s}(t), \tilde{A}(t - 1) = \tilde{a}(t - 1)).
\]

For example, any treatment mechanism \( g^*(A \mid X) = \prod_{t=0}^{K} g(A(t) \mid \tilde{A}(t - 1), \tilde{S}(t)) \) satisfies this condition: that is, if treatment assignment is only based on the \( S(t) \) process, then the counterfactual history-adjusted mean equals the observed history-adjusted mean. Thus, our model (3) can also be viewed as the following model:

\[
E_{P_{X_0},g^*}(Y_{\tilde{a}(t)} | \tilde{S}(t)) = \tilde{s}(t), \tilde{A}(t - 1) = \tilde{a}(t - 1)) \quad (7)
\]

where \( \beta \) is arbitrary functions \( h, g, Q \). Thus, we have

\[
E_{P_{X_0},g^*}D_{h,g^*,Q}(O \mid \beta) = 0 \text{ for all } h, g^*, Q.
\]

The choice \( Q = 0 \) corresponds with the class of IPTW-estimating functions for the HA-MSM (7) based on sampling from \( P_{X_0,g^*} \). Let’s denote the latter DR-IPTW estimating functions with \( D_{h,g^*,Q}(O \mid \beta) \) indexed by arbitrary functions \( h, g^*, Q \). The typical choice we recommend is

\[
D_{h,g^*}(O \mid \beta) = \sum_{t=0}^{K(m)} h(t, \tilde{A}, \tilde{S}(t)) \{ Y(t, m) - m_\beta(t, A(t) \mid \tilde{A}(t - 1), \tilde{S}(t)) \},
\]

where

\[
g^*(A(t) | \tilde{A}(t - 1), X) = \prod_{j=t}^{K} g^*(A(j) | \tilde{A}(j - 1), \tilde{S}(j)).
\]

The typical choice we recommend is

\[
h^*(t, \tilde{A}, \tilde{S}(t)) \equiv \frac{d}{d\beta} m_\beta(t, A(t) | \tilde{A}(t - 1), \tilde{S}(t))g^*(A(t) | \tilde{A}(t - 1), \tilde{S}(t)), \quad (8)
\]
van der Laan and Petersen: Origin-Specific Statically Optimal Individualized Treatment Rules

where
\[ g^*(A(t) \mid \bar{A}(t-1), S(t)) \equiv \prod_{j=t}^{K} g^*(A(j) \mid \bar{A}(j-1), S(t)). \]

This implies now the following class of IPTW-estimating functions based on sampling from the actual true probability distribution \( P_{X_0, g_0} \) of \( O \):
\[ D_{b=(h, g^*, Q), IPTW}(O \mid g, \beta) = D_{h, g^*, Q}(O \mid \beta) \frac{g^*(\bar{A} \mid X)}{g(\bar{A} \mid X)}. \]  

This class of IPTW-estimating functions is indexed by arbitrary functions \( b = (h, g^*, Q) \). If we set \( Q = 0 \), then we obtain the following class of IPTW estimating functions indexed by \( (h, g^*) \)
\[ D_{b=(h, g^*)}, IPTW(O \mid g, \beta) = \sum_{t=0}^{K(m)} \frac{h_{t, A}(\bar{S}(t))}{g^*(A(t) \mid A(t-1), X)} \{ Y(t, m) - m_{\bar{A}}(t, A(t) \mid \bar{A}(t-1), \bar{S}(t))^2 \}. \]

If we choose \( h = h^* \) (8), then the corresponding IPTW-estimator defined as the solution of \( 0 = \sum_{t=1}^{n} D_{h, g^*}, IPTW(O \mid g_n, \beta) = 0 \) is the following weighted least squares estimator:
\[ \beta_{n, IPTW} = \arg\min_{\beta} \sum_{t=1}^{n} \sum_{i=1}^{K(m)} w_i(t) \left\{ Y_i(t, m) - m_{\bar{A}}(t, A_i(t) \mid \bar{A}_i(t-1), \bar{S}_i(t)) \right\}^2 \]

with weights given by
\[ w_i(t) = \frac{g^*_n(\bar{A}_i \mid X_i) g^*_n(A_i(t) \mid \bar{A}_i(t-1), \bar{S}_i(t))}{g_n(A_i \mid X_i) g^*_n(A_i(t) \mid A_i(t-1), X_i)}. \]

Note, however, that this weight can be re-written:
\[ w_i(t) = \frac{g^*_n(\bar{A}_i \mid \bar{S}_i) g^*_n(A_i(t) \mid \bar{A}_i(t-1), \bar{S}_i(t))}{g_n(A_i \mid X_i) g^*_n(A_i(t) \mid A_i(t-1), S_i)}. \]
\[ = \frac{g_n(\bar{A}_i \mid \bar{S}_i)}{g_n(A_i \mid X_i)}. \]

Thus this alternative mapping gives back the original IPTW estimating function, given in section 3.1. In the special case that \( S \) is such that \( g^* = g \), then this estimator reduces to the IPTW-estimator proposed in van der Laan et al. (2005) for HA-MSM models (the first ratio now equals 1 in \( w_i(t) \)).

As in the previous subsection, these IPTW estimating functions can be mapped into DR-IPTW estimating functions.
References


Published by The Berkeley Electronic Press, 2007
