Treatment-competing events in dynamic regimes

Abstract A dynamic treatment regime is a sequence of decision rules for assigning treatment based on a patient's current need for treatment. Dynamic regimes are viewed, by many, as a natural way of treating patients with chronic diseases; that is, treating patients with adaptive, complex, longitudinal treatment regimens. In developing dynamic treatment strategies, treatment-competing events may play an important role in the overall treatment strategy, and their effects on subsequent treatment decisions and eventual outcome should be considered. Treatment-competing events may be defined generally as patient-specific, random events which interrupt the ongoing treatment decision process in a dynamic regime. Treatment-competing events censor later treatment decisions that would otherwise be made on a particular dynamic treatment regime had the competing events not occurred. For example, in therapeutic studies of HIV, physicians may assign treatment based on a patient's current level HIV1-RNA; this defines a treatment assignment rule. However, the presence of opportunistic infections or severe adverse events may preclude a strict adherence of the treatment assignment rule. In other contexts, the "censoring"-by-death phenomenon may be viewed as an example of a treatment-competing event for a particular dynamic treatment regime. Treatment-competing events can be built into the dynamic treatment regime framework and counting processes are a natural mechanism to facilitate this development. In this paper, we develop treatment-competing events in a dynamic infusion policy, a random dynamic treatment regime where multiple infusion treatments are initiated simultaneously and given continuously over time subject to the presence/absence of a treatment-competing event. We illustrate how our methodology may be used to suggest an estimator for a particular causal estimand of recent interest. Finally, we exemplify our methods in a recent study of patients undergoing coronary stent implantation.

Keywords: Adaptive treatment strategies · "Censoring"-by-death · Counting processes · Infusion trial · Observational data.

1 Introduction

Dynamic treatment regimes are adaptive, subject-specific treatment schedules with well-defined rules for treatment assignment (Murphy, van der Laan, and Robins, 2001). The rules for treatment assignment in a dynamic treatment regime may be stochastic or nonstochastic but are always specified before treatment begins. In *dynamic infusion policies*, physicians use intermediate, time-dependent measures to determine infusion length. An important component of a dynamic infusion policy includes a requirement that patients who experience particular adverse events during the course of the study immediately discontinue infusion. Then, the course of infusion treatment is described as follows: if a patient has not already experienced such an adverse event, infusion continues until the attending physician discontinues the infusion by choice or the adverse event occurs. The attending physician's decision to "continue or stop infusing" at time t is the treatment decision of interest; a competing event which "stops infusion by necessity" is an intermediate event which censors the treatment decision. In a dynamic regime, these treatment-competing events are an important part of the overall treatment strategy and not regarded as nuisance events. Moreover, they are an integral part of the parameter of interest. Murphy et al. (MLR hereafter) develop a general framework for longitudinal studies where treatment assignment occurs at a finite number of fixed clinic visits. A major goal of this paper is to generalize concepts in MLR for estimating parameters in dynamic infusion policies or, more generally, experiments where multivariate treatment regimens run continuously in time.

Following major surgeries, patients are often infused with multiple drugs running "in parallel." For patients undergoing coronary stent implantation, those medications may include multiple blood thinners, to prevent coagulation near the stent, and pain medications. During the course of an infusion trial, any number of adverse events may require a discontinuation of some or all the infusion

treatments. A heart attack may require an immediate termination of all infusions so the patient may receive appropriate medical attention, including a second surgery. However, bleeding may signal the blood is too thin and, hence, a blood thinner (e.g. Integrilin) infusion may be stopped immediately while other infusion treatments allowed to continue. The infusions cannot continue indefinitely; they must be terminated before some protocol-defined upper limit, say τ. At some time point beyond τ, an endpoint is observed. The endpoint could be observed either before the patient leaves the hospital or at a follow-up visit. One goal of this paper is to illustrate how one can estimate the mean endpoint had all physicians followed a dynamic infusion policy using data from an observational study. The data generating process in the observational study differs from that in the dynamic infusion policy by the mechanism which assigns infusion treatment to patients for the next interval of time. In a dynamic infusion policy, infusion treatment is stopped or continued at time t based on intermediate measures observed just prior to time t. In routinely collected data and observational studies, however, physicians use their best judgement and all available patient information — potentially including auxiliary patient data, medical history, treatment history — in deciding when to stop the infusion treatment. MLR provide a method of relating estimators in the observational data to estimators in their dynamic treatment regime. The details of this approach in the context of our dynamic infusion policy are given in Sect. 3.

In describing a physician's decision to stop or continue infusing at time t in a dynamic infusion policy, we find that counting processes lend themselves naturally. In our setup, the individual counting process takes a jump of size +1 at a time when the attending physician stops the infusion by choice. Along these lines, other authors have used counting processes in causal inference or the analysis of observational data. In an unpublished PhD thesis, Lok (2001) used multivariate counting processes to facilitate her analysis of structural nested failure time models (Lok, Gill, van der Vaart, and Robins, 2004). The structural approach taken in these papers are fundamentally different from

the marginal approach formulated below. More importantly, our use of counting processes below applies to arbitrary endpoints, not only to potentially truncated or censored endpoints. Recently, Johnson and Tsiatis (JT; 2004, 2005) used counting processes to argue consistency and asymptotic normality for estimators of mean response modeled as functions of treatment duration policies, which may considered a special case of our dynamic infusion policy. Our use of counting processes is most similar to that presented in Johnson and Tsiatis (2005, Appendix), whose estimation problem was also motivated by infusion trial applications.

The framework described in this paper applies to a general class of problems where multiple treatments are given in parallel, continuously over time. In Sect. 2, we introduce data from an infusion trial where a single infusion treatment is administered. We develop a general framework in Sect. 3 in anticipation of data from several substantive efforts between the authors and collaborators (two of which are described in Sect. 6). Of course, the general framework also applies to the univariate infusion treatment problem. However, such generality is superfluous if univariate infusion treatment is the only motivation. Our goal will be to present the general and simplified framework jointly but only illustrate the methods in the simple setting. This is a direct consequence of the available data at the time of this writing. In Sect. 4, we use our framework to propose a local linear estimator for the same causal estimand in Johnson and Tsiatis (2004; 2005). Our estimator is easy to implement and will consistently estimate the causal estimand under weaker conditions than those presented in Johnson and Tsiatis. We conclude with an analysis of data collected in an infusion trial conducted at Duke University Medical Center in Sect. 5.

2 Motivating dynamic regime

The methods discussed in this paper are motivated, in part, by a scientific problem from the ESPRIT (Enhanced Suppression of the Platelet Receptor IIb/IIIa with Integrilin Therapy) infusion trial conducted at Duke University Medical Center. The main objective of ESPRIT was to compare eptifibatide (Integrilin) therapy to placebo on the basis of the composite binary endpoint Y of death, myocardial infarction (MI), or urgent target vessel revascularization within 30 days treatment initiation. We note that patients were initially randomized to receive either the experimental treatment regimen or placebo regimen. The ESPRIT investigators found that the experimental Integrilin infusion regimen reduced significantly the proportion of 30-day primary endpoints when compared to the placebo regimen. Then, investigators became keenly interested in their secondary endpoint, that was, to understand how the average 30-day endpoint changes as function of Integrilin infusion length. The answer to this scientific question is complex because physicians treated their patients on a case-by-case basis according to patient medical histories, medical experience and expertise. In addition to well-known problems of potential confounding in observational data, we note that an infusion-censoring event, if observed, may be significantly related to a patient's response. In fact, in the ESPRIT trial, the events which define the composite 30-day endpoint may also be one of the adverse events that would necessitate infusion termination (e.g. death and MI).

Definition 1 (Infusion-length policy for ℓ units of time) *Infuse for ℓ units of time or until an infusion-competing event occurs, whichever comes first.*

Johnson and Tsiatis (JT; 2004, 2005) argued that one could address the goals of the ESPRIT investigators through "infusion-length policies" defined above. In particular, they argued that the inferential goal may be formalized as estimating the mean response, i.e. 30-day composite

endpoint, had the entire study population followed an infusion-length policy for ℓ units of time. An "infusion-length policy for ℓ units of time" is a special case of our dynamic infusion policy introduced in Sect. 1, which is in turn a special case of MLRs dynamic treatment regime. We defer a detailed discussion of dynamic infusion policies to Sect. 3 and focus on the infusion-length policy here. Such a policy may be viewed as a *nonrandom* dynamic treatment regime, that is, a random dynamic treatment regime with degenerate treatment selection probabilities. As mentioned in Sect. 1, a dynamic treatment regime is a sequence of decision rules over a time period where treatment assignments and decisions are admissible. In order to apply results from the dynamic treatment regime framework proposed by MLR to JTs infusion-length policy, a formal decision rule is required. To this end, we must introduce some preliminary notation.

We let X denote the observed infusion length and Δ be the indicator where the event $\{\Delta = 0\}$ indicates a patient experienced an infusion-competing event. Define $S(t)$ as an intermediate response variable; here, $S(t)$ is scalar and indicates whether an infusion-competing event has occurred by time t; in particular, define $S(t) = I(X \leq t, \Delta = 0)$. Given the intermediate response just prior to time, i.e. $S(t^-)$, nonrandom dynamic treatment regimes assign treatment at time t deterministically. In this example, our treatment assignment (decision) rule, $r\{t, S(t^-)\}$, tells the physician whether to stop or continue infusing based on whether the patient has experienced an infusion-terminating event prior to time t. Note, that the decision rule must be properly defined for all time t in the admissible range $[0, \tau]$. The decision rule (Johnson and Tsiatis, 2004, Appendix) is given by

$$r\{t, S(t^-)\} \quad = \quad \begin{cases} 0 & \text{if } t < \ell \text{ and } S(t^-) = 0 \\ 1 & \text{otherwise.} \end{cases}, \quad \text{for } t \leq \tau \tag{1}$$

This nonrandom dynamic treatment regime (hence, the infusion-length policy) is denoted by $\bar{r}_\tau$ and represents the entire sequence of decision rules $r\{t, S(t^-)\},\ 0 \leq t \leq \tau$.

From our description of the ESPRIT infusion trial above, it is evident that physicians were neither forced nor encouraged to follow infusion-length policies, as defined through the decision rules $r\{t, S(t^-)\},\ 0 \leq t \leq \tau$. In ESPRIT (as in other infusion trials), physicians naturally continued to infuse patients until they deemed it appropriate to stop, given a patient did not experience an infusion-terminating event at some earlier time. It is not uncommon, for example, that a physician would consult a patient's medical chart when deciding whether to stop or continue the infusion process. However, according to the dynamic treatment regime $\bar{r}_\tau$ defined through the rules in (1), additional patient-specific data including medical charts are considered auxiliary information. Hence, auxiliary variables, denoted by $\{\mathbf{Z}(u),\ u \leq X\}$, affect directly a physician's decision to stop or continue infusion in the observed data but not in the infusion-length policy. In both the infusion-length policy and the observed data, auxiliary variables may affect the infusion-competing events and, consequently, the endpoint. In short, auxiliary variables are an important component of the estimand for both the infusion-length policy and observational data although they affect treatment decisions only in the observational study. In ESPRIT, auxiliary variables included baseline measures such as diabetes, heparin use, angina and time-dependent measures such as enzyme-level at time $t,\ t \leq X$. In Sect. 3, we discuss under what assumptions one can estimate the mean 30-day endpoint if the population had followed the target infusion-length policy for ℓ units of time.

Note: As with many developments in the dynamic treatment regime genre, there are two studies considered: a sequentially randomized study, i.e. the dynamic treatment regime, and the observational study. In this paper, we will always use dynamic treatment regime to refer to MLRs general concept and "policy" to refer to our specialized application of dynamic treatment regime with infusion treatment(s). In particular, we use dynamic infusion policy for the general concept in Sect. 3, and infusion-length policy for the special case. When referring to data collected from the observational study, we will say observed or observational and avoid terms such as "observed

infusion policy" and "observed infusion-length policy." Finally, we refer to an infusion-length policy for ℓ units of time as the *target* infusion-length policy when it is the particular infusion-length policy under investigation (Sects. 4–5).

The ESPRIT infusion-length problem is considered a special case of the general dynamic infusion policy methodology in several ways. First, the only intermediate outcome in ESPRIT is the binary indicator "infusion-competing event prior to time t." For general dynamic infusion policies, a physician's decision to stop or continue may be a function of an objective, time-dependent measure such as blood flow or score on a nurse-administered exam. Second, the ESPRIT treatment regimen is univariate whereas the general framework accommodates multiple infusion treatments running in parallel. Third, the infusion-length policy is an example of a random dynamic regime with degenerate selection probabilities. The general setup will consider non-trivial treatment selection probabilities.

3 General setup

3.1 Notation

We let X_k be the observed infusion length on treatment k and all observed infusion lengths $\mathbf{X} = (X_k,\ k = 1, \ldots, K)$. In defining our "failure-indicator" Δ_k, it is easier to describe its complement; we let $\Delta_k = 0$ if infusion on the k-th treatment was stopped due to a treatment-competing event. We define the counting process $N_k(t) = I(X_k \leq t, \Delta_k = 1)$, where $I(B) = 1$ if the event B occurs, and the multivariate counting process $\mathbf{N}(t) = \{N_k(t),\ k = 1, \ldots, K\}$. In the dynamic infusion policy, a physician's decision to stop or continue infusing treatment k at time t depends intermediate measures of a patient's need for continued treatment k; define such intermediate measures as $\mathbf{S}_k(t)$ (also called tailoring variables). A minimum requirement of $\mathbf{S}_k(t)$ is that it *contains the indicator*

of whether a treatment-competing event for treatment k has occurred up to and including time t, e.g. $\mathbf{S}_k(t) = \{I(X_k \leq t, \Delta_k = 0), \ldots\}$. We define the filtration in the dynamic infusion policy as

$$\mathcal{F}_D(t) = \sigma\{N_k(u), \mathbf{S}_k(u),\ u \leq t,\ k = 1, \ldots, K\}.$$

Assume $\mathbf{S}_k(t)$ is right-continuous and let $\mathbf{S}(t) = \{\mathbf{S}_k(t),\ k = 1, \ldots, K\}$, we define the following cause-specific hazard function for treatment k in the dynamic infusion policy

$$\lambda_{D,k}\{t, \mathbf{S}(t^-)\} = \lim_{\epsilon \to 0} \epsilon^{-1} P_D[t \leq X_k < t + \epsilon, \Delta_k = 1 | X_k \geq t, \mathbf{S}(t^-)],$$

where $P_D(\cdot)$ is the probability distribution in the dynamic infusion policy. In general, we will always use the subscript "D" when we refer to concepts, probabilities, distribution functions, etc. in the dynamic infusion policy and the corresponding concepts without "D" refer to the observed data. The integrated hazard function and its related integrated intensity process in the dynamic infusion policy are:

$$\begin{aligned} \Lambda_{D,k}\{t, \mathbf{S}(t^-)\} &= \int_0^t \lambda_{D,k}\{u, \mathbf{S}(u^-)\}\, d\mu(u) \\ A_{D,k}\{t, \mathbf{S}(t^-)\} &= \int_0^t I(X_k \geq u)\, d\Lambda_{D,k}\{u, \mathbf{S}(u^-)\}, \end{aligned}$$

where μ is the dominating measure of policy-conforming treatment decisions $(X_k, \Delta = 1)$. We prefer to use a single definition of integrated hazard through general measure μ to allow for the possibility that policy-conforming treatment decisions occur at finite number of time points (as in Johnson and Tsiatis, 2004), in continuous time (as in Johnson and Tsiatis, 2005), or a mixture of continuous and discrete time points. The probability that a physician decides to discontinue treatment k in the interval $[t, t + dt)$ is, informally,

$$P_D[dN_k(t) = 1 | \mathcal{F}_D(t^-)] = dA_{D,k}\{t, S_k(t^-)\}. \tag{2}$$

Because the rule for treatment assignment in (2) is stochastic, our dynamic infusion policy is what MLR refer to as a random dynamic treatment regime. In the dynamic infusion policy, we define

the observed, individual counting process $N_k^t = \{N_k(u),\ 0 \leq u \leq t\}$. The counting process N_k^τ takes a jump of size $+1$ at time t when a physician stopped the infusion on treatment k given that a patient has been infused on treatment k continuously up to time t. In our setup, the multivariate counting process $\mathbf{N}^t = \{N_k^t,\ k = 1, \ldots, K\}$ describes the *history of treatment assignment* up to and including time t. The sample paths, $\mathbf{n}^t = \{n_k(u),\ 0 \leq u \leq t,\ k = 1, \ldots, K\}$, are right-continuous and referred to as the *history of treatment* in the literature. Here, we note that the entire history of treatment assignment $\mathbf{N}^\tau$ contains less information than the collection of *jump times* and *jump marks*, i.e. $\{(X_k, \Delta_k),\ k = 1, \ldots, K\}$.

After τ units of time, we observe the entire treatment path $\mathbf{n}^\tau$. The collection of all possible treatment paths is defined as $\mathcal{N}$. We conceptualize the potential endpoint $Y(\mathbf{n}^\tau)$ (Rubin, 1974) as the endpoint we would observe, possibly contrary to fact, if a physician's sequence of treatment decisions is consistent with the path $\mathbf{n}^\tau$. The collection of all potential responses is $\{Y(\mathbf{n}^\tau),\ \mathbf{n}^\tau \in \mathcal{N}\}$. Hence, we assume that one patient's treatment does not affect another patient's treatment or potential endpoint; this assumption is often called the stable unit treatment value assumption (SUTVA; Rubin, 1986).

In a dynamic infusion policy, time-dependent auxiliary variables $\mathbf{Z}(t)$ are also collected in the dynamic infusion policy. At time t, the event $\{\mathbf{S}(t), \mathbf{Z}(t)\}$ is an intermediate response of treatment received prior to time t. Thus, SUTVA implies that for every treatment path $\mathbf{n}^\tau \in \mathcal{N}$, the intermediate outcome at time t may be written as the potential event $\{\mathbf{S}(t, \mathbf{n}^t), \mathbf{Z}(t, \mathbf{n}^t)\}$. Assuming SUTVA, the observed data at time t are $\{\mathbf{S}(t, \mathbf{N}^t), \mathbf{Z}(t, \mathbf{N}^t)\}$. As in MLR, we write the observed data at time t as $\mathbf{S}(t) = \mathbf{S}(t, \mathbf{N}^t)$ and $\mathbf{Z}(t) = \mathbf{Z}(t, \mathbf{N}^t)$ and the observed endpoint as Y. A physician's decisions to stop or continue infusion treatments at time t are denoted by $\Delta\mathbf{N}(t) = \mathbf{N}(t) - \mathbf{N}(t^-)$. For a randomly selected patient, the observed data are

$$\mathcal{D} = \{\mathbf{Z}(0), \ldots, \Delta\mathbf{N}(t), \mathbf{S}(t), \mathbf{Z}(t), \ldots, \Delta\mathbf{N}(\tau^-), \mathbf{S}(\tau^-), \mathbf{Z}(\tau^-), \Delta\mathbf{N}(\tau), Y\}.$$

We note that the observed data $\mathcal{D}$ are the same in both the dynamic infusion policy as well as in the observational study. In fact, the marginal distribution of the data $\mathcal{D}$ is also the same in both the dynamic infusion policy and the observational study. The main difference between the two experiments is the manner in which physicians assign infusion treatments. In the observational study, the probability that a physician decides to discontinue treatment k in the interval $[t, t + dt)$ will not be equal to $dA_{D,k}\{t, S_k(t^-)\}$ as is in the dynamic infusion policy via (2).

3.2 Identifying the causal estimand from the observed data

Our goal is to estimate the mean outcome had, contrary to fact, the entire study population followed the dynamic infusion policy. In the observational study, we do not observe a random sample of individuals from the study population following the dynamic infusion policy. Recall, in the observational study, the probability that a physician stops treatment at time t given the patient has been continuously treated up to time t depends on the need for continued treatment, and the histories of past treatment and auxiliary covariables. Therefore, we impose the following assumption to identify causal effects and causal parameters.

Assumption. *No unmeasured confounders*
At time t, the treatment decision $\Delta\mathbf{N}(t)$ is independent of potential intermediate events $\{\mathbf{S}(u, \mathbf{n}^u),\ u > t\}$ and potential responses $\{Y(\mathbf{n}^\tau),\ \mathbf{n}^\tau \in \mathcal{N}\}$ given the observed history $\{\mathbf{N}(u), \mathbf{S}(u), \mathbf{Z}(u),\ u < t\}$.

The no unmeasured confounders assumption implies that the potential random variables $\mathbf{S}(t)$ — including potential treatment censoring — and potential responses do not vary systematically within levels of the observed history. Such an assumption is plausible if information about an individual through time t, which may be prognostic and which an investigator may use to make decisions on treatment duration, are captured in the auxiliary data $\mathbf{Z}(u),\ u \leq t$ (cf. Robins, 1986).

Now, we discuss how we may use MLRs Lemma 4.1 to identify $E_D(Y)$, where $E_D(\cdot)$ and $P_D(\cdot)$ denotes expectation and probability, respectively, in the dynamic infusion policy. Define the filtration

$$\mathcal{F}(t) = \sigma\{\mathbf{N}(u), \mathbf{S}(u), \mathbf{Z}(u),\ u < t\}.$$

In the dynamic infusion policy, the probability that a physician discontinues treatment k in the interval $[t, t + dt)$ is

$$P_D[dN_k(t) = 1|\mathcal{F}(t^-)] = P_D[dN_k(t) = 1|\mathcal{F}_D(t^-)] = (2).$$

This follows because $\mathcal{F}_D(t) \subseteq \mathcal{F}(t)$ which implies that any stochastic process $W(t)$ which is $\mathcal{F}_D(t)$-measurable is also $\mathcal{F}(t)$-measurable. In the dynamic infusion policy, the auxiliary information $\{\mathbf{Z}(u),\ u < t\}$ is not used in assigning treatment at time t. In the observational study, however,

$$P[dN_k(t) = 1|\mathcal{F}(t^-)] = dA_k\{t, \mathcal{L}(t^-)\} = I(X_k \geq t)d\Lambda_k\{t, \mathcal{L}(t^-)\}, \tag{3}$$

where $\Lambda_k\{t, \mathcal{L}(t^-)\} = \int_0^t d\Lambda_k\{u, \mathcal{L}(u^-)\}$, $\mathcal{L}(t^-) = \{\mathbf{S}(u), \mathbf{Z}(u),\ u < t\}$ and $E(\cdot)$ and $P(\cdot)$ denote expectation and probability, respectively, in the observational study. Define

$$\eta\{\mathbf{N}^\tau, \mathcal{L}(\tau^-)\} = \mathcal{P}_0^\tau \left[\prod_{k=1}^{K} \left\{ \frac{dA_{D,k}\{t, \mathbf{S}(t^-)\}}{dA_k\{t, \mathcal{L}(t^-)\}} \right\}^{dN_k(t)} \left\{ \frac{1 - dA_{D,k}\{t, S(t^-)\}}{1 - dA_k\{t, \mathcal{L}(t^-)\}} \right\}^{1-dN_k(t)} \right], \tag{4}$$

where $\mathcal{P}_0^\tau$ denotes the product integral (cf. Kalbfleisch and Prentice, 2002, p. 10; Andersen et al. § II.6). Under regularity conditions, including $\mathbf{S}(t)$ is *cadlag* such that $\mathcal{F}(t)$ is right-continuous, the no unmeasured confounders assumption and

$$P\left[d\Lambda_k\{t, \mathcal{L}(t^-)\} > 0 \text{ for all } t \text{ such that } d\Lambda_{D,k}\{t, S(t^-)\} > 0,\ k = 1, \ldots, K\right] = 1, \tag{5}$$

the distribution of $[Y, \mathcal{S}(\tau^-), \mathbf{N}^\tau]$ under P_D is absolutely continous with respect to the distribution of $[Y, \mathcal{S}(\tau^-), \mathbf{N}^\tau]$ under P and a version of the Radon-Nikodym derivative is

$$E\left[\eta\{\mathbf{N}^\tau, \mathcal{L}(\tau^-)\}|Y = y, \mathcal{S}(\tau^-) = s(\tau^-), \mathbf{N}^\tau = \mathbf{n}^\tau, \right]. \tag{6}$$

The Radon-Nikodym derivative in (6) is a natural extension of Lemma 4.1 from MLR to the situation where multiple treatments are observed continuously over time. The assumption (5) implies that treatment policies occurring in the dynamic infusion policy must also have positive probability of occurring in the observational study.

To write the causal estimand $E_D(Y)$, we assume, without loss of generality, that $P[\mathcal{F}(0)] = 1$. Also, to describe the probabilistic mechanisms over the interval $[0, \tau]$, we use the intuitive result (cf. Kalbfleisch and Prentice, 2002, § 6.3) that

$$P[\mathcal{F}(t^- + dt)|\mathcal{F}(t^-)] = P[\mathbf{N}(dt)|\mathcal{F}(t^-)]P[\mathbf{S}(dt), \mathbf{Z}(dt)|\mathbf{N}(dt), \mathcal{F}(t^-)],$$

where the first expression on the right-hand side is the treatment assignment probability in the interval $[t, t+dt)$. Now, it is well-known that because $\eta\{\mathbf{N}^\tau, \mathcal{L}(\tau^-)\}$ is the Radon-Nikodym derivative, we can write causal estimand in terms of the observational data, i.e.

$$\begin{aligned} E_D(Y) &= E\left[\eta\{\mathbf{N}^\tau, \mathcal{L}(\tau^-)\}Y\right] = \\ &\int_{\mathcal{S}_0} \cdots \int_{\mathcal{S}_\tau} \int_{y \in \text{Supp}[Y|\mathcal{F}(\tau)]} y P[Y \in dy|\mathcal{F}(\tau)] \qquad (7) \\ &\times \mathcal{P}_0^\tau P[\mathbf{S}(dt) \in d\mathbf{s}(dt), \mathbf{Z}(dt) \in d\mathbf{z}(dt)|\mathbf{N}(dt), \mathcal{F}(t^-)]\left\{\mathcal{P}_0^\tau P_D[\mathbf{N}(dt)|\mathcal{F}(t^-)]\right\} \qquad (8) \end{aligned}$$

where $\mathcal{S}_t = \{(\mathbf{s}(dt), \mathbf{z}(dt)) : (\mathbf{s}(dt), \mathbf{z}(dt)) \in \text{Supp}[\mathbf{S}(dt), \mathbf{Z}(dt)|\mathbf{N}(dt), \mathcal{F}(t^-)]\}$ and "Supp" refers to the support of the data in the observational study (Gill and Robins, 2001). In the dynamic infusion policy, the treatment assignment probability in the interval $[t, t + dt)$ is

$$P_D[\mathbf{N}(dt)|\mathcal{F}(t^-)] = \prod_{k=1}^{K} [dA_{D,k}\{t, \mathbf{S}(t^-)\}]^{dN_k(t)}[1 - dA_{D,k}\{t, \mathbf{S}(t^-)\}]^{1-dN_k(t)}.$$

In addition to deriving the causal estimand, the Radon-Nikodym derivative is also useful for constructing estimators in the observational study. First, let $\mathbf{U}(\boldsymbol{\beta}, \mathcal{D})$ be an unbiased estimating function in the dynamic infusion policy for a parameter of interest, $\boldsymbol{\beta}$, and the observed data $\mathcal{D}$;

that is, $\mathbb{P}_n \mathbf{U}(\boldsymbol{\beta}, \mathcal{D})$ is an unbiased estimator for $E_D \mathbf{U}(\boldsymbol{\beta}, \mathcal{D})$ and $E_D \mathbf{U}(\boldsymbol{\beta}_0, \mathcal{D}) = 0$ under the truth, $\boldsymbol{\beta} = \boldsymbol{\beta}_0$. [Note: $\mathbb{P}_n h(W) = m^{-1} \sum_{i=1}^{m} h(W_i)$.] The population parameter $\boldsymbol{\beta}$ may refer to dose effects or effects of subpopulations on the mean outcome, for example. Second, for simplicity, suppose we write the causal estimand $E_D(Y) = \mu(\boldsymbol{\beta}, \mathcal{D})$. One may construct an unbiased estimating function in the observational study by weighting $\mathbf{U}(\boldsymbol{\beta}, \mathcal{D})$ with the Radon-Nikodym derivative. Namely, the following estimating function may be used to consistently estimate $\boldsymbol{\beta}$ using data from the observational study,

$$\mathbb{P}_n[\eta\{\mathbf{N}^{\tau}, \mathcal{L}(\tau^-)\}\mathbf{U}(\boldsymbol{\beta}, \mathcal{D})]. \tag{9}$$

As a final note on inference for the causal estimand $E_D(Y)$, we comment on efficient estimation. Using arguments along the lines of Kalbfleisch and Prentice (2002; § 6.3.2), we construct the observed data likelihood as the following:

$$\left\{ P[Y \in dy | \mathcal{F}(\tau)] \mathcal{P}_0^{\tau} P[\mathbf{S}(dt), \mathbf{Z}(dt) | \mathbf{N}(dt), \mathcal{F}(t^-)] \right\} \mathcal{P}_0^{\tau} P[\mathbf{N}(dt) | \mathcal{F}(t^-)]. \tag{10}$$

In the observational study,

$$\mathcal{P}_0^{\tau} P[\mathbf{N}(dt) | \mathcal{F}(t^-)] = \mathcal{P}_0^{\tau} \prod_{k=1}^{K} [dA_k\{t, \mathcal{L}(t^-)\}]^{dN_k(t)} [1 - dA_k\{t, \mathcal{L}(t^-)]^{1 - dN_k(t)}. \tag{11}$$

The causal estimand $E_D(Y)$ only depends on the expression in curly brackets $\{\cdot\}$ in (10) while (11) are nuisance parameters for the parameter of interest. If we further assume that $\Lambda_k\{t, \mathcal{L}(t^-)\} = \int_0^t \lambda_k\{u, \mathcal{L}(u^-)\}\, du$, then (11) simplifies to

$$\prod_{0 \leq t \leq \tau} \prod_{k=1}^{K} [\lambda_k\{t, \mathcal{L}(t^-)\} dt]^{dN_k(t)} \exp\left[- \int_0^{\tau} I(X_k \geq u) \lambda_k\{u, \mathcal{L}(u^-)\}\, du \right].$$

An optimal estimating function for β will be orthogonal to the score function for the treatment selection probabilities, $\lambda_k\{t, \mathcal{L}(t^-)\}$, $k = 1, \ldots, K$. Under this setup, the treatment selection probabilities $\lambda_k\{t, \mathcal{L}(t^-)\}$ are cause-specific hazard functions and could be modeled parametrically (e.g. Gamma, Gompertz, Weibull) or semi-parametrically (e.g. relative risk model).

4 Extended nonparametric estimators for $E_D(Y)$

In this section, we present a family of local linear estimators for $E_D(Y)$ under the simplified setup discussed in Sect. 2. The proposed estimators are not strictly nonparametric as they involve estimated propensities, functions of data and parameter estimates derived from parametric or semi-parametric models for the amount of "treatment received." The estimators may be considered semi-parametric, in a loose sense, because they contain a nonparametric and parametric component. However, we do not wish to confuse our estimator with Johnson-Tsiatis (2005) "semi-parametric" estimator for the same estimand. We term this class of estimators "extended nonparametric estimators" because they extend nonparametric kernel-type estimators straightforwardly via weighting. Our estimators are sensible when the treatment is continuously observed or assigned in time.

For the decision rule $r\{t, S(t^-)\}$ in (1), Johnson and Tsiatis (2004) argued the following equality

$$\mathcal{P}_0^\tau \left\{ \frac{I[dN(t) = r\{t, S(t^-)\}]}{P[N(dt)|\mathcal{F}(t^-)]} \right\} = \frac{I(X = \ell, \Delta = 1)}{f\{X, \mathcal{L}(X)\}} + \frac{I(X < \ell, \Delta = 0)}{\bar{F}\{X, \mathcal{L}(X)\}}, \tag{12}$$

where $\Lambda\{t, \mathcal{L}(t^-)\} = \int_0^t \lambda\{u, \mathcal{L}(u^-)\}\, du$, $\bar{F}\{t, \mathcal{L}(t^-)\} = \exp[-\Lambda\{t, \mathcal{L}(t^-)\}]$, $f\{t, \mathcal{L}(t^-)\} = \lambda\{t, \mathcal{L}(t^-)\}\bar{F}\{t, \mathcal{L}(t^-)\}$. We define $E_D(Y) = \mu(\ell)$ as the mean potential outcome for a patient following the target infusion-length policy (i.e. an infusion-length policy for ℓ units of time). Again, the infusion-length policy is an example of a dynamic infusion policy; hence, our choice of notation $E_D(Y) = \mu(\ell)$ is a special case of the more general $\mu(\beta, \mathcal{D})$ given in Sect. 3.2. To estimate the mean potential endpoint assuming all patients followed the target infusion-length policy, (9) suggests the following simple estimating function for $\mu(\ell)$:

$$\mathbb{P}_n \left[(Y - \mu(\ell)) \left\{ \frac{I(X = \ell, \Delta = 1)}{f\{X, \mathcal{L}(X)\}} + \frac{I(X < \ell, \Delta = 0)}{\bar{F}\{X, \mathcal{L}(X)\}} \right\} \right]. \tag{13}$$

However, $I(X = \ell, \Delta = 1)$ is an event of probability measure zero when the observed infusion length X is observed continuously in time. To overcome this difficulty, Johnson and Tsiatis (2004)

discretized the data and subsequently redefined $\bar{F}\{t, \mathcal{L}(t^-)\}$ through the product of discrete hazard functions. Johnson and Tsiatis (2005) parameterized the causal estimand using a linear logistic model and polynomials in treatment duration. Below, we suggest simple local linear estimators for $\mu(\ell)$.

Following (13), it is natural to believe that the estimator defined by the following estimating function would be consistent under suitable regularity conditions:

$$\mathbb{P}_n \left[(Y - \mu(\ell)) \left\{ \frac{b^{-1} I(X \in [\ell - b, \ell + b), \Delta = 1)}{f\{X, \mathcal{L}(X)\}} + \frac{I(X < \ell, \Delta = 0)}{\bar{F}\{X, \mathcal{L}(X)\}} \right\} \right], \tag{14}$$

for b an unknown bandwidth decreasing to zero. Now, we define the scaled kernel function $K_b(t) = b^{-1}K(t/b)$, and assume throughout that the kernel function $K(t)$ is a density on $(-\infty, \infty)$ satisfying

$$\int sK(s)ds = 0, \ \int s^2 K(s)ds = \sigma_K^2, 0 < \sigma_K^2 < \infty. \tag{15}$$

The estimator in (14) is a special case of the following kernel estimator

$$\mathbb{P}_n \left[(Y - \mu(\ell)) \left\{ \frac{\Delta K_b(\ell - X)}{f\{X, \mathcal{L}(X)\}} + \frac{I(X < \ell, \Delta = 0)}{\bar{F}\{X, \mathcal{L}(X)\}} \right\} \right]. \tag{16}$$

We note that both (14) and (16) include data from patients whose attending physician discontinued (by choice) the infusion process in the interval $[t - b, t + b)$. However, it may be wise to place more weight on data from patients whose attending physician stopped the infusion process close to time t; this is the primary role of the kernel function in (16). For suitably defined kernel function, e.g. Epanechnikov kernel,

$$K(u) = 0.75(u - 1)^2 I(-1 \leq u \leq 1),$$

the estimating function in (16) places less weight on data where a physician's decision to stop treatment is farther from the treatment time of interest t, as desired. However, it is well-known that the estimator defined as the solution to the estimating function in (16) has bias $O(b^{-1})$. Local polynomial estimators (Fan and Gijbels, 1996) offer a simple way for constructing estimators with

a smaller order bias than the estimator in (16). Let the bivariate vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$. The new estimator is the estimated intercept $\hat{\alpha}_0$, where $\hat{\alpha} = \arg\min Q_b(\boldsymbol{\alpha})$ and

$$Q_b(\boldsymbol{\alpha}) = \mathbb{P}_n \left[\Delta\{Y - \alpha_0 - \alpha_1(\ell - X)\}^2 \frac{K_b(\ell - X)}{f\{X, \mathcal{L}(X^-)\}} + (Y - \alpha_0)^2 \frac{I(X < \ell, \Delta = 0)}{\bar{F}\{X, \mathcal{L}(X^-)\}} \right]. \tag{17}$$

Note that when $\alpha_1 = 0$, $(\partial/\partial\alpha_0) Q_b(\boldsymbol{\alpha}) = (16)$. Hence, the solution to the estimating function (16) is often referred to as a zero-degree local polynomial estimator.

In our definition of the objective function $Q_b(\boldsymbol{\alpha})$, the aggregate quantities $f\{X, \mathcal{L}(X)\}$ and $\bar{F}\{X, \mathcal{L}(X)\}$ are unknown and must be estimated from the observed data. As defined earlier, $f\{X, \mathcal{L}(X)\}$ and $\bar{F}\{X, \mathcal{L}(X)\}$ are functions of the treatment selection probabilities $\lambda\{t, \mathcal{L}(t^-)\}$. In our application, the treatment selection probabilities $\lambda\{t, \mathcal{L}(t^-)\}$ are functions of the observed data $\mathcal{L}(t^-)$ through a finite number of parameters, say ψ. In practice, we model the cause-specific hazard $\lambda\{t, \mathcal{L}(t^-); \psi\}$ as a function $\mathcal{L}(t^-)$ and estimate the unknown parameters ψ using regression models for event-time data (cf. Kalbfleisch and Prentice, 2002, chs. 2-4; cf. Fleming and Harrington, 1991, ch. 4). Our estimator $\hat{\mu}(t)$ is defined as the estimated intercept $\hat{\alpha}_0$ derived from minimizing $Q_b(\boldsymbol{\alpha})$ when $f\{X, \mathcal{L}(X)\}$ and $\bar{F}\{X, \mathcal{L}(X)\}$ are replaced with the estimated quantities $f\{X, \mathcal{L}(X); \hat{\psi}\}$ and $\bar{F}\{X, \mathcal{L}(X); \hat{\psi}\}$, respectively, in (17). Additional practical details, including a data-driven cross-validation procedure for selecting the bandwidth b, are outlined in the Appendix.

4.1 Statistical inference

In general, in addition to technical regularity conditions and the no unmeasured confounders assumption, the asymptotic properties of estimators for marginal mean models in dynamic regimes rely on two conditions: (i) correct parametric model specification for $E_D(Y) = \mu(\boldsymbol{\beta}, \mathcal{D})$ (See Subsection 3.2), and (ii) correct (semi-)parametric model specification for the treatment selection probabilities, $\lambda\{t, \mathcal{L}(t^-); \psi\}$. We note that some authors attempt to relax the requirement that both (i)

and (ii) are needed to achieve consistency. One can relax (i) through modeling nested functions of the treatment selection probabilities in the dynamic treatment regime, say $\xi\{t, \mathcal{L}(t^-)\}, t \leq \tau$ (these functions are called $g_t(\cdot)$ in MLR). Then, one may construct an estimator which is consistent and asymptotically normal if either $\lambda\{t, \mathcal{L}(t^-)\}$ or $\{\xi\{t, \mathcal{L}(t^-)\}, t \leq \tau\}$ are modeled correctly. This is called the "double robustness property" (Scharfstein, Rotnizky, and Robins, 1999). In practice, it is difficult to posit consistent models for $\xi(\cdot)$ due to its nested structure (See Murphy et al., 2001, 5.1-5.2). On the other hand, $\lambda\{t, \mathcal{L}(t^-)\}$ and $\mu(\boldsymbol{\beta}, \mathcal{D})$ may be modeled with conventional regression methods. Hence, many statisticians simply work with assumptions (i) and (ii) directly or indirectly.

For estimating the causal estimand $E_D(Y) = \mu(\ell)$, Johnson and Tsiatis (2005) parameterize $\mu(\ell) = \mu(\boldsymbol{\beta}, \ell)$ and estimate $\lambda\{t, \mathcal{L}(t^-)\}$ semi-parametrically. In this paper, we estimate $\lambda\{t, \mathcal{L}(t^-)\}$ semi-parametrically and $\mu(\ell)$ nonparametrically. That is, Johnson and Tsiatis (2005) assume both conditions (i) and (ii) whereas our estimator $\hat{\mu}(\ell)$ only assumes condition (ii). In place of condition (i), we make weaker assumptions, including $\mu(\ell)$ is twice-differentiable for every $\ell,\ 0 \leq \ell \leq \tau$, kernel conditions (15), and $b_n \sim n^{-1/5}$. In fact, we conjecture that, under suitable regularity conditions, one can show that $n^{1/2}\{\hat{\mu}(\ell) - \mu(\ell)\}$ converges in distribution to a Gaussian random variable with mean $B(\ell)$ and variance $\sigma^2(\ell)$. Because our estimator requires condition (ii) — a correctly specified parametric or semi-parametric model for the treatment selection probabilities — where general nonparametric statistics do not impose such restrictions, a rigorous proof of asymptotic normality may be of modest interest. Heuristically, the argument is outlined as follows: in addition to regularity conditions, assume $n^{1/2}\|\hat{\psi}_n - \psi_0\| = O_p(1)$. Then, because the nonparametric estimators converge much slower, it is sufficient to consider the distribution of $n^{1/2}\{\hat{\mu}(\ell) - \mu(\ell)\}$ with $\psi = \psi_0$, the true value. Now, assuming (ii) with $\psi = \psi_0$, allows us to appeal to the usual asymptotic arguments for local linear estimators (Fan, Heckman, and Wand, 1995). We note that the bias $B(\ell)$ and variance $\sigma^2(\ell)$ terms will involve integrals with respect to the joint distribution of the

all the data $(Y, X, \Delta, Z(u),\ u \leq X)$, and not simply the joint distribution of $(Y, X, \Delta = 1)$. In simulation studies (not shown but available from the authors), we found several scenarios where our local linear estimator $\hat{\mu}(\ell)$ gave competitive estimates of $\mu(\ell)$, in terms of integrated mean square error (IMSE), when compared with JT (2005) estimator $\mu(\hat{\boldsymbol{\beta}}, \ell)$ under simple linear logistic models for $\mu(\ell)$. However, the JT (2005) semi-parametric estimator performed poorly (in terms of IMSE when compared with $\hat{\mu}(\ell)$) when the true curve $\mu(\ell)$ could not be adequately described by three- and four-degree polynomials. Our simulation results are not surprising; rather, they confirm our intuition. Our experience suggests that there is room for improvement in estimating $\mu(\ell)$ because, even in the simplest cases, $\mu(\ell)$ is a complicated integral with respect to the joint distribution of several random variables and may not be easily characterized by the usual parametric formulations.

In this paper, we advocate bootstrap methods to obtain standard error estimates for $\hat{\mu}(\ell)$. Namely, draw a sample of m indices with replacement from $\{1, \ldots, m\}$. Each index corresponds to a row in the data set and all the rows corresponding to the first collection of indices forms the first bootstrap data set. Construct estimates of $\mu(\ell)$ from the first bootstrap data set, say $\hat{\mu}_1^*(\ell)$. Repeat the process $B = 1000$ times to form the sample of bootstrap estimates $\{\hat{\mu}_1^*(\ell), \ldots, \hat{\mu}_B^*(\ell)\}$. Finally, the sample percentiles or sample standard deviation from the bootstrap estimates may be used to construct pointwise confidence intervals for $\hat{\mu}(\ell)$.

5 Application to the ESPRIT Infusion Trial

We now apply our methods to data from the ESPRIT infusion trial conducted at Duke University Medical Center. A total of 1036 patients were randomized to the Integrilin arm of the study. Of these, 106 patients ($\sim$ 10% of 1036) experienced a treatment-competing event. The random variable X denotes the observed infusion length and Δ is an indicator realizing the value one if

the attending physician stopped the infusion process by choice and zero if the infusion was stopped by an infusion-terminating event, such as abrupt closure, no reflow, or coronary thrombosis. To facilitate comparisons between our results to those of Johnson and Tsiatis (2004; 2005), we include the following potential confounders $\mathbf{Z}$ in our analysis — diabetes (0/1), percutaneous transluminal coronary angioplasty (PTCA,0/1), angina (0/1), heparin (0/1), and weight (in kilograms).

The results of our analyses are presented in Figure 1. We illustrate our methods through three estimators: $\hat{\mu}_1(\ell)$, $\hat{\mu}_2(\ell)$, and $\hat{\mu}_2^{(0)}(\ell)$. The treatment selection probabilities $\lambda\{t, \mathcal{L}(t^-)\}$ are modeled through the Weibull family of distributions in $\hat{\mu}_1(\ell)$ and through the Cox model in $\hat{\mu}_2(\ell)$, $\hat{\mu}_2^{(0)}(\ell)$. The estimator $\hat{\mu}_2^{(0)}(\ell)$ does not use any of the potential confounding variables $\mathbf{Z}$ in models of the treatment selection probabilities. All estimators use the Epanechnikov kernel. We exemplify our methods with a variety of bandwidths, chosen largely for comparison purposes. One figure illustrates our proposed data-based bandwidth selection procedure. Point estimates from earlier analyses, denoted with Xs and Ps, are taken from Johnson and Tsiatis (2004).

[Figure 1 about here]

Figure 1: Estimates of $\mu(\ell)$ from the ESPRIT Infusion Trial. (a) A comparison of $\hat{\mu}_1(\ell)$ (solid line), $\hat{\mu}_2(\ell)$ (dashed line), and $\hat{\mu}_2^{(0)}(\ell)$ (dotted line) for a single bandwidth; (b) A comparison of $\hat{\mu}_2(\ell)$ using different bandwidths, dotted line (b=0.65), solid line (b=1.0), dashed line (b=1.35); (c) A display of $\hat{\mu}_2(\ell)$ with an approximate 95% pointwise confidence interval and compared to point estimates (marked by an "X") and estimated standard errors from Johnson and Tsiatis (2004); (d) A comparison of the Integrilin (solid line) versus placebo (dashed line) groups using $\hat{\mu}_2(\ell)$ with point estimates from an earlier analysis, "X" for the Integrilin and "P" for placebo.

In Figure 1(a), we plot the mean response as a function of policy using the different estimators, $\hat{\mu}_1(\ell)$ (dashed line), $\hat{\mu}_2(\ell)$ (solid line) and $\hat{\mu}_2^{(0)}(\ell)$ (dotted line). Here, we use a bandwidth $b = 1.0$ which corresponds roughly to the grouping rule used by Johnson and Tsiatis (2004). Not surprisingly, then, all estimated curves agree with the point estimates. Here, we note $\hat{\mu}_2(\ell)$ and $\hat{\mu}_2^{(0)}(\ell)$ follow a similar trend. It is interesting to note that $\hat{\mu}_1(\ell)$ appears modestly higher for policies around 16

hours. Figure 1(b) is a display of $\hat{\mu}_2(\ell)$ using three different bandwidths: $b = 0.65$ (dotted line), $b = 1$ (solid line), and $b = 1.35$ (dashed line). Compared to curve with larger bandwidths, $\hat{\mu}_2(\ell)$ with bandwidth $b = 1.0$ now seems to lead to a fitted curve with too many peaks and valleys. Now, we interpret the overall trend of $\hat{\mu}_2(\ell)$ with bandwidth $b = 1.35$ — the average 30-day endpoint increases gradually with longer infusion-length policies up to about 20 hours, then flattens out for roughly two hours, and then increases steadily. At the same time, we note that the overall difference in estimated mean response from 16 to 24 hours is less than 0.08.

Next, we wish to summarize the standard error estimates in Figure 1(c). We plot the pointwise 95% confidence intervals using bootstrap standard errors for $\hat{\mu}_2(\ell)$ with bandwidth $b = 1.0$. We also include the five 95% confidence intervals given by the Johnson-Tsiatis (2004) point estimates and their respective standard errors. As expected, the pointwise confidence intervals for $\hat{\mu}_2(\ell)$ match well the Johnson-Tsiatis confidence intervals for the particular bandwidth $b = 1.0$. Finally, in Figure 1(d), we plot $\hat{\mu}_2(\ell)$ for both the placebo and Integrilin arms of the ESPRIT trial using the cross-validated bandwidth $b = 2.52$. We plot the Johnson-Tsiatis point estimates for the Integrilin arm (denoted with an "X") as well as the point estimates for the placebo arm (denoted with a "P"). Now, we see somewhat simpler duration-response trend than the one presented in Figure 1(c), and perhaps easier to interpret. Now, we observe the duration-response trend as relatively flat from 16 to 22 hours and then gradually increasing to 24 hours. Johnson and Tsiatis (2004) interpret this observation as mild evidence that there may be little benefit infusing patients longer than 16 hours and that infusing longer than 22 hours "may actually be harmful." Our analyses in panel (d) suggests a cautious interpretation, one that appears to depend somewhat on the chosen bandwidth. In particular, the local linear estimate $\hat{\mu}_2(\ell)$ with cross-validated bandwidth appears to be over-smoothing.

6 Remarks

We discuss treatment-competing events in the context of dynamic infusion policies, i.e. random dynamic treatment regimes where physicians make decisions when to stop or continue infusion based on time-dependent measures. When treatment decisions occur over time, treatment-competing events dictate how much of the (treatment) decision process will be observed. We argued that treatment-competing events may be incorporated into dynamic regimes, as they were incorporated into dynamic infusion policies. When treatment-competing events play some role in the dynamic regime but they are inadvertently or explicitly ignored, Johnson and Tsiatis (2004; 2005) argued this will generally lead to biased estimates of the causal estimand of interest. Furthermore, the marginal model formulation has a "population-average" interpretation even though the dynamic treatment regimes (and dynamic infusion policies) may be defined through an individual patient's intermediate response to current therapies. In the case of infusion-length policies, our parameter estimates have the interpretation of the mean 30-day endpoint had the population followed the policy for ℓ units of time.

Our paper extends some of the main concepts in Murphy et al. (2001) to the random dynamic treatment regime subject to treatment-competing events. In the case of infusion trials, such as the ESPRIT infusion trial, we termed this dynamic regime a dynamic infusion policy. The counting process methodology facilitated our development and allowed for a precise definition of otherwise complicated concepts. We used our framework to define the nonrandom dynamic treatment regime "infusion-length policy for t units of time" (Johnson and Tsiatis, 2004, 2005). We proposed a local linear estimator $\hat{\mu}(\ell)$ which we argued, heuristically, leads to consistent estimators of the causal estimand $\mu(\ell)$ under weaker assumptions than the semi-parametric estimator proposed by Johnson and Tsiatis (2005).

We believe the methods considered in this article may well extend beyond the specific infusion application described above. The following two collaborative efforts between the authors and relevant substantive co-investigators exemplify treatment-competing events in dynamic regimes.

1. In the Adult Clinical Trials Group (ACTG) study A5095 (Gulick, 2004), patients are started on one of three antiretroviral (ARV) regimens: (a) Abacavir (ABC)+Lamivudine (3TC)+Zidovudine (ZDV), (b) 3TC+ZDV+Efarirenz (EFV), or (c) ABC+3TC+ZDV+EFV. After the first virologic failure, defined as two subsequent assays where HIV-RNA levels exceed 200 copies/mL, patients are allowed to switch ARV regimens. An extremely important question in the HIV community is whether to "switch-early" or "switch-late" for optimal long-term suppression of viral load levels. There are compelling biological arguments to believe a patient should switch soon after the viral load exceeds the 200 copies/mL threshold; however, different but equally compelling arguments suggest one should allow the viral load levels to increase to much higher levels, e.g. 8,000-10,000 copies/mL, before switching ARV regimens. Our methods are relevant here since several AIDS-defining events — including patient refusal, opportunistic infections, and other adverse events — preclude switching ARV-regimens.

2. Similarly, the treatment of Tourette's syndrome (TS) includes a complex course of drug regimens, including anti-hypertensives and anti-psychotics, and behavioral interventions (cf. Olson, 2006). Psychiatrists often subscribe drug regimens based on a patient's tic (involuntary body movements) severity. A particular dynamic regime of interest is to subscribe treatment regimens that address the most severe tic first, the second most severe tic second, and so on. However, co-morbidity is a serious consideration as patients who suffer from TS also often suffer from other neurological disorders such as obsessive compulsive disorder (OCD) or schizophrenia. It is common that a particular OCD or schizophrenic event will trigger an immediate change in treatment regimens, and the usual treatment regime following tic

severity must be discontinued or modified significantly. Hence, if a schizophrenic, OCD, or other adverse event occurs at time t, then all treatment decisions after t that would have occurred on the original dynamic regime following tic severity are unobserved.

Therefore, in these two examples, it seems that practical dynamic regimes will be defined as following dynamic regime A if the competing event does not occur and following regime B if the event occurs. The whole regime could be define "A or B" depending on the treatment-competing event. An important distinction between the infusion policies and dynamic treatments for HIV and TS is that drugs are often switched and substituted when intermediate, adverse events occur in the latter examples. In infusion policies, a treatment-competing event discontinues the infusion abruptly and the infusion is not restarted at some later time. Switching times and treatments within a dynamic regime is not covered by the methods discussed in this paper but an important topic for future research.

Acknowledgements

The authors gratefully acknowledge the help of the associate referee and one anonymous referee whose comments significantly improved the presentation.

References

P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, Statistical Models Based on Counting Processes, Springer-Verlag: New York, 1993.

J. Fan and I. Gijbels, Local Polynomial Modelling and Its Applications, Chapman & Hall: London,

1996.

J. Fan, N. E. Heckman, and M. P. Wand, "Local polynomial kernel regression for generalized linear models and quasi-likelihood functions," J. Am. Stat. Assoc. vol. 90 pp. 141–150, 1995.

T. A. Fleming and D. P. Harrington, Counting Processes and Survival Analyses, Wiley: New York, 1991.

R. M. Gulick et al., "Triple-nucleoside regimens versus efavirenz-containing regimens for the initial treatment of HIV-1 infection," New Eng. J. Med. vol. 350 1850–1861, 2004.

B. A. Johnson and A. A. Tsiatis, "Estimating mean response as a function of treatment duration in an observational study, where duration may be informatively censored," Biometrics vol. 60 pp. 315–323, 2004.

B. A. Johnson and A. A. Tsiatis, "Semiparametric inference in observational duration-response studies, with duration possibly right-censored," Biometrika vol. 92 pp. 605–618, 2005.

J. D. Kalbfleisch and R. L. Prentice, The Statistical Analysis of Failure Time Data, Wiley: New York, 2002.

J. J. Lok . "Statistical modelling of causal effects in time," Ph. D. Thesis, Division of Mathematics and Computer Science, Free University, Amsterdam, 2001.

J. Lok, R. Gill, A. van der Vaart and J. M. Robins, "Estimating the causal effect of a time-varying treatment on time-to-event using structural nested failure time models," Stat. Neer. vol. 58 pp. 1–25, 2004.

S. A. Murphy, M. J. van der Laan and J. M. Robins, "Marginal mean models for dynamic regimes," J. Am. Stat. Assoc. vol. 96 pp. 1410–23, 2001.

L. L. Olson et al., "Tourette's syndrome: diagnosis, strategies, therapies, pathogenesis, and future research directions," J. Child Neur. vol. 21 pp. 630–41, 2006.

H. Ramlau-Hansen, "Smoothing counting process intensities by means of kernel functions," Ann. Stat. vol. 11 pp. 453–66, 1983.

A. G. Rebeiz et al., "Optimal duration of eptifibatide infusion in percutaneous coronary intervention (an ESPRIT substudy)," Am. J. Cardiol. vol. 94 pp. 926–929, 2004.

J. M. Robins, "A new approach to causal inference in mortality studies with sustained exposure periods — application to control of the healthy worker survivor effect," Math. Mod. vol. 7 pp. 1393–1512, 1986.

D. B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies," J. Ed. Psych. vol. 6 pp. 688–701, 1974.

D. B. Rubin, "Comment: 'Which Ifs Have Causal Answers,'" J. Am. Stat. Assoc. vol. 81 pp. 961–962, 1986.

D. O. Scharfstein, A. Rotnizky, and J. M. Robins, "Adjusting for nonignorable drop-out using semiparametric nonresponse models," J. Am. Stat. Assoc. vol. 94 pp. 1096–1120, 1999.

Appendix A: Modeling $\lambda\{t, \mathcal{L}(t^-)\}$

In our estimator $\hat{\mu}(t)$, the functions $f\{X, \mathcal{L}(X)\}$ and $\bar{F}\{X, \mathcal{L}(X)\}$ are defined through the unknown cause-specific hazard function $\lambda\{t, \mathcal{L}(t^-)\}$. In this paper, we consider two models of practical interest. First, define $\lambda\{t, \mathcal{L}(t^-); \psi\}$ parametrically through the Weibull family of distribuions ψ and second, define $\lambda\{t, \mathcal{L}(t^-); \psi\}$ semi-parametrically via the relative-risk (Cox) model. We distinguish the two resulting estimators for $\mu(t)$ through different subscripts: $\hat{\mu}_1(t)$ and $\hat{\mu}_2(t)$ when the cause-specific hazard functions are modeled parametrically and semi-parametrically, respectively. In this case of $\hat{\mu}_1(t)$, we define

$$\lambda\{t, \mathcal{L}(t^-); \psi\} = \alpha\rho t^{\alpha-1} \exp\{\alpha\gamma'\mathbf{Z}(t)\},$$

where $\psi = (\alpha, \rho, \gamma')'$. In the case of $\hat{\mu}_2(t)$,

$$\lambda\{t, \mathcal{L}(t^-); \psi\} = \lambda_0(t) \exp\{\gamma'\mathbf{Z}(t)\},$$

where $\psi = (\gamma, \lambda_0(u),\ u \leq \tau)$ and $\lambda_0(t)$ is an unspecified baseline hazard function. While $\hat{\mu}_2(t)$ is more flexible than $\hat{\mu}_1(t)$ in modeling the probability of continued treatment duration, $\hat{\mu}_2(t)$ now requires an estimate of the baseline hazard function $\lambda_0(t)$. In keeping with our use of kernel-based estimators, we propose to use the Ramlau-Hansen (1983a, 1983b) to estimate the baseline hazard function in $\hat{\mu}_2(t)$, i.e.

$$\hat{\lambda}_0(t) = n^{-1} \sum_{i=1}^{n} \int K_b(t-s) d\widehat{\Lambda}_0(s; \gamma),$$

where

$$\widehat{\Lambda}_0(t; \gamma) = \sum_{i=1}^{n} \int_0^t \frac{dN_i(u)}{\sum_j \exp\{\gamma'\mathbf{Z}_j(u)\} R_j(u)}.$$

Once the baseline hazard function is estimated, (17) may be evaluated using the observed data.

Appedix B: Data-based bandwidth selection

For our local linear estimator, $\hat{\mu}(t)$, we need to select a bandwidth b. Here, we propose an extension of V-fold cross-validation for regression problems with censored predictors. Let $\mathcal{D}$ denote the full data set and $\mathcal{D} - \mathcal{D}^v$ and $\mathcal{D}^v$ denote the cross-validation training and test data sets, respectively. For each b and $v = 1, \ldots, V$, we find the estimator $\hat{\mu}_b^{(v)}(t)$ of $\mu(t)$ using the training set $\mathcal{D} - \mathcal{D}^v$. Let $h(W)$ be some function of the observed data $\mathbf{O} = (Y, X, \Delta, \mathbf{Z}(u),\ u \leq X)$. We let $\mathbb{P}_v h(W)$ denote the average statistic using the observations in the v-th cross-validated data set, i.e.

$$\mathbb{P}_v h(W) = m_v^{-1} \sum_{\{k: \mathbf{O}_k \in \mathcal{D}^v\}} h(W_k),$$

and $m_v = |\mathcal{D}^v|$, the number of subjects in the v-th cross-validated data set. Our goal is to construct a natural extension of the following cross-validated statistic from nonparametric regression, i.e.

$$\mathrm{CV}_F(b) = \sum_{v=1}^{V} \mathbb{P}_v \left[Y - \hat{\mu}_b^{(v)}(X) \right]^2,$$

if we had observed the random sample $(Y_i, X_i),\ i = 1, \ldots, m$ without any discussion of dynamic regimes and/or confounding. In this paper, we propose the following cross-validation criterion to select the bandwidth parameter b,

$$\mathrm{CV}(b) = \sum_{v=1}^{V} \mathbb{P}_v \left[\eta\{\mathbf{N}, \mathcal{L}(\tau^-)\} Q_b^{(v)}(\mathbf{O}) \right] \tag{18}$$

where $\eta\{\mathbf{N}, \mathcal{L}(\tau^-)\}$ is the Radon-Nikodym derivative and

$$Q_b^{(v)}(\mathbf{O}) = \Delta_k \{Y - \hat{\mu}_b^{(v)}(X)\}^2 + \frac{(1 - \Delta_k)}{\bar{H}(X)} \int_X^{\tau} \{Y - \hat{\mu}_b^{(v)}(t)\}^2 h(t)\, dt,$$

where $h(t)$ is any probability density function with support over $[0, \tau]$, and $\bar{H}(t) = \int_t^{\infty} h(u)\, du$. The criterion is motivated through the following heuristic reasoning. The objective function $Q_b^{(v)}(\mathbf{O})$ can be motivated through an "idealized randomized design" (Johnson and Tsiatis, 2005), that is, a completely randomized study where the physicians intend to infuse patients for T units of time

but cannot due to an infusion-censoring event at time $X^c < T$. In the idealized randomized design, we wish to construct statistics in terms of the random variables (Y, X, Δ) where $X = \min(T, X^c)$ and $\Delta = I(T \leq X^c)$. In this design, we assume that $h(t)$ is the probability density function of the intended, randomly selected infusion length T. Using arguments similar to Johnson and Tsiatis (2005, p.608),

$$E_D\left[g\{Y - \mu(T)\}\right] = E_D\left[\Delta g\{Y - \mu(T)\} + \frac{(1-\Delta)}{\bar{H}(X)}\int_X^\tau g\{Y - \mu(t)\}h(t)\,dt\right], \tag{19}$$

where $E_D(\cdot)$ is the expectation in the "idealized randomized design." [Note: that the idealized randomized design is an example what we call an dynamic infusion policy.] Applying the function $g(z) = z^2$ in (19) and substituting the estimator $\hat{\mu}_b^{(v)}(t)$ for $\mu(t)$ leads to $Q_b^{(v)}(\mathbf{O})$. We weight the objective function $Q_b^{(v)}(\mathbf{O})$ by the Radom-Nikodym derivative to account for the potential confounding in the observational data. Johnson and Tsiatis used the function $h(t) = \lambda_0(t)$ and we implemented the same function in our cross-validation criterion.

Appedix C: Pseudo code for implementation

In this section, we offer the following steps in pseudo code to implement the methods discussed above. Define a coarse grid of interest $\mathcal{T} = \{t_1, \ldots, t_m\}$.

S1. Estimate ψ using standard survival analysis methods briefly outlined in Appendix A.

S2. Calculate $\hat{\pi}_i^{-1}$ where $\hat{\pi}_i = \Delta_i f\{X_i, \mathcal{L}(X_i); \hat{\psi}\} + (1-\Delta_i)\bar{F}\{X_i, \mathcal{L}(X_i); \hat{\psi}\}$, and $f\{X_i, \mathcal{L}(X_i); \hat{\psi}\}$ and $\bar{F}\{X_i, \mathcal{L}(X_i); \hat{\psi}\}$ were introduced in (12).

S3. For each $t \in \mathcal{T}$,

S3(i). Calculate $w_i(t) = \hat{\pi}_i^{-1}\{\Delta_i K_b(t - X_i) + (1 - \Delta_i)I(X_i < t)\}$.

S3(ii). Calculate $\hat{\mu}(t) = \sum_i w_i(t) Y_i / \sum_i w_i(t)$.

Steps S1-S3 compute a zero-degree polynomial curve estimate $\hat{\mu}(t)$. For arbitrary degree polynomial curves, Fan and Gijbels (1996, §3.1) describe local polynomial estimators through weighted regression. For our problem, replace step S3(ii) with

T3. For each $t \in \mathcal{T}$,

T3(i) . Define $\mathbf{W} = \text{diag}[w_1(t), \ldots, w_n(t)]$

T3(ii). Define the solution vector $\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{Y}$, where

$$\mathbf{X} = \begin{pmatrix} 1 & \Delta_1(t - X_1) & \cdots & \Delta_1(t - X_1)^p \\ \vdots & \vdots & & \vdots \\ 1 & \Delta_n(t - X_n) & \cdots & \Delta_n(t - X_n)^p \end{pmatrix},$$

$\mathbf{Y} = (Y_1, \ldots, Y_n)$. The estimate $\hat{\mu}(t)$ is defined as the estimated intercept $\hat{\alpha}_0$.

The methods described in Sect. 4 use a local linear estimator, thus $p = 1$. The bandwidth is chosen using methods in Appendix B.