Chapter 26
Individualized Antiretroviral Initiation Rules

Romain Neugebauer, Michael J. Silverberg, Mark J. van der Laan

In this chapter, TMLE is illustrated with a data analysis from a longitudinal observational study to investigate "when to start" antiretroviral therapy to reduce the incidence of AIDS-defining cancer (ADC), defined as Kaposi sarcoma, non-Hodgkin's lymphoma, or invasive cervical cancer, in a population of HIV-infected patients. A key clinical question regarding the management of HIV/AIDS is when to start combination antiretroviral therapy (ART), defined in the Department of Health and Human Services (2004) guidelines as a regimen containing three or more individual antiretroviral medications. CD4+ T-cell count levels have been the primary marker used to determine treatment eligibility, although other factors have also been considered, such as HIV RNA levels, history of an AIDS-defining illness (Centers for Disease Control and Prevention 1992), and ability of the patient to adhere to therapy. The primary outcomes considered in ART treatment guidelines described above have always been reductions in HIV-related morbidity and mortality. Until recently, however, guidelines have not considered the effect of CD4 thresholds on the risk of specific comorbidities, such as ADC. In this analysis, we therefore evaluate how different CD4-based ART initiation strategies influence the burden of ADC. We are analyzing ADC here since it is well established that these malignancies are closely linked to immunodeficiency.

We compare the effectiveness in delaying onset of ADC of two clinical guidelines regarding when to start ART. Specifically, the following research question is addressed: Should ART be initiated when a patient's CD4 count drops below 350 cells/μl (current guideline) or should ART initiation be instead delayed until his/her CD4 count drops below 200 cells/μl (official guideline for years 2001–2007)? The target population where this effect is of interest is composed of all patients who are HIV-infected, aged 18 years or older, ART-naive, never diagnosed with ADC and engaged in medical care as demonstrated by receipt of a CD4 test.

M.J. van der Laan and S. Rose, *Targeted Learning: Causal Inference for Observational and Experimental Data*, Springer Series in Statistics, DOI 10.1007/978-1-4419-9782-1_26,

Addressing this research question involves the estimation of the effect of two personalized ART intervention rules (each based on the patients' CD4 count profile over time) on the distribution of the resulting failure times defined as the patients' times to cancer onset. A dynamic marginal structural model (dMSM) provides an adequate causal model for such an effect since dMSMs are models for the distribution of rule-specific counterfactual outcomes. More precisely, each of the two decision rules of interest for when to start ART are indexed by a CD4 count threshold denoted by θ (equal to 200 or 350) and can be described as follows: "Only initiate ART once the patient's CD4 count drops below θ and continue treatment with ART without interruption thereafter."

26.1 Longitudinal Data Structure

This analysis was conducted within Kaiser Permanente of Northern California (KPNC), a large integrated health care delivery system that provides comprehensive medical services to approx. 3.2 million members in a 14-county region in northern California, representing 30% of the surrounding population (N. Gordon, pers. comm.). KPNC maintains complete databases on hospitalizations, outpatient visits, laboratory tests, and prescriptions. Numerous disease registries are maintained at the KPNC Division of Research, including HIV and cancer. For additional details on KPNC's registries and members we refer readers to Selby et al. (2005) and our accompanying technical report: Neugebauer et al. (2010).

KPNC's databases were used to retrospectively identify a cohort of adult HIV-infected patients within KPNC followed between 1996 and 2007 who met the following eligibility criteria: HIV-infected, at least 18 years old, KPNC member during years 1996–2007, never previously treated with antiretrovirals, at least one CD4 count measurement available in the previous year, and never previously diagnosed with an ADC. Based on these eligibility criteria, a total of 6,250 HIV-infected patients were identified. The start of follow-up for patients was the first date at which they met all of the eligibility criteria defined above. Patients were then followed until they achieved the outcome of interest, i.e., incident ADC, or until right censored due to occurrence of a competing event: death, discontinuation of KPNC health insurance, or administrative censoring at the end of the study on December 31, 2007. Small gaps in KPNC health insurance of less than 3 months were ignored, which more likely represented administrative glitches as opposed to lack of health plan coverage. In addition, ART discontinuation lasting less than 6 months was also ignored.

The data for this analysis are viewed as realizations of 6250 i.i.d. copies of the following random longitudinal data structure:

$$O = (\tilde{T}, \Delta, L(0), A(0), L(1), A(1), \ldots, L(t), A(t), \ldots, L(\tilde{T}), A(\tilde{T}), L(\tilde{T}+1)),$$

where $\tilde{T}$ denotes the follow-up time, Δ denotes the indicator of $\tilde{T}$ being equal to the time till cancer onset, $L(0)$ denotes the baseline covariates, $L(t)$ denotes intermediate time-dependent covariates at time t, and $A(t) = (A_1(t), A_2(t) = I(\tilde{T} \leq t, \Delta = 0))$ denotes the indicator of receiving ART at time t ($A_1(t)$) and the indicator of being right-censored at time t ($A_2(t) = I(\tilde{T} \leq t, \Delta = 0)$). Let P_0 be the probability distribution of O.

In this data analysis the time scale is discretized in the sense that t denotes a discrete time stamp that indexes the consecutive intervals of $\tau = 180$ d following a patient's study entry. In particular, $t = 0$ represents the first τ days of a patient's follow-up, and $\tilde{T}$ is measured in units of τ days. The right-censoring time C is defined as the minimum time to a competing event (all expressed in units of τ days since study entry). Thus Δ is the indicator that cancer onset occurs prior to the right-censoring time C.

A patient was defined to initiate ART on the first day of a sequence of at least 2 consecutive months during which the patient was treated with three or more individual antiretroviral medications. The values for all treatment variables before ART initiation were set to 0, i.e., the value for each $A_1(t)$ such that t represents an interval of τ days all of which precede ART initiation was set to 0. Values for all other treatment variables, $A_1(t)$, were mapped to 1 except for all time points t that include or follow the first day of a sequence of 6 consecutive months when the patient had discontinued ART (discontinuation of ART is defined as being treated with less than three antiretroviral medications). For such time points, treatment with ART was deemed interrupted and the value for $A_1(t)$ was set to 0 for t representing the interval of τ days when the first discontinuation of ART was deemed to occur and was set to NA (i.e., considered missing) for all subsequent time points. Note that this missing treatment information could have been recovered but is irrelevant for the estimation of the causal estimand described below and was thus left indeterminate. We also note that $L(t)$ represents subject-matter attributes that occur before the action $A(t)$ at time t and otherwise are assumed not to be affected by the actions at time t or thereafter. In particular, the covariate $L(t)$ contains information on the failure time through the outcome variable $Y(t) = I(\tilde{T} \leq t - 1, \Delta = 1) \in L(t)$, which denotes the indicator of failure at or before time $t - 1$ for $t > 0$ and $Y(0) \equiv 0$ by convention since all patients are cancer free at study entry. The outcome variable, $Y(\tilde{T} + 1)$, is the only covariate relevant at time $\tilde{T} + 1$ for this analysis and information on other attributes at that time is thus ignored, i.e., $L(\tilde{T} + 1) \equiv Y(\tilde{T} + 1)$.

The number of time-independent (e.g., sex, race) and time-dependent (CD4 count, viral load, indicator of past clinical AIDS-defining events, indicator of past ADC diagnosis) covariate attributes are denoted respectively by $p = 15$ and $q = 4$. The covariates in $L(t)$ that represent time-independent attributes are denoted by $L_j(t)$ for $j = 1, \ldots, p$, where j represents an arbitrary order of the time-independent attributes. The covariates in $L(t)$ that represent time-dependent attributes are denoted by $L_j(t)$ for $j = p + 1, \ldots, p + q$ such that j represents an order of the time-dependent attributes, where $L_{p+1}(t)$ and $L_{p+2}(t)$ represent the indicator of past failure (i.e., $L_{p+1}(t) \equiv Y(t)$) and the CD4 count variable, respectively. We thus have $L_j(0)_{j=1,\ldots,p+q} \subset L(0)$, where $L_j(0)$ for $j = 1, \ldots, p$ represent the time-independent

covariates collected at baseline and $L_j(0)$ for $j = p+1, \ldots, p+q$ represent the time-dependent covariates collected at baseline. Similarly, we have $L_j(t)_{j=p+1,\ldots,p+q} \subset L(t)$ for $t = 1, \ldots, \tilde{T}$, where $L_j(t)$ for $j = p+1, \ldots, p+q$ represent the time-dependent covariates collected at time t. Finally, we have $L(\tilde{T}+1) = L_{p+1}(\tilde{T}+1) = Y(\tilde{T}+1)$. Table 26.1 summarizes the link between the measurements of the subject-matter attributes and the notation adopted above to represent these data with the covariates in $L(t)$ for $t = 0, \ldots, \tilde{T}+1$. See the accompanying technical report for a full description of variables.

To respect the time-ordering assumption that imposes that covariates $L(t)$ are not affected by actions at time t and thereafter, the daily data on time-dependent attributes during follow-up were mapped to an observation of $L_j(t)$ for $j \in \{p+1, \ldots, p+q\}$. We will use the notation $[a, b[$ for $\{x : a \leq x < b\}$; thus all points between a and b, including a, but excluding b. For all time points t representing intervals $[t \times \tau, (t+1) \times \tau[$ that do not contain the day when ART is deemed to be initiated, $L_j(t)$ represents the last measurement for attribute j available: (1) at time

Table 26.1 Mapping of the subject-matter attribute measurements into the covariates, $L(t)$, and actions, $A(t)$, of the observed data process O

Attribute	Variable	Number of levels
sex	$L_1(0)$	2
race	$L_2(0)$	4
censusEdu	$L_3(0)$	4
censusPov	$L_4(0)$	4
censusInc	$L_5(0)$	4
riskHIV	$L_6(0)$	4
enrollYear	$L_7(0)$	12
yearsHIV	$L_8(0)$	4
ageAtEntry	$L_9(0)$	3
everSmoke	$L_{10}(0)$	2
everAlcohol	$L_{11}(0)$	2
everDrug	$L_{12}(0)$	2
everHepatitisB	$L_{13}(0)$	2
everHepatitisC	$L_{14}(0)$	2
everObese	$L_{15}(0)$	2
Y	$L_{16}(t)$ for $t = 0, \ldots, \tilde{T}+1$	$n(t, 16) \equiv 2$
CD4	$L_{17}(t)$ for $t = 0, \ldots, \tilde{T}$	$n(t, 17) \equiv 4$
VL	$L_{18}(t)$ for $t = 0, \ldots, \tilde{T}$	$n(t, 18) \equiv 4$
clinicalAIDS	$L_{19}(t)$ for $t = 0, \ldots, \tilde{T}$	$n(t, 19) \equiv 2$
I.CD4	$L_{20}(t)$ for $t = 0, \ldots, \tilde{T}$	$n(t, 20) \equiv 2$
I.VL	$L_{21}(t)$ for $t = 0, \ldots, \tilde{T}$	$n(t, 21) \equiv 2$
I.race	$L_{22}(0)$	2
I.censusEdu	$L_{23}(0)$	2
I.censusPov	$L_{24}(0)$	2
I.censusInc	$L_{25}(0)$	2
I.riskHIV	$L_{26}(0)$	2
ART	$A_1(t)$ for $t = 0, \ldots, \tilde{T}$	2

$t-1$ (i.e., the last measurement obtained during interval $[(t-1)\times\tau, t\times\tau[$) if $t > 0$ and (2) within the year preceding study entry if $t = 0$. For the time point t representing the interval $[t \times \tau, (t + 1) \times \tau[$ that contains the day when ART is deemed to be initiated, $L_j(t)$ represents the last measurement for attribute j available (1) at time $t - 1$ or time t but always prior to the actual day when ART was initiated if $t > 0$ and (2) within the year preceding study entry or at time point 0 but always prior to the actual day when ART was initiated if $t = 0$.

Once this mapping was implemented, some of the observations for $p' = 5$ time-independent attributes (race, censusEdu, censusPov, censusInc, and riskHIV) were missing. Similarly, some of the observations for $q' = 2$ time-dependent attributes (CD4 and VL) were missing. Such observations were imputed (e.g., with the mode or with the average of the nonmissing observations), and indicators of imputation (named I.censusEdu, I.censusPov, I.censusInc, I.race, I.riskHIV, I.CD4, and I.VL, respectively) were created. For the covariates race and riskHIV, the imputation was implemented conditional on the covariate sex. Each of these imputation indicators are denoted by $L_j(t)$ such that $L_j(t)$ for $j = p + q + 1, \ldots, p + q + q'$ represent the indicators of imputation for the time-dependent attributes ordered arbitrarily (CD4 and VL), and $L_j(t)$ for $j = p+q+q'+1, \ldots, p+q+q'+p'$ represent the indicators of imputation for the time-independent attributes ordered arbitrarily (race, censusEdu, censusPov, censusInc, and riskHIV). These imputation indicators are included in the definition of $L(t)$.

The following forward imputation method was used for missing observations of time-dependent covariates at any time $t > 0$: the missing observation for $L_j(t)$ was imputed with the last nonmissing and nonimputed observation of the covariates $(L_j(0), \ldots, L_j(t-1))$ if available and otherwise with the imputed observation for $L_j(0)$. For additional detailed descriptions of the variables and a variety of summary statistics for the data, we refer to the accompanying technical report. Here, we suffice with mentioning that two thirds of the patients who experienced cancer onset before right censoring did so during the first two years of follow-up and that from the 118 patients who experienced a cancer-onset event in the first two years of follow-up, 13 occurred in patients who only followed the rule for starting ART indexed by a CD4 count threshold of 200, 12 occurred in patients who only followed the rule for starting ART indexed by the CD4 count threshold of 350, 51 occurred in patients who followed both rules for starting ART, and 42 occurred in patients who followed neither of these two when-to-start rules. This simple summary provides an indication that the data are sparse with respect to the scientific question of interest, and that the data lack a strong signal for favoring one dynamic treatment over the other. Nevertheless, it is of interest to determine if these data imply a narrow confidence interval around a zero treatment effect.

For conciseness, we adopt the following shorthand notation to represent the history of measurements on a given subject-matter attribute between time point 0 and t: for a time-dependent process $X()$, $\bar{X}(t) \equiv (X(0), \ldots, X(t))$ and by convention $X(t)$ is nil for $t < 0$. Using this notation, the observed data for one subject can be summarized as follows: $O = (\tilde{T}, \Delta, \bar{L}(\tilde{T} + 1), \bar{A}(\tilde{T}))$.

26.2 Likelihood of the Observed Data Structure

Following the time ordering of actions and covariates encoded in the directed acyclic graph implied by Fig. 26.1, the likelihood of the observed data under a probability distribution P can be factorized as

$$P(O) = \prod_{t=0}^{\tilde{T}+\Delta} P(L(t) \mid \bar{L}(t-1), \bar{A}(t-1)) \prod_{t=0}^{\tilde{T}} P(A(t) \mid \bar{A}(t-1), \bar{L}(t)).$$

Note that the first product of conditional probabilities ends with the conditional probability of $L(\tilde{T})$ given all past actions, $\bar{A}(\tilde{T}-1)$, and past covariates, $\bar{L}(\tilde{T}-1)$, when the patient's data are right-censored, i.e., $\Delta = 0$, since $L(\tilde{T}+1) \equiv Y(\tilde{T}+1)$ is then 0 with probability one. If the patient's cancer onset is observed (i.e., $\Delta = 1$), the first product ends with the conditional probability of $L(\tilde{T}+1)$ given past actions, $\bar{A}(\tilde{T})$, and past covariates, $\bar{L}(\tilde{T})$.

The two products of the likelihood above are referred to as the Q-factor and g-factor of the likelihood. The Q-factor of the likelihood is composed of the product of the conditional probabilities of covariates given past covariates and actions, whereas the g-factor of the likelihood is composed of the product of the conditional probabilities of actions given past actions and covariates. The g-factor of the likelihood is also referred to as the action mechanism. Following this terminology, the notation for the likelihood of the observed data can be summarized as

$$P = \prod_{t=0}^{K} Q_{L(t)} \prod_{t=0}^{K} g_{A(t)}, \tag{26.1}$$

where the conditional probability distributions $P(L(t) \mid \bar{L}(t-1), \bar{A}(t-1))$ and $P(A(t) \mid \bar{A}(t-1), \bar{L}(t))$ are denoted by $Q_{L(t)}$ and $g_{A(t)}$, respectively, and K is the last possible time point so that $P(\tilde{T} < K) = 1$. After $t > \tilde{T} + \Delta$, all conditional distributions $Q_{L(t)}$

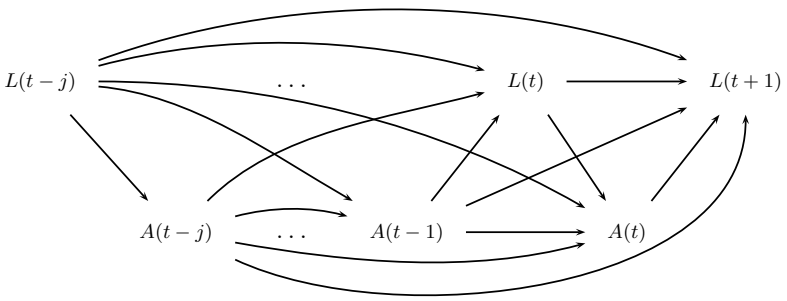

Fig. 26.1 Template of the directed acyclic graph that encodes the time ordering of all variables of the observed data process O. The complete graph can be derived by sequentially drawing the nodes and arcs implied by this template for $t = 0, \ldots, \tilde{T}$ and $j \leq t$

and $g_{A(t)}$ are degenerate at an arbitrarily defined value, such as the last measured value.

At each time point t, the covariate $L(t)$ is composed of a collection of discrete covariates denoted by $L_j(t)$. Specifically, $L(0) \equiv (L_j(0) : j = 1, \ldots, p + q + q' + p')$, $L(t) \equiv (L_j(t) : j = p+1, \ldots, p+q+q')$ for $t = 1, \ldots, \tilde{T}$, and $L(\tilde{T}+1) \equiv L_{p+1}(\tilde{T}+1)$, where, as described in Table 26.1, for $p = 15$, $q = 4$, $q' = 2$, and $p' = 5$: $j = 1, \ldots, p$ indexes the covariates that represent the time-independent attributes listed; $j = p+1, \ldots, p+q$ indexes the covariates that represent the time-dependent attributes listed such that $p+1$ represents Y and $p+2$ represents CD4; $j = p+q+1, \ldots, p+q+q'$ indexes the indicators of imputations for the q' time-dependent attributes that have missing observations; $j = p+q+q'+1, \ldots, p+q+q'+p'$ indexes the indicators of imputations for the p' time-independent attributes that have missing observations.

The factors $Q_{L(t)}$ of the likelihood for $t = 1, \ldots, \tilde{T} + \Delta$ can thus be factorized as $Q_{L(t)}(O) = \prod_{j=1}^{n_{\tilde{T}}(t)} Q_{L_{p+j}(t)}(O)$, where $n_{\tilde{T}}(t) \equiv q + q'$ for $t = 1, \ldots, \tilde{T}$, $n_{\tilde{T}}(\tilde{T}+1) \equiv 1$, and $Q_{L_{p+j}(t)}(O) \equiv P(L_{p+j}(t) \mid Pa(L_{p+j}(t)))$ with $Pa(L_{p+j}(t)) \equiv (\bar{L}(t-1), L_{p+1}(t), \ldots, L_{p+j-1}(t), \bar{A}(t-1))$. Note that the notation above makes implicit use of the convention that $(L_j(t), \ldots, L_{j'}(t))$ for $j' < j$ is nil.

Similarly, at each time point t, the action $A(t)$ is composed of a treatment, $A_1(t)$, and an indicator of right censoring, $A_2(t)$. The factors $g_{\bar{A}(t)}$ of the likelihood can thus be factorized as $g_{A(t)} = g_{A_1(t)} g_{A_2(t)}$, where $g_{A_1(t)}(O) \equiv P(A_1(t) \mid Pa(A_1(t)))$, $Pa(A_1(t)) = (\bar{A}(t-1), \bar{L}(t), A_2(t)))$, $g_{A_2(t)}(O) \equiv P(A_2(t) \mid Pa(A_2(t)))$, and $Pa(A_2(t)) = (\bar{A}(t-1), \bar{L}(t))$. This yields the following factorization of likelihood (26.1):

$$P(O) = Q_{L(0)}(L(0) \prod_{t=1}^{\tilde{T}+\Delta} \prod_{j=1}^{n_{\tilde{T}}(t)} Q_{L_{p+j}(t)}(O) \prod_{t=0}^{\tilde{T}} g_{A_1(t)}(O) g_{A_2(t)}(O). \tag{26.2}$$

Since each covariate $L_{p+j}(t)$ for $t = 1, \ldots, \tilde{T} + \Delta$ and $j = 1, \ldots, n_{\tilde{T}}(t)$ is discrete with $n(t, p+j)$ categories, it can be recoded with $n(t, p+j) - 1$ binary variables: $L_{p+j,m}(t) \equiv I(L_{p+j}(t) = m)$ for $m = 1, \ldots, n(t, p+j) - 1$, i.e. $L_{p+j}(t) = (L_{p+j,m}(t))_{m=1,\ldots,n(t,p+j)-1}$. This recoding of the information in $L_{p+j}(t)$ leads to the following factorization $Q_{L_{p+j}(t)} = \prod_{m=1}^{n(t,p+j)-1} Q_{L_{p+j,m}(t)}$, where $Q_{L_{p+j,m}(t)}$ represents the conditional probability of $L_{p+j,m}(t)$ given $Pa(L_{p+j}(t))$ and $L_{p+j,l}(t)$ for $l = 1, \ldots, m-1$. Note that this conditional probability is degenerate, i.e., equal to 1 at $L_{p+j,m}(t) = 0$, if one of the indicators $L_{p+j,l}(t)$ in $Pa(L_{p+j,m}(t))$ is 1. Note also that if $L_{p+j}(t)$ is binary ($n(t, p+j) = 2$), then $L_{p+j,1}(t) = L_{p+j}(t)$ and $Q_{L_{p+j,1}(t)} = Q_{L_{p+j}(t)}$. This provides us with the following factorized likelihood in terms of conditional distributions of binary variables:

$$P(O) = Q_{L(0)}(L(0)) \prod_{t=1}^{\tilde{T}+\Delta} \prod_{j=1}^{n_{\tilde{T}}(t)} \prod_{m=1}^{n(t,p+j)-1} Q_{L_{p+j,m}(t)}(O) \prod_{t=0}^{\tilde{T}} g_{A_1(t)}(O) g_{A_2(t)}(O). \tag{26.3}$$

26.3 Target Parameter

In this section, the causal effect of interest in this analysis is defined as the effect of dynamic ART interventions on the cumulative risk of cancer in the first two years of follow-up. Under an identifiability assumption, this causal effect is expressed as a function of the likelihood of the observed data. This latter estimand is referred to as the target parameter. We can define an SCM for the full data (U, O) in terms of a nonparametric structural equation model: $L(0) = f_{L(0)}(U_{L(0)})$, $L(t) = f_{L(t)}(Pa(L(t)), U_{L(t)})$, $t = 1, \ldots, \tilde{T} + \Delta$, $A(t) = f_{A(t)}(Pa(A(t)), U_{A(t)})$, $t = 0, \ldots, \tilde{T}$. This SCM can also be defined for $t = 1, \ldots, K$, by defining the equations as degenerate for $t > \tilde{T} + \Delta$, and, it can also be extended to one equation for each univariate time-dependent covariate.

Clinical practice often involves treatment decisions that are continuously adjusted to the patient's evolving medical history (e.g., new diagnoses and laboratory values) and are not set a priori at baseline. Thus, it may often be less clinically relevant to compare the health effect of static treatment interventions than to compare the effectiveness of competing medical guidelines, i.e., adaptive treatment strategies that map the patient's unfolding medical history to subsequent treatment decisions. Following such treatment strategies leads to treatment interventions over time which are referred to as dynamic interventions since the treatment experienced by each patient at any point in time is not set a priori at baseline but is rather adjusted based on the patient's current circumstances.

Our aim was to evaluate the comparative effectiveness between ART initiation strategies guided by the patient's evolving CD4 count. These adaptive treatment strategies are referred to as individualized action rules. The individualized action rules of interest are each indexed by a CD4 count threshold, denoted by $\theta \in \Theta$, and are each defined as a vector function $d_\theta = (d_\theta(0), \ldots, d_\theta(K))$ where each function, $d_\theta(t)$ for $t = 0, \ldots, K$, is a decision rule for determining the action (treatment and right censoring) to be experienced by a patient during time interval t. A decision rule $d_\theta(t)$ maps the action and covariate history measured up to a given time interval t to an action regimen (i.e., an intervention) during time interval t: $d_\theta(t) : (\bar{L}(t), \bar{A}(t-1)) \mapsto (a_1(t), a_2(t))$. In this analysis, the decision rules of interest are defined such that $d_\theta(t)((\bar{L}(t), \bar{A}(t-1))$ is:

- $(a_1(t), a_2(t)) = (0, 0)$ (i.e., no ART use and no right censoring) if and only if the patient was not previously treated with ART [i.e. $\bar{A}(t-1) = 0$] and the previous CD4 count measurement was greater than or equal to the threshold θ (i.e., $L_{p+1}(t) \geq \theta$);
- $(a_1(t), a_2(t)) = (1, 0)$ (i.e., ART use and no right censoring) otherwise.

The individualized action rules, d_θ for $\theta \in \Theta$, implied by the time-specific decision rules above, $d_\theta(t)$ for $t = 0, \ldots, K$, are monotone in the sense that if a patient follows one of these rules, d_θ, then s/he is not treated with ART until his/her CD4 count falls below the threshold θ for the first time and from then on the patient remains treated with ART.

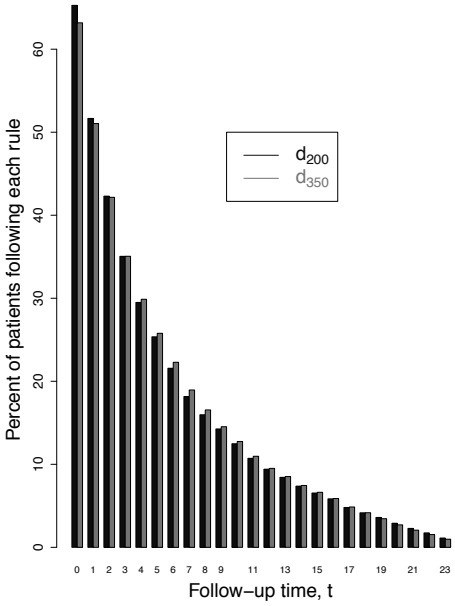

Fig. 26.2 Percentages of patients following each of the two individualized action rules of interest indexed by a CD4 count threshold of 200 and 350 cells per μl (reminder: $t = 0$ represents the first 180 d of follow-up)

The counterfactual covariate process that could be observed on a patient in an ideal experiment where interventions on the action process according to a decision rule d_θ are carried out through time K is denoted by $\bar{L}_{d_\theta}(K+1)$. Note that such dynamic treatment interventions through time K according to the adaptive treatment strategy d_θ are only functions of the observed covariate process $\bar{L}(K)$ and are thus denoted by $d_\theta(\bar{L}(K))$. Failure may occur during follow-up under any such dynamic treatment intervention. Such *counterfactual* failure times are denoted by $T_{d_\theta} \leq K$ and defined by $(\bar{Y}_{d_\theta}(T_{d_\theta}) = 0, Y_{d_\theta}(T_{d_\theta}+1) = 1, \ldots, Y_{d_\theta}(K+1) = 1)$. This random process $(T_{d_\theta}, L_{d_\theta} = \bar{L}_{d_\theta}(K+1))$ is defined in terms of the postintervention distribution of the above-stated SCM for the full-data (U, O). A causal effect can now be defined as a target parameter for the full-data SCM:

$$P(Y_{d_\theta}(t+1) = 1) - P(Y_{d_{\theta'}}(t+1) = 1) = P(T_{d_\theta} \leq t) - P(T_{d_{\theta'}} \leq t)$$

for $t = 0, \ldots, K$ and any two different individualized action rules d_θ with $\theta \in \Theta$ and $d_{\theta'}$ with $\theta' \in \Theta$.

For our research question, the comparative effectiveness between the two adaptive ART strategies indexed by the CD4 thresholds 200 and 350 is of interest. Figure 26.2 represents the percentages of patients following the corresponding two individualized action rules over time denoted by d_{200} and d_{350}. The empirical distri-

bution of observed events indicates that two thirds of the observed cancers occurred in the first four time points of follow-up. To ease the computing time required for implementation of the TMLE procedure while illustrating the computing steps involved without loss of generality, we restrict the focus of this analysis to the evaluation of a single causal contrast between the cumulative risks of failure within 2 years of study entry under the individualized action rules d_{350} and d_{200}:

$$\begin{aligned}\psi^F &= P(T_{d_{350}} \leq 3) - P(T_{d_{200}} \leq 3)\\ &= P(Y_{d_{350}}(4) = 1) - P(Y_{d_{200}}(4) = 1).\end{aligned} \tag{26.4}$$

Note that this parameter is a function of the full-data distribution as is made explicit with the notation $\psi^F = \Psi^F(P_{X,U})$. This target parameter would be identifiable under this full-data SCM if we observed the full-data (U, O). However, we only observe the O-component of the full data.

Under the sequential randomization assumption (SRA)

$$A(t) \perp (Y_{d_\theta}, L_{d_\theta}) \mid \bar{L}(t), \bar{A}(t-1), \text{ for all } t = 0, \ldots, K,$$

the marginal distribution of the counterfactual process $(\bar{L}_{d_\theta}(t+1))$ with $t \leq K$ is identified by the g-computation formula:

$$P_{\bar{L}_{d_\theta}(t+1)} = Q_{L(0)} \prod_{t'=1}^{t+1} \prod_{j=1}^{n_t(t')} \prod_{m=1}^{n(t',p+j)-1} Q_{L_{p+j,m}(t'),d_\theta}, \tag{26.5}$$

where the conditional probabilities $Q_{L_{p+j,m}(t'),d_\theta}$ are the conditional probabilities $Q_{L_{p+j,m}(t')}$ defined in Sect. 26.2 where the action values are set according to the decision rule d_θ in the conditioning events. Recall that the covariate $L_{p+1,d_\theta}(t)$ of $L_{d_\theta}(t)$ corresponds with the outcome variable also denoted by $Y_{d_\theta}(t)$ and that $P(T_{d_\theta} = t) = P(\bar{Y}_{d_\theta}(t) = 0, Y_{d_\theta}(t+1) = 1)$. By integrating over all covariate values that are consistent with $T_{d_\theta} = t$ of the g-computation formula (26.5), the following probability $P(T_{d_\theta} = t)$ is obtained:

$$P(T_{d_\theta} = t) = \sum_{\{\bar{l}(t+1):y(t)=0,y(t+1)=1\}} Q_{L(0)}(l(0)) \prod_{t'=1}^{t+1} \prod_{j=1}^{n_t(t')} \prod_{m=1}^{n(t',p+j)-1} Q_{L_{p+j,m}(t'),d_\theta}(\bar{l}(t+1)). \tag{26.6}$$

We denote the right-hand side of equality (26.6) with $\Phi(Q, t, \theta)$ to make explicit that this parameter of P only depends on P through its $Q(P)$-factor. Using that $P(Y_{d_\theta}(t+1) = 1) = \sum_{t'=0}^{t} P(\bar{Y}_{d_\theta}(t') = 0, Y_{d_\theta}(t'+1) = 1)$ and equality (26.6), the causal estimand defined by (26.4) can be expressed as a function of the Q-factor of the observed data likelihood under the SRA:

$$\psi = \sum_{t=0}^{3} \Phi(Q, t, 350) - \sum_{t=0}^{3} \Phi(Q, t, 200). \tag{26.7}$$

The parameter $\psi = \Psi(Q)$, as defined by equality (26.7), is the target parameter in this analysis. It is a mapping from the Q-part of a probability distribution P of the observed data structure O into a one-dimensional euclidean parameter, as made explicit by the notation $\psi = \Psi(Q)$. We denote the true value of this target parameter by $\psi_0 = \Psi(Q_0)$ where Q_0 denotes the Q-part of the likelihood of the observed data under the true distribution P_0. Note that under the SRA, the target parameter value ψ_0 can be interpreted causally since it then corresponds to the causal parameter $\psi_0^F = \Psi^F(P_{X,U,0})$ from equality (26.4).

26.4 IPCW-R-TMLE

In this section, we develop an inverse probability of action-weighted reduced-data targeted maximum likelihood estimator (IPAW-R-TMLE) for estimation of the above-defined target parameter ψ_0. Implementation of this IPAW-R-TMLE is illustrated with data from the KPNC electronic medical record. The IPAW-R-TMLE is a weighted TMLE applied to so-called reduced data that corresponds to the original data where the time-dependent covariates are ignored, with the exception of the time-dependent CD4 count and the outcome process (van der Laan 2008b). The weights applied to the reduced-data TMLE (R-TMLE) permit adjustment for the other time-dependent confounders that could not be accounted for by the R-TMLE because of this data-reduction step.

The IPAW-R-TMLE is a targeted minimum-loss-based estimator as presented in Appendix A, where we provide detailed explanation and understanding of IPAW-R-TMLE (see also Chap. 24). The technical report provides additional detail regarding the specific implementation carried out here. It involves the following steps. Firstly, the observed data structure is reduced by replacing the time-dependent covariates L in O by a reduced-data time-dependent covariate L^r, resulting in a reduced observed data structure O^r. The causal quantity of interest is represented as a parameter of the probability distribution Q_0^r, a function of $O^r = (L^r, A)$, where $Q_0^r(l^r, a) = \prod_t Q_0^r(l^r(t) \mid \bar{l}^r(t-1), \bar{a}(t-1))$ represents the probability that the counterfactual L_a^r equals l^r. Thus Q_0^r is a function of the full data distribution of the counterfactuals $(L_a : a)$. A possible loss function for Q_0^r is given by $-g^r/g_0 log Q$, but more stable time-dependent weighting schemes will be employed, resulting in a specified loss function $L_{w_0}(Q)$ relying on a weight function $(w_0(t) : t)$, where $w_0(t) = \prod_{s \le t-1} g_{0,A(s)}^r / \prod_{s \le t-1} g_{0,A(s)}$ is indexed by g^r and g_0. Specifically, the minus log of the conditional distribution of $L_a^r(t)$, given $\bar{L}_a^r(t-1)$, in $Q^r(l^r, a) = \prod_t Q^r(l^r(t) \mid \bar{l}^r(t-1), \bar{a}(t-1))$ is weighted by the corresponding $w_0(t)$ for that time point. This loss function $L_{w_0}(Q)$ is valid for each choice of g^r, but it relies on correct specification of the true action mechanism g_0. Let w_n be an estimator of the weight function w_0, thereby using the actual observed data $O_1, \ldots, O_n$. We select a parametric model $(Q(\epsilon) : \epsilon)$ so that $\frac{d}{d\epsilon} \log Q(\epsilon)$ at $\epsilon = 0$ spans the efficient influence curve of the target parameter for the reduced data in the special case that there is no time-dependent confounding beyond L^r, or, equivalently, $w_0 = 1$. This

now defines the quantity Q_0^r, the target parameter $\psi_0^r = \Psi(Q_0^r)$, the loss function $L_{w_0}(Q)$ for Q_0^r, and the parametric working model $\{Q(\epsilon) : \epsilon\}$, so that the targeted minimum-loss-based estimator is defined, and it will solve the so-called IPAW-R-efficient influence curve estimating equation:

$$0 = P_n \frac{d}{d\epsilon} L_{w_n}(Q_n^*(\epsilon))\Big|_{\epsilon=0}.$$

Section 26.4.1 describes the R-TMLE on which the IPAW-R-TMLE is based. The R-TMLE is identical to the actual targeted maximum likelihood estimator applied to the reduced data, relying on the log-likelihood loss function, using the parametric working logistic regression models defined by clever time-dependent covariates in order to fluctuate the conditional distributions of binary variables. This TMLE applied to reduced data would be consistent if $w_0 = 1$, but it is biased in general, due to ignoring the time-dependent covariates that were removed from the data structure. Section 26.4.2 describes the implementation of the corresponding IPAW-R-TMLE of ψ_0, which simply involves applying the estimated weights w_n to the R-TMLE. The results of the application of both the R-TMLE as well as the IPAW-R-TMLE are provided.

26.4.1 R-TMLE Implementation and Results

The R-TMLE described below corresponds with the TMLE applied to the simplified data where the only time-dependent covariates considered past baseline represent outcome and CD4 count measurements. Below, we describe the steps involved in the implementation of this R-TMLE starting with the presentation of the IPAW-estimating function, the corresponding derivation of the efficient influence curve, the parametric working model used to fluctuate the initial estimator defined in terms of logistic regression models for each binary conditional distribution, using a clever covariate, and, finally, the implementation of the iterative TMLE algorithm as defined by this parametric fluctuation model and the log-likelihood loss function.

For clarity, we abuse the notation previously introduced and associated with the original data structure to describe the R-TMLE based on the reduced-data structure. All reference to the covariates $L_{p+j}(t)$ for $j = 3, \ldots, q + q'$ and $t > 0$ in the notation below should thus be ignored since such variables are considered nil in this subsection.

IPAW estimating function. An IPAW estimating function for the target parameter ψ is defined as

$$D_{\text{IPAW}}(g, \psi) \equiv \frac{I(\bar{A}(\check{T}) = d_{350}(\bar{L}(\check{T})))}{\prod_{t=0}^{\check{T}} g_{A_1(t)} g_{A_2(t)}} Y(\check{T} + 1)$$

$$-\frac{I(\bar{A}(\check{T}) = d_{200}(\bar{L}(\check{T})))}{\prod_{t=0}^{\check{T}} g_{A_1(t)} g_{A_2(t)}} Y(\check{T}+1) - \psi, \tag{26.8}$$

where $\check{T}$ is defined as the minimum between the follow-up time and 3, i.e., $\check{T} = \min(3, \tilde{T})$. Recall that the outcome variable $Y(t+1)$ is also denoted by $L_{p+1}(t+1)$ and $L_{p+1,1}(t+1)$ (Sect. 26.2). We use these three notations for the same variable interchangeably in the following sections.

Efficient influence curve and clever covariate. As described in Sect. 3.2 of van der Laan (2010a), the IPAW estimating function can be mapped into the efficient (relative to the reduced data) influence curve for ψ, denoted by $D^*(Q, g, \psi)$, by projecting it onto the tangent space of Q: $D^*(Q, g, \psi) = \Pi(D_{\text{IPAW}} \mid T_Q)$, where D_{IPAW} is shorthand notation for $D_{\text{IPAW}}(g, \psi)$ [definition (26.8)]. Theorem 2 in van der Laan (2010a) applied to factorization (26.3) of the likelihood of the reduced observed data leads to the following result:

$$\Pi(D_{\text{IPAW}} \mid T_Q) = \Pi(D_{\text{IPAW}} \mid T_{L(0)}) + \sum_{t=1}^{\tilde{T}} \sum_{j=1}^{2} \sum_{m=1}^{n(t,p+j)-1} \Pi(D_{\text{IPAW}} \mid T_{L_{p+j,m}(t)}) + \Delta\Pi(D_{\text{IPAW}} \mid T_{L_{p+1,1}(\tilde{T}+1)}),$$

where we have that $\Pi(D_{\text{IPAW}} \mid T_{L(0)}) = E(D_{\text{IPAW}} \mid L(0))$ and $\Pi(D_{\text{IPAW}} \mid T_{L_{p+j,m}(t)}) = H^*_{L_{p+j,m}(t)}(L_{p+j,m}(t) - Q_{L_{p+j,m}(t)}(1))$ for $t = 1, \ldots, \tilde{T}$, $j = 1, 2$ and $m = 1, \ldots, n(t, p+j)-1$ or $(t, j, m) = (\tilde{T}+1, 1, 1)$, with $Q_{L_{p+j,m}(t)}(1)$ representing the conditional probability of $L_{p+j,m}(t)$ defined in Sect. 26.2 and evaluated at $L_{p+j,m}(t) = 1$. Here $H^*_{L_{p+j,m}(t)}$ is defined as the following function of $Pa(L_{p+j,m}(t))$:

$$E(D'_{\text{IPAW}} \mid L_{p+j,m}(t) = 1, Pa(L_{p+j,m}(t))) - E(D'_{\text{IPAW}} \mid L_{p+j,m}(t) = 0, Pa(L_{p+j,m}(t))), \tag{26.9}$$

where $D'_{\text{IPAW}}(g)$ equals $D_{\text{IPAW}}(g, \psi)$ with $\psi = 0$. Note that $D'_{\text{IPAW}}(g)$ is only a function of the observed data collected up to $\check{T}+1$ which excludes $L_{p+j,m}(t)$ for $t > \check{T}+1$. As a result, equality (26.9) for $t > \check{T}+1$ becomes $H^*_{L_{p+j,m}(t)} = 0$. Note also that $D'_{\text{IPAW}}(g)$ can be represented as

$$D'_{\text{IPAW}}(O \mid g) = \frac{D_1(\bar{A}(\check{T}), \bar{L}(\check{T}+1))}{\prod_{t=0}^{\check{T}} g_{A_1(t)} g_{A_2(t)}},$$

with $D_1(\bar{A}(\check{T}), \bar{L}(\check{T}+1))$ defined as

$$\left(I(\bar{A}(\check{T}) = d_{350}(\bar{L}(\check{T}))) - I(\bar{A}(\check{T}) = d_{200}(\bar{L}(\check{T})))\right) Y(\check{T}+1).$$

From Theorem 2 in van der Laan (2010a), equality (26.9) can thus be rewritten as:

$$H^*_{L_{p+j,m}(t)} = \frac{1}{\prod_{t'=0}^{t-1} g_{A_1(t')} g_{A_2(t')}} \times \tag{26.10}$$

$$\Big(E(\sum_{\bar{a}(t,3)} D_1 \mid L_{p+j,m}(t)=1, Pa(L_{p+j,m}(t))) - E(\sum_{\bar{a}(t,3)} D_1 \mid L_{p+j,m}(t)=0, Pa(L_{p+j,m}(t)))\Big),$$

where D_1 is shorthand notation for $D_1(\bar{A}(t-1), \bar{a}(t,\check{T}), \bar{L}(\check{T}+1))$ and $\bar{a}(t,t') = (a(t),\ldots,a(t'))$ for $t' \geq t$ and nil otherwise. In addition, note that the second expectation in (26.10) is 0 when $j=1$ and $t=4$, and we have

$$H^*_{L_{p+1,1}(4)} = \frac{I(\bar{A}(\check{T}) = d_{350}(\bar{L}(\check{T}))) - I(\bar{A}(\check{T}) = d_{200}(\bar{L}(\check{T})))}{\prod_{t'=0}^{\check{T}} g_{A_1(t')} g_{A_2(t')}}.$$

Finally, note that both expectations in (26.10) are equal when $j=2$ and $t=4$. We thus have $H^*_{L_{p+2,m}(4)} = 0$.

From all the results above, the efficient (relative to the reduced data) influence curve for ψ, $D^*(Q,g,\psi)$ is defined as

$$\begin{aligned} &\Pi(D_{\text{IPAW}} \mid T_{L(0)}) + \sum_{t=1}^{\check{T}} \sum_{j=1}^{2} \sum_{m=1}^{n(t,p+j)-1} \Pi(D_{\text{IPAW}} \mid T_{L_{p+j,m}(t)}) \\ &+ \Delta^{I(\tilde{T}\leq 3)} \Pi(D_{\text{IPAW}} \mid T_{L_{p+1,1}(\check{T}+1)}), \end{aligned} \tag{26.11}$$

with $I(\tilde{T} \leq 3)$ representing the indicator that the follow-up time $\tilde{T}$ is lower than or equal to 3,

$$\Pi(D_{\text{IPAW}} \mid T_{L(0)}) = E(D_{\text{IPAW}} \mid L(0)),$$

$$\Pi(D_{\text{IPAW}} \mid T_{L_{p+j,m}(t)}) = H^*_{L_{p+j,m}(t)}(L_{p+j,m}(t) - Q_{L_{p+j,m}(t)}(1)), \text{ where}$$

for $t = 1, 2, 3$

$$H^*_{L_{p+j,m}(t)} = \frac{1}{\prod_{t'=0}^{t-1} g_{A_1(t')} g_{A_2(t')}} \times \tag{26.12}$$

$$\Big(E_Q(\sum_{\bar{a}(t,3)} D_1 \mid L_{p+j,m}(t)=1, Pa(L_{p+j,m}(t))) - E_Q(\sum_{\bar{a}(t,3)} D_1 \mid L_{p+j,m}(t)=0, Pa(L_{p+j,m}(t)))\Big),$$

and

$$H^*_{L_{p+1,1}(4)} = \frac{I(\bar{A}(3) = d_{350}(\bar{L}(3))) - I(\bar{A}(3) = d_{200}(\bar{L}(3)))}{\prod_{t'=0}^{3} g_{A_1(t')} g_{A_2(t')}}. \tag{26.13}$$

The variables $H^*_{L_{p+1,1}(t)}$ and $H^*_{L_{p+2,m}(t)}$ are the clever covariates that are used for updating the initial estimators of $Q_{L_{p+1,1}(t)}$ and $Q_{L_{p+2,m}(t)}$, respectively, during the implementation of the R-TMLE of the target parameter ψ_0.

Obtain an initial estimate Q_n^0 of Q_0. The efficient influence curve defined by equality (26.11) is a function of only a subset of the Q components of the reduced, observed data likelihood [see equality (26.3) tailored to the reduced-data structure]. Specifically, the following 14 components of Q are relevant for implementation of the R-TMLE of the target parameter and thus need to be estimated:

- $Q_{L(0)} \equiv P(L(0))$.
- $Q_{L_{p+1,1}(t)} \equiv P(Y(t) \mid \bar{L}(t-1), \bar{A}(t-1))$ for $t = 1, 2, 3, 4$ [only relevant at $\bar{Y}(t-1) = 0$, $\bar{A}_2(t-1) = 0$, and nonmissing $\bar{A}_1(t-1)$]. Recall that treatment is coded as missing after first ART discontinuation; see Sect. 26.1.
- $Q_{L_{p+2,m}(t)} \equiv P(I(L_{p+2}(t) = m) \mid \bar{L}(t-1), Y(t), \bar{A}(t-1), I(L_{p+2}(t) = 1), \ldots, I(L_{p+2}(t) = m-1))$ for $t = 1, 2, 3$ and $m = 1, 2, 3$ [only relevant at $\bar{Y}(t) = 0$, $\bar{A}_2(t-1) = 0$ and nonmissing $\bar{A}_1(t-1)$, and only unknown at $I(L_{p+2}(t) = 1) = 0, \ldots, I(L_{p+2}(t) = m-1) = 0$].

Thus, estimation of the target parameter $\psi_0 = \Psi(Q_0)$ with the R-TMLE relies on the initial estimation of the corresponding 14 Q_0-components of the true reduced-data-generating distribution, P_0. The initial estimate of $Q_{0,L(0)}$ is denoted by $Q^0_{L(0),n}$. It is defined based on nonparametric estimation of $Q_{0,L(0)}$ with the empirical distribution of $L(0)$. The initial estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ are denoted by $Q^0_{L_{p+1,1}(t),n}$ and $Q^0_{L_{p+2,m}(t),n}$, respectively. They are defined based on sieve estimation of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ with the DSA algorithm as described below.

The initial estimate of $Q_{0,L_{p+1,1}(1)}$ (at $A_2(0) = 0$ and nonmissing $A_1(0)$) is obtained separately. The DSA algorithm was set so that it searched among main term logistic regression fits and reports the best fit of size s, $s = 1, \ldots, 10$, using deletion, substitution, and addition moves. Cross-validation was used to determine the best choice of size s.

This DSA algorithm was also used to obtain the remaining 12 estimates based on the following three data-pooling schemes:

- The initial estimates of $Q_{0,L_{p+1,1}(t)}$ for $t = 2, 3, 4$ (at $\bar{Y}(t-1) = 0$, $\bar{A}_2(t-1) = 0$, and nonmissing $\bar{A}_1(t-1)$) are obtained simultaneously through a DSA-selected, pooled estimator over the three time intervals, t.
- The initial estimates of $Q_{0,L_{p+2,m}(1)}$ for $m = 1, 2, 3$ (at $Y(1) = 0$, $A_2(0) = 0$, nonmissing $A_1(0)$, and $I(L_{p+2}(t) = 1) = 0, \ldots, I(L_{p+2}(t) = m-1) = 0$) are obtained simultaneously through a DSA-selected, pooled estimator over the three CD4 count levels, m.
- The initial estimates of $Q_{0,L_{p+2,m}(t)}$ for $t = 2, 3$ and $m = 1, 2, 3$ [at $\bar{Y}(t) = 0$, $\bar{A}_2(t-1) = 0$, nonmissing $\bar{A}_1(t-1)$, and $I(L_{p+2}(t) = 1) = 0, \ldots, I(L_{p+2}(t) = m-1) = 0$] are obtained simultaneously through a DSA-selected, pooled estimator over the three time intervals, t, and the three CD4 count levels, m.

We refer to the table in our technical report that lists the variables of the reduced data that were considered as candidate main terms in each of the four DSA estimators described above. With the exception of the variables enrollyear and m, no categorical variable with $l > 2$ levels was directly considered as a candidate main term. Instead, l binary variables, each of which represents the indicator that X is equal to m (denoted by $I(X = m)$) for $m = 1, \ldots, l$, were considered as potential main terms. The time-variable t was not only treated as a categorical variable but was also considered directly as a candidate main term. The initial estimates resulting from the application of the four DSA estimators described above are summarized in the accompanying technical report.

Calculate the optimal fluctuation. Estimation of the target parameter $\psi_0 = \Psi(Q_0)$ with the R-TMLE involves the fluctuation of the initial estimators of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ obtained previously. The optimal fluctuation of each of these initial estimators is based on the clever covariates, H^*, defined by equalities (26.12) and (26.13). R-TMLE implementation thus requires calculation of these clever covariates. They are functions of $Q_{0,L_{p+1,1}(t)}$ for $t = 1, 2, 3, 4$ and $Q_{0,L_{p+2,m}(t)}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$ but also the following components of the action mechanism defined in Sect. 26.2: $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ for $t = 0, 1, 2, 3$. Therefore, we first need to estimate $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ prior to computing the clever covariates. These estimates, combined with the initial estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$, can then be mapped into an estimate of the clever covariates using Monte Carlo simulations.

Obtain an estimate g_n of g_0. The following eight components of the action mechanism, i.e., the g part of the reduced, observed data likelihood [see equality (26.3) tailored to the reduced-data structure], are relevant for implementation of the R-TMLE of the target parameter ψ_0 and need to be estimated:

- $g_{A_1(t)} \equiv P(A_1(t) \mid \bar{A}(t-1), \bar{L}(t), A_2(t))$ for $t = 0, 1, 2, 3$ [only relevant at $\bar{Y}(t) = 0$, $\bar{A}_2(t) = 0$, and nonmissing $\bar{A}_1(t)$];
- $g_{A_2(t)} \equiv P(A_2(t) \mid \bar{A}(t-1), \bar{L}(t))$ for $t = 0, 1, 2, 3$ [only relevant at $\bar{Y}(t) = 0$, $\bar{A}_2(t-1) = 0$, and nonmissing $\bar{A}_1(t-1)$].

Estimation of the target parameter $\psi_0 = \Psi(Q_0)$ with the R-TMLE relies on estimation of the corresponding 8 g_0-components of the true reduced-data-generating distribution, P_0. The 8 estimates of $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ are denoted by $g_{A_1(t),n}$ and $g_{A_2(t),n}$, respectively. They are obtained based on sieve estimation with the same estimator selection procedure adopted for initial estimation of the 14 Q_0 components described in the previous section. The following four data stratification/pooling schemes were applied to derive each of the 8 estimates:

- The estimate of $g_{0,A_1(0)}$ [at $A_2(0) = 0$ and nonmissing $A_1(0)$] is obtained separately.
- The estimates of $g_{0,A_1(t)}$ for $t = 1, 2, 3$ [at $\bar{Y}(t) = 0$, $\bar{A}_2(t) = 0$, and nonmissing $\bar{A}_1(t)$] are obtained simultaneously through a DSA-selected, pooled estimator over the three time intervals, t.
- The estimate of $g_{0,A_2(0)}$ is obtained separately.
- The estimates of $g_{0,A_2(t)}$ for $t = 1, 2, 3$ [at $\bar{Y}(t) = 0$, $\bar{A}_2(t-1) = 0$, and nonmissing $\bar{A}_1(t-1)$] are obtained simultaneously through a DSA-selected, pooled estimator over the three time intervals, t.

We refer to the table in the technical report that lists the variables of the reduced data that were considered as candidate main terms in each of the four DSA estimators described above. With the exception of the variables enrollyear and m, no categorical variable with $l > 2$ levels was directly considered as a candidate main term. Instead, l binary variables, each of which represents the indicator that X is equal to m (denoted by $I(X = m)$) for $m = 1, \ldots, l$, were considered as candidate main terms.

The variable t was not only treated as a categorical variable but was also considered directly as a candidate main term. The estimates resulting from the application of the DSA estimators are presented in our technical report.

Monte Carlo simulation based on g_n^0 and Q_n^0. The 13 clever covariates that need to be computed are defined by equalities (26.12) and (26.13): $H^*_{L_{p+1,1}(t)}$ for $t = 1, 2, 3, 4$ and $H^*_{L_{p+2,m}(t)}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$. Each of these clever covariates are used to fluctuate the estimators of $Q_{0,L_{p+1,1}(t)}$ for $t = 1, 2, 3, 4$ and $Q_{0,L_{p+2,m}(t)}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$ respectively.

By extending the definition of the observed outcome, $Y(t) = I(\tilde{T} \leq t - 1, \Delta = 1)$, to time points t beyond the time of an event when it is observed, i.e., for $t > \tilde{T}$ when $\Delta = 1$, the clever covariate for updating the initial estimator of $Q_{0,L_{p+1,1}(t)}$ at $\bar{Y}(t-1) = 0, \bar{A}_2(t-1) = 0$, and nonmissing $\bar{A}_1(t-1)$ for $t = 1, 2, 3, 4$ can be rewritten

$$
\begin{aligned}
H^*_{L_{p+1,1}(t)} &= \frac{1}{\prod_{t'=0}^{t-1} g_{A_1(t')} g_{A_2(t')}} \qquad (26.14)\\
&\times \Big[I(\bar{A}(t-1) = d_{350}(\bar{L}(t-1)))\Big(1 - E_Q(Y_{d_{350}}(4) \mid \bar{L}(t-1), \bar{A}(t-1), Y(t) = 0)\Big)\\
&- I(\bar{A}(t-1) = d_{200}(\bar{L}(t-1)))\Big(1 - E_Q(Y_{d_{200}}(4) \mid \bar{L}(t-1), \bar{A}(t-1), Y(t) = 0)\Big)\Big].
\end{aligned}
$$

Note that at $t = 4$, equality (26.14) does indeed simplify to equality (26.13). Similarly, the clever covariate for updating the initial estimator of $Q_{0,L_{p+2,m}(t)}$ at $\bar{Y}(t) = 0$, $\bar{A}_2(t-1) = 0$, nonmissing $\bar{A}_1(t-1)$, and $L_{p+2,1}(t) = 0, \ldots, L_{p+2,m-1}(t) = 0$ for $t = 1, 2, 3$ and $m = 1, 2, 3$ can be rewritten as

$$
\begin{aligned}
H^*_{L_{p+2,m}(t)} &= \frac{1}{\prod_{t'=0}^{t-1} g_{A_1(t')} g_{A_2(t')}} \Big[I(\bar{A}(t-1) = d_{350}(\bar{L}(t-1)))\\
&\times \Big(E_Q(Y_{d_{350}}(4) \mid Pa(L_{p+2,m}(t)), L_{p+2,m}(t) = 1) - E_Q(Y_{d_{350}}(4) \mid Pa(L_{p+2,m}(t)), L_{p+2,m}(t) = 0)\Big)\\
&- I(\bar{A}(t-1) = d_{200}(\bar{L}(t-1)))\Big(E_Q(Y_{d_{200}}(4) \mid Pa(L_{p+2,m}(t)), L_{p+2,m}(t) = 1)\\
&\qquad - E_Q(Y_{d_{200}}(4) \mid Pa(L_{p+2,m}(t)), L_{p+2,m}(t) = 0)\Big)\Big],
\end{aligned}
$$

where $Pa(L_{p+2,m}(t)) \equiv (\bar{L}(t-1), \bar{A}(t-1), Y(t), L_{p+2,1}(t), \ldots, L_{p+2,m-1}(t))$.

The above formulation of the clever covariates makes explicit how they can be computed by first approximating each of the conditional expectations of $Y_{d_\theta}(4)$ by Monte Carlo simulations. Following the factorization of the reduced, observed data likelihood according to the time ordering of actions and covariates, 10,000 observations of the potential outcomes $Y_{d_\theta}(4)$ (for $\theta = 200, 350$) were simulated by sequentially generating future covariates starting at the fixed covariate and action history specified by the conditional event in each expectation and by setting future actions to the interventions implied by the individualized action rule d_θ. This simulation process ends when the outcome at time point 4 is simulated or earlier if the simulated event occurs before time point 4. The averages of these simulated potential outcomes approximate the desired conditional expectations of $Y_{d_\theta}(4)$. The value of the clever covariates needs to be calculated at each of the time points t, and for each

of the n subjects in the sample. These values could also be computed analytically, using the above analytical expressions instead of the reliance on simulations.

The TMLE. The optimal fluctuations of the initial estimators of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ based on the clever covariates computed previously, $H^*_{L_{p+1,1}(t)}$ and $H^*_{L_{p+2,m}(t)}$, result in the definition of updated, one-step estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ denoted by $Q^1_{L_{p+1,1}(t),n}$ and $Q^1_{L_{p+2,m}(t),n}$, respectively.

Specifically in this analysis, each updated estimate $Q^1_{L_{p+1,1}(t),n}$ for $t = 1, 2, 3, 4$ is defined by a separate, t-specific maximum likelihood regression of the outcome at time t, $L_{p+1,1}(t)$, on the clever covariate $H^*_{L_{p+1,1}(t)}$ based on a logistic model with offset equal to the logit transformation of the initial estimate $Q^0_{L_{p+1,1}(t),n}$ and based on the same observations at time t that contributed to the initial estimate of $Q_{0,L_{p+1,1}(t)}$, i.e., the updated estimate is defined by

$$Q^1_{L_{p+1,1}(t),n} = \frac{1}{1 + \exp\left(- (\text{logit}(Q^0_{L_{p+1,1}(t),n}) + \epsilon_n H^*_{L_{p+1,1}(t)})\right)},$$

where ϵ_n is the maximum likelihood estimate. Similarly, each updated estimate $Q^1_{L_{p+2,m}(t),n}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$ is defined by a separate, (t, m)-specific maximum likelihood regression of the indicator of CD4 count at time t equal to level m, $L_{p+2,m}(t)$, on the clever covariate $H^*_{L_{p+2,m}(t)}$ based on a logistic model with offset equal to the logit transformation of the initial estimate $Q^0_{L_{p+2,m}(t),n}$ and based on the same observations at time t that contributed to the initial estimate of $Q_{0,L_{p+2,m}(t)}$. The value of the coefficient in front of the clever covariate in each of the logistic models defining the one-step estimates above is given in Table 26.2.

Implementation of the (iterative) R-TMLE relies on iteration of the updating process above until a convergence criterion is met. Starting with $k = 1$, the clever covariates are first recalculated by Monte Carlo simulations based on the latest updated estimates $Q^k_{L_{p+1,1}(t),n}$ and $Q^k_{L_{p+2,m}(t),n}$. Second, these latest updated estimates are fluctuated with the newly computed clever covariates to define newly updated estimates $Q^{k+1}_{L_{p+1,1}(t),n}$ and $Q^{k+1}_{L_{p+2,m}(t),n}$ using the updating process above where the initial estimates $Q^0_{L_{p+1,1}(t),n}$ and $Q^0_{L_{p+2,m}(t),n}$ are replaced with the latest updated estimates $Q^k_{L_{p+1,1}(t),n}$ and

Table 26.2 Estimates ϵ_n of the coefficients in front of the clever covariates in the logistic models defining the one-step, updated estimates of $Q_{0,L_{p+1,1}(t)}$ for $t = 1, 2, 3, 4$ and $Q_{0,L_{p+2,m}(t)}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$

t	$L_{p+1,1}(t)$	$L_{p+2,1}(t)$	$L_{p+2,2}(t)$	$L_{p+2,3}(t)$
1	0.114	7.544	−24.694	34.977
2	−0.005	5.061	−20.949	29.879
3	0.031	1.886	−2.207	−1.892
4	0.013			

$Q^k_{L_{p+2,m}(t),n}$. Third, k is incremented by 1. The three-step process just described is repeated until a convergence criterion is met. The last updated estimates are referred to as the targeted estimates. A sensible convergence criterion is that the ϵ_n-values approximate zero, or that the empirical mean of the efficient influence curve at the current update approaches a value close enough to zero, taking into account the standard error of the estimator. In this analysis, only one step was carried out, so that the estimates $Q^1_{L_{p+1,1}(t),n}$ and $Q^1_{L_{p+2,m}(t),n}$ are deemed the targeted estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ denoted by $Q^*_{0,L_{p+1,1}(t),n}$ and $Q^*_{0,L_{p+2,m}(t),n}$, respectively. Practical evidence has suggested that most bias reduction occurs in the first step (Chap. 18). We also refer readers to the forthcoming Stitelman and van der Laan (2011a).

A substitution estimator of the parameter of interest. The R-TMLE estimate of the target parameter $\psi_0 = \Psi(Q_0)$ defined by equality (26.7) is derived by substitution of the relevant distributions Q_0 in the right hand-side of equality (26.7), i.e. $Q_{0,L(0)}$, $Q_{0,L_{p+1,1}(t)}$ for $t = 1, 2, 3, 4$, and $Q_{0,L_{p+2,m}(t)}$ for $t = 1, 2, 3$ and $m = 1, 2, 3$, with the empirical distribution $Q_{L(0),n}$, and the targeted estimates $Q^*_{L_{p+1,1}(t),n}$ and $Q^*_{L_{p+2,m}(t),n}$, respectively.

Concretely, this substitution estimate can be calculated using the following two-step procedure. First, the conditional expectations $E(Y_{d_\theta}(4) \mid L(0))$ (for $\theta = 200, 350$ and each unique observation of $L(0)$) are approximated by Monte Carlo simulation based on the targeted estimates $Q^*_{L_{p+1,1}(t),n}$ and $Q^*_{L_{p+2,m}(t),n}$ using the general simulation protocol described previously for the computation of the clever covariates. The resulting estimates of $E(Y_{d_\theta}(4) \mid L(0))$ are denoted by $E_{Q^*_n}(Y_{d_\theta}(4) \mid L(0))$. Second, these estimates are mapped into the R-TMLE estimate of ψ_0 denoted by $\Psi(Q^*_{0,n})$ using the formula

$$\Psi(Q^*_{0,n}) = \frac{1}{n}\sum_{i=1}^{n} E_{Q^*_{0,n}}(Y_{d_{350}}(4) \mid L(0) = l_i(0)) - \frac{1}{n}\sum_{i=1}^{n} E_{Q^*_{0,n}}(Y_{d_{200}}(4) \mid L(0) = l_i(0)).$$

This resulted in an R-TMLE estimate $\psi^*_n = 1.98\text{e}{-03}$ of ψ_0.

Influence curve based inference. Under regularity conditions, and under the assumption that g_n is a consistent estimator of action mechanism g_0, the R-TMLE estimator is asymptotically linear with influence curve that has a variance that is smaller than or equal to the variance (under P_0) of $D^*(Q^*, g_0, \psi_0)(O)$, where Q^* denotes the possibly misspecified limit of Q^*_n. A consistent estimator of the variance of the R-TMLE is thus obtained as follows:

$$Var_n(\Psi(Q^*_n)) = \frac{1}{n^2}\sum_{i=1}^{n}\{D^*_n(o_i) - \bar{D}^*_n\}^2, \tag{26.15}$$

where D^*_n is the estimated efficient influence curve, $\bar{D}^*_n$ is its empirical mean [which would equal zero if we plugged in the fully iterated TMLE (Q^*_n, g_n)]. The projection $\Pi(D_{IPAW} \mid L(0))$-term in the influence curve was estimated using the targeted estimates $Q^*_{L_{p+1,1}(t),n}$ and $Q^*_{L_{p+2,m}(t),n}$, while for the other (clever covariate) terms, we

used the initial estimator Q_n^0. This simplification permits straightforward calculation of the influence curve evaluated at each observation i using the intermediate results from the previous computing steps, i.e., the clever covariate calculations based on the initial estimate of Q_0, and the Monte Carlo simulations based on the targeted estimate of Q_0 to derive the substitution estimator, without the need for additional computation. Based on this approach, the estimate of the standard error associated with the R-TMLE is $\sigma_{0,n} = 2.71\text{e}{-03}$ resulting in the following 95% confidence interval for ψ_0: [−3.32e−03, 7.28e−03].

Diagnosing sparse data bias. To mitigate the higher variability of the R-TMLE resulting from practical violation of the ETA assumption, truncation of the IPA weights can be used as part of the R-TMLE implementation to improve the mean squared error associated with R-TMLE estimation of the target parameter. Based on the distributions of the IPA weights in this analysis, as presented in the accompanying technical report, we used a truncation level of 20 for the implementation of the R-TMLE, i.e., the clever covariates on which implementation of the R-TMLE is based were computed based on IPA weights that were set to 20 if their values implied by the estimates $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ were greater than 20. The point estimate and estimate of the standard error associated with the R-TMLE based on truncated IPA weights is $\Psi(Q_n^*) = 1.44\text{e}{-03}$ and $\sigma_n = 2.68\text{e}{-03}$ respectively which results in the following 95% confidence interval for ψ_0: [−3.81e−03,6.69e−03].

26.4.2 IPAW-R-TMLE Implementation and Results

To account for potential time-dependent confounding that was ignored by the R-TMLE as a consequence of the data-reduction step, the IPAW-R-TMLE relies on an estimate of the true action mechanism, i.e., the components of the action mechanism, denoted by $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$. These conditional distributions should not be confused with the components of the reduced-data action mechanism that were estimated in the previous section as part of the R-TMLE implementation and that we now denote by $g^r_{0,A_1(t)}$ and $g^r_{0,A_2(t)}$. The approach based on the DSA algorithm to derive the estimates of $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ is described in the accompanying technical report.

The initial estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$ of the R-TMLE were recomputed, but now using weights equal to the plug-in estimator $w_n(t)$ of

$$w_0(t) \equiv \frac{\prod_{j=0}^{t-1} g^r_{0,A_1(t)} g^r_{0,A_2(t)}}{\prod_{j=0}^{t-1} g_{0,A_1(t)} g_{o,A_2(t)}}.$$

Similarly, the R-TMLE of the ϵ-coefficients was recomputed by now using these time-dependent weights $w_n(t)$. The resulting substitution estimate of the target parameter ψ_0 is equal to 1.41e-03 and corresponds to the one-step IPAW-R-TMLE point estimate. Note that the only difference in implementation of the R-TMLE point

estimate vs. that of the IPAW-R-TMLE point estimate is in the use of the weights $w_n(t)$ to obtain the initial and updated estimates of $Q_{0,L_{p+1,1}(t)}$ and $Q_{0,L_{p+2,m}(t)}$. As explained in Appendix A, the weighted log-likelihood for the reduced data actually represents a valid loss function for the conditional counterfactual distributions of the reduced-data components $L_a^r(t)$ of the counterfactual $L_a(t)$, if g_0 is consistently estimated (i.e., w_n is consistent for w_0), or if the SRA holds with respect to the reduced data (i.e., there is no time-dependent confounding beyond the time-dependent covariates included in the reduced-data structure).

Inference with the IPAW-R-TMLE can be derived based on its influence curve (26.11) evaluated at the estimator of the action mechanism g_0 and the targeted estimator of Q_0 defined by the procedure above. Note that evaluation of formula (26.11) to derive inference with the IPAW-R-TMLE involves the estimates of $g_{0,A_1(t)}$ and $g_{0,A_2(t)}$ instead of the estimates of $g^r_{0,A_1(t)}$ and $g^r_{0,A_2(t)}$, i.e., the clever covariates used for fluctuation of the initial estimators of Q_0 in the implementation of the IPAW-R-TMLE should be multiplied by $w(t)$ to derive the clever covariates that appear in (26.11). Based on this approach and the same implementation shorcut employed earlier to simplify calculation of the R-TMLE influence curve evaluated at each observation i, the estimate of the standard error associated with the IPAW-R-TMLE is $\sigma_{0,n} = 2.48\text{e}{-03}$ resulting in the following 95% confidence interval for ψ_0: [−3.45e−03, 6.27e−03].

A few observations are characterized by relatively large IPA weights (> 20) which suggests some practical violation of the ETA assumption. To mitigate the higher variability of the IPAW-R-TMLE resulting from practical violation of the ETA assumption, the IPAW-R-TMLE was implemented as described above with the difference that the clever covariates were computed based on reduced-data IPA weights, $(\prod_{t=0}^{\check{T}} g^r_{A_1(t)} g^r_{A_2(t)})^{-1}$, that were truncated at 20. The point estimate and estimate of the standard error associated with this truncated IPAW-R-TMLE is $\Psi(Q^*_{0,n}) = 7.6\text{e}{-04}$ and $\sigma_{0,n} = 2.46\text{e}{-03}$ respectively which results in the following 95% confidence interval for ψ_0: [−4.07e−03, 5.59e−03].

Table 26.3 summarizes the results from the application of each estimator of the target parameter ψ_0 implemented in this analysis. Note that all inferences are consistent. The null hypothesis of a null effect ψ_0 may not be rejected based on the data

Table 26.3 Comparison of the results for each estimator of the target parameter ψ_0

Estimator	Estimate	SE	95% CI	p-value
R-IPAW (based on g^r)	1.42e−03	2.69e−03	[−3.86e−03, 6.69e−03]	0.60
R-TMLE	1.98e−03	2.71e−03	[−3.32e−03, 7.28e−03]	0.46
Truncated R-TMLE	1.44e−03	2.68e−03	[−3.81e−03, 6.69e−03]	0.59
IPAW (based on g)	6.9e−04	2.48e−03	[−4.16e−03, 5.54e−03]	0.78
IPAW-R-TMLE	1.41e−03	2.48e−03	[−3.45e−03, 6.27e−03]	0.57
Truncated IPAW-R-TMLE	7.6e−04	2.46e−03	[−4.07e−03, 5.59e−03]	0.76

from the KPNC electronic medical record and the three assumptions the IPAW-R-TMLE relies upon: SRA, positivity assumption, and consistent estimation of g_0.

The 95% confidence intervals suggest that the absolute value of the true causal risk difference is less than 1%. If the three assumptions on which these estimators rely for drawing a valid causal inference indeed hold, the null result may reflect a true null effect, a bias due to the erroneous inclusion of patients with a prevalent ADC at study entry, or a lack of power to detect a relatively small causal risk difference. In the accompanying technical report we comment on these three possible explanations before discussing possible reasons for potential bias due to violation of one or more of the three assumptions the IPAW-R-TMLE relies upon.

26.5 Discussion

In order to make progress in individualized medicine and comparative effectiveness research, one will need to understand outcome distributions under dynamic treatments. This chapter represents one very important application of dynamic treatments in comparative effectiveness research to inform the decision of when to start treatment in HIV patients.

The development of robust and efficient estimation methods that allow the data analyst to target clinically relevant causal quantities, and corresponding user friendly software implementations, will allow these methods to become prominent tools for analyzing longitudinal data. Furthermore, the roadmap for causal inference allows honest and careful interpretation of the results.

In response to the need to compare dynamic treatments, sequentially randomized controlled trials are also becoming more popular and provide a way to consistently estimate the causal effect of dynamic treatments such as dynamic treatments indexed by a choice of first line therapy, a cutoff for an intermediate biomarker, and a second line therapy to be assigned if the biomarker exceeds the cut-off (Thall et al. 2007; Bembom and van der Laan 2007b).

Possible important extensions of the analysis carried out in this chapter are to target dose-response curves defined as the survival curves under the when-to-start rule d_θ that starts HIV-treatment when the CD4-count drops below θ for a range of θ.By posing a working model for this dose-response curve in θ, one will be able to obtain more precise estimators of the projection of the true dose response curve on the working model, since all individuals will now contribute to the fit of this working model. Such an approach still allows for a valid test of a null hypothesis of interest about a certain contrast of this class of treatment rules as long as the working model is valid under the null hypothesis of interest. We refer to van der Laan (2010b) for details on formulation and the TMLE of the unknown parameters defined by this working model.