On estimating optimal regime for treatment initiation time based on restricted mean residual lifetime

Xin Chen¹,² | Rui Song³ | Jiajia Zhang⁴ | Swann Arp Adams⁴,⁵ | Liuquan Sun⁶ | Wenbin Lu³

¹ School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai, China
² Interdisciplinary Research Institute of Data Science, Shanghai Lixin University of Accounting and Finance, Shanghai, China
³ Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA (Email: lu@stat.ncsu.edu)
⁴ Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, South Carolina, USA
⁵ College of Nursing, University of South Carolina, Columbia, South Carolina, USA
⁶ Institute of Applied Mathematics, Chinese Academy of Science, Beijing, China

Correspondence Xin Chen, School of Statistics and Mathematics, Interdisciplinary Research Institute of Data Science, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China. Email: chenxin@amss.ac.cn

Funding information NIH, Grant/Award Numbers: P01 CA142538, NSF-DMS 1555244; National Natural Science Foundation of China, Grant/Award Numbers: 11690015, 11771431, 12001371; Key Laboratory of RCDS; CAS, Grant/Award Number: 2008DF173182; Shanghai Sailing Program, Grant/Award Number: 20YF1434200

Abstract
When to initiate treatment on patients is an important problem in many medical studies such as AIDS and cancer. In this article, we formulate the treatment initiation time problem for time-to-event data and propose an optimal individualized regime that determines the best treatment initiation time for individual patients based on their characteristics. Different from existing optimal treatment regimes where treatments are undertaken at a pre-specified time, here new challenges arise from the complicated missing mechanisms in treatment initiation time data and the continuous treatment rule in terms of initiation time. To tackle these challenges, we propose to use restricted mean residual lifetime as a value function to evaluate the performance of different treatment initiation regimes, and develop a nonparametric estimator for the value function, which is consistent even when treatment initiation times are not completely observable and their distribution is unknown. We also establish the asymptotic properties of the resulting estimator in the decision rule and its associated value function estimator. In particular, the asymptotic distribution of the estimated value function is nonstandard, which follows a weighted chi-squared distribution. The finite-sample performance of the proposed method is evaluated by simulation studies and is further illustrated with an application to a breast cancer data.

KEYWORDS
individualized treatment regime, kernel estimation, optimal treatment initiation time, time-to-event data, value function

1 | INTRODUCTION
Finding the optimal time to initiate treatment is a critical issue in many medical studies including AIDS and cancer. For example, in the treatment of patients with tuberculosis and newly identified infection with human immunodeficiency virus (HIV), antiretroviral therapy (ART) must be started during the treatment for tuberculosis.
Another example is the initiation of adjuvant therapy for patients diagnosed with breast cancer. In practice, adjuvant chemotherapy or radiotherapy is routinely recommended to breast cancer patients after definitive surgery and within 24 weeks from the surgery (Lohrisch et al., 2006). However, the optimal time to initiate adjuvant therapy during the 24 weeks after surgery is controversial. Due to the heterogeneity in patients and diseases, several retrospective studies evaluating the role of early or delayed initiation of adjuvant chemotherapy reported conflicting results (Yu et al., 2017). Similar treatment initiation time problems also arise in finding the timing of neurosurgery for medically refractory epilepsies (Sugano and Arai, 2015) and the timing for cardiovascular surgeries (Jung et al., 2019).

The goal of this article is to develop a method to search for the optimal individualized treatment initiation regime (OTIR) that selects the best treatment initiation time based on each individual’s characteristics. Specifically, we focus on the case where the outcome of interest is time-to-event data, such as the death time of patients, and initiating treatment at the proper time within a pre-specified time range may decrease the hazard rate of the failure event. In the example of breast cancer patients, the pre-specified time range could be the 24 weeks after surgery. In the area of precision medicine, although numerous efforts have been made in finding optimal individualized treatment regime (OTR) for discrete choices of treatment options (Watkins and Dayan, 1992; Blatt et al., 2004; Murphy, 2003; Qian and Murphy, 2011; Zhang et al., 2012; Zhao et al., 2012), the optimal individualized treatment initiation problem for time-to-event data has been seldom studied, and the estimation of OTIR has to tackle several new challenges.

One main challenge in the estimation of OTIR rises from the missing of the treatment initiation time. In clinical studies, although every patient would be assigned with a treatment initiation time, the value of the assigned treatment initiation time would be missing if the patient does not survive to the assigned treatment initiation time. Therefore, the potential outcome of a given treatment initiation time assignment regime cannot be evaluated directly.

Another challenge in the searching for OTIR comes from the continuity of treatment initiation time. In the treatment initiation time data, since the set of decision options is a time period containing an infinite number of time points, the relationship among the treatment options, covariates and outcome of interest could be too complicated to be correctly specified, and thus, the regression-based estimation methods may not be suitable. Besides, when treatment initiation time follows a continuous distribution, the probability that the observed treatment initiation time exactly matches a given regime is zero. Thus, existing value search methods (Zhang et al., 2012; Zhao et al., 2012), which evaluate each treatment regime based on samples whose treatment option exactly follows this regime, cannot be directly applied to treatment initiation time data. Lastly, since in practice, when to initiate treatment is usually decided by physicians based upon patients’ status, there could exist unknown dependence between assigned treatment initiation times and covariates, and statistical inference should consider such dependence.

In the literature, several works related to the estimation of OTIR have been conducted. In estimating the effect of treatment initiation time on an outcome measured at a fixed duration after initiation, Lok and DeGruttola (2012) discretized the treatment initiation time into multiple treatment points and developed structural nested mean models to deal with the nonrandom assignment of treatment initiation time in observational data. Zhao et al. (2011) presented an adaptive reinforcement learning approach to discover the optimal individualized treatment regimen that selects the optimal time to initiate second-line therapy in a specially designed clinical reinforcement trial. Hu et al. (2018) proposed a structural proportional hazards model to evaluate the effect of treatment initiation time on the survival time in the absence of baseline covariates. For the case where treatment initiation is not assigned randomly, they fit a semi-parametric model on the distribution of the assigned treatment initiation time and adopted inverse probability weighting techniques to deal with the missing of treatment initiation time.

In this article, we formulated the treatment initiation time decision problem in a meaningful and practical framework so as to overcome the aforementioned challenges and estimate the OTIR via a value search method. Specifically, we proposed a new value function, which is constructed on a restricted mean residual lifetime, to evaluate the performance of the treatment initiation regime and developed a nonparametric kernel-based estimation method for the value function. The proposed method has three important advantages. First, it does not posit any specific model on how the failure time depends on the treatment initiation time and covariates and is thus more robust than regression-based methods. Second, the obtained estimates are consistent even when treatment initiation...
times are not completely observable, and their distribution is unknown. Third, the estimation procedure allows treatment initiation times to depend on covariates arbitrarily, and thus, this method can be applied to datasets from both clinical trials and observational studies.

In the remainder of this article, Section 2 describes a breast cancer study that motivates the treatment initiation time problem. Section 3 formulates the optimal treatment initiation problem and presents the details of the proposed OTIR estimation method. Section 4 establishes the asymptotic properties of the resulting estimator in the decision rule and its associated value function estimator. In particular, since kernels are included in the estimation procedure, the asymptotic distribution of the estimated value function is nonstandard, which follows a weighted chi-squared distribution. Section 5 evaluates the performance of the proposed method via simulation studies, followed by an application to a breast cancer dataset for further illustration in Section 6. Section 7 provides some concluding remarks. Technical details are provided in the Web Appendices in the Supporting Information.

2 DATA

This research is motivated by a breast cancer dataset linked between the South Carolina Central Cancer Registry (SCCCR) and the South Carolina Revenue and Fiscal Affairs Office (RFA). The SCCCR is a population-based cancer surveillance system that collects, processes, analyzes, and publishes cancer incidence for the state of South Carolina. In the linked dataset, a total of 629 diagnosed breast cancer patients who received breast cancer surgery after the age of 45 years old initiated adjuvant chemotherapy or radiotherapy therapy at a specific time after surgery. Among these patients, 19.9% were in stage 0 (noninvasive breast cancers), 54.2% were in stage 1 (localized only), and 25.9% were in stage 2 (regional by direct extension only), stage 3 (regional lymph nodes involved only), stage 4 (regional by both direct extension and lymph), or stage 7 (distant sites/nodes involved). The patients’ age at surgery ranges from 45 to 62 years, and the initiation time of the adjuvant therapy is observed on all the patients. Figure 1a shows the distribution of adjuvant

![Figure 1](image-url)

**Figure 1** Analysis of the breast cancer dataset, including (a) the distribution of adjuvant therapy initiation times, and the comparison of initiating adjuvant therapy within 80 days since surgery and after 80 days since surgery among (b) all patients, (c) patients in earlier stages, and (d) patients in advanced stages.
therapy initiation time since surgery (in days), where the average duration from breast cancer surgery to adjuvant therapy is 49 days, and the majority of patients (619 of 629 patients) started adjuvant therapy within 24 weeks since surgery. The dataset also includes the time from surgery to death or the loss of follow-up (in days) for each patient, where the censoring rate of survival time is about 94% and the observed survival time ranges from 1110 to 3690 days. Figure 2 shows the Kaplan–Meier estimates of the survival function, where the range of treatment initiation time was marked in shadow. As shown in the plot, all the patients in this dataset initiated adjuvant therapy within a relatively short time interval compared to their survival times.

For patients diagnosed with breast cancer, numerous randomized trials have demonstrated that adjuvant chemotherapy or radiotherapy after the definitive surgeries could decrease the risk of death caused by breast cancer versus those without adjuvant therapy. However, due to the heterogeneity in patients and their responses to the treatment, finding the optimal time from surgery to the initiation of adjuvant therapy is still challenging (Yu et al., 2017; Riba et al., 2018). In Figure 1, we also compared the censored survival times of patients who initiated adjuvant therapy within 80 days since surgery and of patients who initiated adjuvant therapy after 80 days since surgery among all the patients, patients in earlier stages (stages 0 and 1), and patients in advanced stages (stages 2, 3, 4, and 7), respectively. Although on average, delaying the initiation of adjuvant therapy makes no significant differences on the patients’ survival time distribution (as shown in Figure 1b), it is also clear that, among patients in advanced stages, those initiating adjuvant therapy within 80 days tend to survive longer than those initiating adjuvant therapy after 80 days (Figure 1d), whereas the situation is reversed for patients in earlier stages (Figure 1c). Such an observation indicates that a treatment initiation time that is beneficial for some group of patients may not be a good choice for the others, and an individualized treatment initiation regimen that is based on patients’ characteristics is desired.

3 | METHODOLOGY

3.1 | Definitions

Consider a study with \( n \) patients who started treatment at various time points and were followed up until the event of interest or censoring occurs. Let \([0, a_0]\) be the pre-specified time range for initiating the treatment and let \( \tau \) be the maximum follow-up time. In medical practices, \([0, a_0]\) is usually a relatively short time range compared to the whole follow-up time \( \tau \). For example, in the breast cancer data studies, patients are recommended to start the adjuvant therapy within 168 days from the surgery, while their maximum follow-up time could be more than 10 years.

For the \( i \)th patient, \( i = 1, \ldots, n \), let \( X_i \in \mathcal{X} \) denote the \( p \)-dimensional vector of baseline covariates, \( A_i^* \) denote the assigned treatment initiation time, \( T_i \) denote the event time of interest, and \( C_i \leq \tau \) denote the censoring time. Since \( T_i \) could be censored by \( C_i \), define \( \tilde{T}_i = \min(T_i, C_i) \) and \( \Delta_i = I(T_i \leq C_i) \). Moreover, since \( A_i^* \) could be unobserved if the patient does not survive beyond the assigned treatment time, we define \( \hat{A}_i = A_i^* I[A_i^* \leq \tilde{T}_i] + \infty I[A_i^* > \tilde{T}_i] \) as the observed treatment initiation time. Then the observed data consist of \( \{(X_i, A_i, \tilde{T}_i, \Delta_i), i = 1, \ldots, n\} \), which are independent and identically distributed across \( i \).

A treatment initiation regime \( d(x) \) is a deterministic function that maps the value of covariate \( x \in \mathcal{X} \) to a treatment initiation time \( a \in [0, a_0] \). Let \( T^*(a) \) denote the potential survival time of a patient who is assigned to start the treatment at \( a \) and let \( m(a, x) = \int_{a_0}^a p_r[T^*(a) \geq t | T^*(a) \geq a, X = x]dt \) denote the restricted mean residual lifetime of \( T^*(a) \) at \( a_0 \) for patients with the covariate value \( x \). Then we propose to evaluate the performance of a treatment regime \( d \) by the average of restricted mean residual lifetime among all the patients:

\[
V(d) = E[m[d(X), X]] = \int \int_{a_0}^\tau p_r[T^*(d(x)) \geq t | T^*(d(x))] \geq a_0, X = x]dTdF_X(x),
\]
where \( F_X(x) \) is the cumulative distribution function of \( X \). For simplicity, we assume the probability density function of \( X \) also exists and denote it as \( f_X(x) \). Now, given a collection of treatment regimes \( D \) that are of interest, the optimal treatment regime in \( D \) could be defined as

\[
d_{\text{opt}} = \arg \max_{d \in D} V(d). \tag{2}
\]

**Remark 1.** \( T^*(a) \) is defined as the potential outcome if patients were assigned to initiate treatment at \( a \). Under this definition, if a doctor assigns a patient to initiate treatment at time \( a \), then regardless of whether the patient dies before \( a \), \( T^*(a) \) would still be the potential outcome of the already assigned treatment initiation time \( a \). However, it should be noted that, such treatment assignment could only have effects on a patient’s survival time when \( T^*(a) > a \). In other words, if a patient does not survive to the assigned treatment initiation time, then the patient would not actually initiate the treatment, and thus the assigned treatment initiation time \( a \) has no effect on the patient’s survival time \( T^*(a) \).

In the definition of \( V(d) \), we restricted \( T^*[d(x)] > a_0 \), which means we considered the restricted mean residual lifetime instead of restricted mean survival time. Under this restriction, \( T^*[d(x)] > d(x) \) holds for any \( x \) and \( d \), which ensures that all the samples used to evaluate the treatment initiation regime \( d \) did initiate the treatment. It is noted that, for any \( 0 \leq t_0 < a_0 \), if we replace the restriction \( T^*[d(x)] > a_0 \) in \( V(d) \) with \( T^*[d(x)] > t_0 \), there always exists \( (x, d) \) such that \( t_0 < T^*[d(x)] < d(x) \). We will show later that for such \( (t_0, x, d) \), the conditional survival probability \( pr[T^*[d(x)] \geq t \mid T^*[d(x)] \geq a_0, X = x] \), \( t > t_0 \), is hard to estimate, because in observational studies, the treatment assignment time \( a \) for \( T < a \) is usually unobservable.

In the Web Appendix A1 in the Supporting Information, we also provide an example to show that, under a general class of hazard models, the optimal treatment initiation regime \( d_{\text{opt}} \) defined in (2) does select the optimal treatment initiation time that leads to the largest reduction in the patient’s hazard rate after treatment, while the maximizer of other value functions, including those based on restricted mean survival time, fails to select the optimal treatment initiation times.

### 3.2 Estimation procedure

To ensure that the proposed value function can be estimated using observed data, the following assumptions are required: (A1) (consistency assumption) \( T = T^*(A^*) \); (A2) (no unmeasured confounder assumption) \( \{T^*(a), 0 \leq a \leq a_0 \} \perp A^* \mid X \); (A3) (conditionally independent censoring assumptions) \( C \perp T \mid (A^*, X) \) and \( C \perp A^* \mid X \).

Under these assumptions, the value function \( V(d) \) can be written as

\[
V(d) = \int \int_{a_0} f_X(x) \, dt \, dx.
\]

Since \( \hat{T} \geq a_0 \) implies \( A^* = A \), we have

\[
V(d) = \int \int_{a_0} f_X(x) \, dt \, dx.
\]

Here, \( (T, A, X) \) are observable, \( pr\{T \geq t \mid A = d(x), X = x\} \) can be directly estimated by smoothing techniques such as kernel smoothing, local polynomial fitting, and spline methods, and thus \( V(d) \) is estimable. For example, let \( \hat{f}_X(\cdot) \) denote the kernel density estimator of \( f_X(\cdot) \), and let \( \hat{S}_C(t \mid x) \) be an estimator for the conditional survival function \( S_C(t \mid x) = pr\{C \geq t \mid X = x\} \), then a kernel-based estimator of \( V(d) \) could be

\[
\hat{V}(d) = \int \int \frac{\sum_{i=1}^p I(T_i \geq a_0)K_{h_1}[X_i-x]K_{h_2}[g(A_i)-g(d(x))]}{\sum_{i=1}^p I(T_i \geq a_0)K_{h_1}[X_i-x]K_{h_2}[g(A_i)-g(d(x))]} \times \hat{S}_C(t \mid x) \hat{f}_X(x) \, dt \, dx,
\]

where

\[
K_{h_1}(X_i - x) = \prod_{j=1}^p \frac{1}{h_1(j)} K \left\{ \frac{X_i^{(j)} - x^{(j)}}{h_1^{(j)}} \right\},
\]

\[
K_{h_2}[g(A_i) - g(d(x))] = \frac{1}{h_2} K \left\{ \frac{g(A_i) - g(d(x))}{h_2} \right\},
\]

and

\[
pr\{T \geq t \mid A = d(x), X = x\} = \int \int_{a_0} f_X(x) \, dt \, dx.
\]
$K(\cdot)$ is a kernel function, $h_1 \in \mathbb{R}^p$ and $h_2 \in \mathbb{R}$ are kernel bandwidths, $x^{(j)} (i = 1, \ldots, p)$ denotes the $j$th component of any $p$-dimensional vector $x$, and $g$ is a monotonic increasing transform function that maps interval $(0, a_0)$ into real line.

There are several approaches to obtain $S_C(t \mid x)$. For example, we may construct a semi-parametric model on the censoring time $C$ conditional on $X$, and obtain a model-based estimator of $\Pr(C \geq t \mid X = x)$. Alternatively, for more robust estimation, we could use the kernel-based local Kaplan–Meier method (Dabrowska, 1989) which estimates the conditional survival function nonparametrically. In some applications such as clinical studies with satisfactory follow-up, there is no obvious evidence that censoring events are related to treatment or covariates. For such cases, it is reasonable to make an independent censoring assumption (A4): $C \perp (A^*, X)$, and estimate $\Pr(C \geq t \mid X = x) = \Pr(C \geq t)$ by the standard Kaplan–Meier estimator.

For simplicity, from now on we only consider the case where independent censoring assumption (A4) holds and let $S_C(t) = \Pr(C \geq t)$ denote the survival function of the censoring time $C$. Let $\hat{S}_C(t)$ be the Kaplan–Meier estimator for $S_C(t)$. Then, under assumptions (A1)–(A4), the value of treatment regime $d$ can be estimated by

$$
\hat{V}(d) = \int \sum_{i=1}^n W_n(T_i)I(T_i \geq a_0)K_{h_1}(X_i - x)K_{h_2}[g(A_i) - g(d(x))] \\
\hat{f}_X(x)dx,
$$

where

$$
W_n(t) = \int_{a_0}^t \frac{\hat{S}_C(a_0)}{\hat{S}_C(u)}du \\
= \int_{a_0}^t \prod_{a_0 < u \leq t} \left\{ 1 - \frac{\sum_{j=1}^n I(T_j = u, \Delta_j = 0)}{\sum_{j=1}^n I(T_j \geq u)} \right\}^{-1}du,
$$

for all $t \geq a_0$.

Remark 2. Under the independent censoring assumption, the integrand of our value function can be written as $\Pr(T \geq t \mid \hat{T} \geq a_0, A^* = a, X)S_C(a_0)[S_C(t)^{-1}f_X(x)]$. Here, $\Pr(T \geq t \mid \hat{T} \geq a_0, A^* = a, X)$ is a conditional survival function that focuses on the patients who are alive and uncensored at $a_0$. Since $\hat{T} \geq a_0$ implies $T \geq A^*$, all these patients will have observed treatment initiation times $A = A^*$, and thus $\Pr(\hat{T} \geq t \mid \hat{T} \geq a_0, A^* = a, X) = \Pr(\hat{T} \geq t \mid \hat{T} \geq a_0, A = a, X)$ can be directly estimated by the observed data. Therefore, although in practice, $A^*$ could beensored by both $C$ and $T$ in some complicated mechanism, our estimation procedure avoids the possible problems associated with the missingness of $A^*$. On the other hand, since the other part of the integrand $S_C(a_0)[S_C(t)^{-1}f_X(x)]$ does not concern with the unobservable $A^*$, the corresponding estimators $\hat{S}_C(t)$ and $\hat{f}_X(x)$ can be directly obtained based on all the samples. Therefore, the information for the patients who do not survive to $a_0$ will also be utilized in the estimation of our value function.

### 3.3 Optimal treatment initiation regime

Now we search for the optimal treatment initiation regime in a collection of treatment regimes $D$ indexed by finite-dimensional parameters, which is $D = \{d_\beta : d_\beta(x) = \phi(\tilde{x}^T \beta)\}$. Here, $\tilde{x} = (1, x^T)^T$, $\beta$ is a $(p + 1)$-dimensional parameter and $\phi$ is a fixed function mapping $(-\infty, +\infty)$ to $(0, a_0)$. In practice, $\phi$ could be defined as a logistic link function $\phi(u) = a_0 \exp(u)/(1 + \exp(u))$ or a normal link function $\phi(u) = a_0 \Phi(u)$, where $\Phi$ is the cumulative distribution function of the standard normal distribution.

Let $d^\text{opt} = \arg\max_{d_\beta \in D} V(d_\beta)$ denote the optimal treatment regime among $D$ and let

$$
M(\beta) = V(d_\beta) = \int \int_{a_0}^{t_a} \frac{\Pr(T \geq t \mid A = \phi(\tilde{x}^T \beta), X = x)}{\Pr(T \geq t \mid A = \phi(\tilde{x}^T \beta), X = x)} \\
\times \frac{S_C(a_0)}{S_C(t)}f_X(x)dt dx,
$$

denote the value of the treatment initiation regime indexed by $\beta$. Define $\beta^\text{opt} = \arg\max_{\beta \in \mathbb{R}^{p+1}} M(\beta)$, then we have $d^\text{opt}(x) = \phi(x^T \beta^\text{opt})$. Following the estimation procedure in Section 3.2, we can estimate $M(\beta)$ by

$$
M_n(\beta) = \int \frac{\sum_{i=1}^n W_n(T_i)I(T_i \geq a_0)K_{h_1}(X_i - x)K_{h_2}[g(A_i) - g(\phi(\tilde{x}^T \beta))]}{\sum_{i=1}^n I(T_i \geq a_0)K_{h_1}(X_i - x)K_{h_2}[g(A_i) - g(\phi(\tilde{x}^T \beta))]} \\
f_X(x)dx.
$$

Let $\hat{\beta}^\text{opt} = \arg\max_{\beta} M_n(\beta)$ be the estimate for $\beta^\text{opt}$. Then $\hat{d}^\text{opt}(x) = \phi(x^T \hat{\beta}^\text{opt})$ is an estimator for the optimal treatment initiation regime $d^\text{opt}$, and $M_n(\hat{\beta}^\text{opt})$ estimates the value function of the optimal treatment initiation regime $V(d^\text{opt})$. 


4 | ASYMPTOTIC PROPERTIES

This section provides the asymptotic properties of \( \hat{\beta}^{opt} \) and \( M_n(\hat{\beta}^{opt}) \). Their proofs are given in the Web Appendix B in the Supporting Information. For simplicity, we consider the case where \( A^* \) and \( X \) are continuous with joint probability density function \( f_{A^*X}(a, x) \) on \([0, a_0] \times X\). Define \( S_T(t, a, x) = Pr(T \geq t | A^* = a, X = x) \). We assume the following conditions:

C1 There exists a bounded set \( B \) of \( \beta \) satisfying \( \beta^{opt} \in B \) and \( \sup \{ M(\beta) : \beta \in B, ||\beta - \beta^{opt}||_2 \geq \varepsilon \} < M(\beta^{opt}) \) for any \( \varepsilon > 0 \).

C2 There exist two positive constants \( \varepsilon_c \) and \( \varepsilon_t \), such that \( pr(C \geq \tau | X = x) \geq \varepsilon_c \) and \( inf_{a \in [0, a_0]} S_T(a; a, x) \geq \varepsilon_t \) for all \( x \in X \).

C3 For each fixed \( t, S_T(t; a, x) \) and \( f_{A^*X}(a, x) \) are three-time-differentiable functions of \((a, x)\) with partial derivatives uniformly bounded on \( t \in [0, \tau], a \in [0, a_0] \) and \( x \in X \). Link function \( \phi(u) \) is twice-differentiable with derivative \( \phi(u) = d \phi(u)/du \) bounded on \( \mathbb{R} \). Transform function \( g(a) \) is a four times differentiable for \( a \in [0, a_0] \).

C4 \( K(u) \) is a twice-differentiable symmetric kernel function in \( u \) and \( K(u) \) is a second derivative of \( K(u) \). Define \( \kappa_{ij} = \int u^i K(u)^j du \) and \( \kappa_{ij} = \int u^i K(u)^j du \) for any \( i \geq 0 \) and \( j \geq 0 \). Then \( u^i K(u) \to 0 \) as \( |u| \to \infty \), \( \int |K(u)| du < \infty \), \( sup_u K(u) \leq K_{max} < \infty \), \( sup_u K(u) \leq K_{max} < \infty \) and \( \kappa_{1,1}, \kappa_{0,2} \) and \( \kappa_{0,2} \) are bounded.

C5 \( n|h_1|^2 \to \infty, n|h_1|^2 ||h_1||_2^2 + h_2^2 \to 0 \) as \( n \to \infty \). Here \( h_1 = \prod_{i=1}^{s} h_1(i), ||h_1||_2^2 = \sum_{i=1}^{s} ||h_1(i)||^2 \) for \( h_1 = (h_1(1), ..., h_1(p)) \).

C6 Let \( D(\beta) \) denote the Hessian matrix of \( M(\beta) \). There exists a small neighborhood \( N_\varepsilon \) of \( \beta \) such that \( -D(\beta) \) is positive definite for any \( \beta \in N_\varepsilon \).

Condition C1 is proposed to prove the strong consistency for \( \beta^{opt} \). For a compact set \( B \) and continuous function \( M(\beta) \), the uniqueness of \( \beta^{opt} \) as a maximizer of \( M(\beta) \) also implies this condition. In Condition C2, the positivity of \( S_C(\tau) \) is commonly assumed in survival analysis to ensure the uniform convergence of \( S_C(t) \) on \( t \in [0, \tau] \). Also, Condition C2 implies \( inf_{u \in [0, a_0]} Pr(A^* \leq T, A^* \leq C | A^* = a, X) > 0 \) almost surely, which indicates any candidate treatment initiation time on \([0, a_0] \) has a potential chance to be observed in our setting. The boundedness of derivatives in Condition C3 is postulated for the uniform convergence of \( M_n(\beta) \) and its derivatives on \( B \). In addition, it is not hard to verify that both logistic link function and normal link function satisfy the existence and boundedness assumption for \( \phi \). Conditions C4 and C5 are posited for the asymptotic properties of the Kernel estimators. It is not hard to verify that most of the kernel functions, including Gaussian and all the bounded symmetric kernels, satisfy Condition C4. Condition C5 also provides a guide to select bandwidth \( h_1 \) and \( h_2 \). For example, when \( p = 1 \) and \( h_1 = h_2 \), an appropriate bandwidth should range from \( n^{-1/4} \) to \( n^{-1/6} \). Condition C3 implies that \( M(\beta) \) is twice-differentiable on \( B \), and thus, the Hessian matrix \( D(\beta) \) defined in Condition C6 exists for any \( \beta \in B \). Moreover, by Condition C6, \( D(\beta) \) is invertible in a neighborhood of \( \beta^{opt} \), which is needed for deriving the asymptotic distribution of \( \beta^{opt} \).

Based on the above conditions, we can establish the asymptotic properties of the proposed estimators.

**Theorem 1.** Under Conditions C1–C6, \( sup_{\beta \in \mathbb{R}^{p+1}} |M_n(\beta) - M(\beta)| \) converges to zero almost surely, and \( \beta^{opt} \) is a consistent estimator for \( \beta^{opt} \).

**Theorem 2.** Under Conditions C1–C6, \( (n|h_1|^2)^{1/2}(\beta^{opt} - \beta^{opt}) \) converges in distribution to a normal variable with mean zero and variance \( \Sigma = D^{-1}(\beta^{opt})\Sigma U D^{-1}(\beta^{opt}) \). Here \( D(\beta) \) is a \( p \times p \) matrix defined in Condition C6, \( D^{-1}(\beta) \) denotes the inverse of matrix \( D(\beta) \), and \( U \) is a \( p \times p \) matrix with expression given in the Web Appendix B1 in the Supporting Information.

Let \( det(\Sigma) \) denotes the determinant of matrix \( \Sigma \) and let \( a=(1, ..., s) \) denote the distinct eigenvalues of \( -\Sigma U D^{-1}(\beta^{opt}) \) satisfying \( det[\lambda I + \Sigma U D^{-1}(\beta^{opt})] = \prod_{i=1}^{s} (\lambda - a_i)^{r_i} \) and \( \sum_{i=1}^{s} r_i = p \). Then the asymptotic distribution of the estimated value, \( M_n(\beta^{opt}) \), of the derived optimal treatment initiation regime is stated in the following theorem.

**Theorem 3.** Under Conditions C1–C6, \( n|h_1|^2|M_n(\hat{\beta}^{opt}) - M(\beta^{opt})| \) converges in distribution to \( \sum_{i=1}^{s} a_i \chi^2(r_i)/2 \), where \( \chi^2(r_i), i = 1, ..., s \) are mutually independent chi-square distributions with degree of freedom \( r_i \).

Theorem 3 shows that the asymptotic distribution of the estimated value function is a weighted chi-squared distribution. To illustrate this, we provide some heuristic arguments below. Consider a general case where the M-estimation function \( M_n(\beta) \) and its derivative \( U_n(\beta) = dM_n(\beta)/d\beta \) satisfy \( c_n[M_n(\beta) - M(\beta)] \rightsquigarrow N(0, \Sigma_n(\beta)) \) and \( d_n[U_n(\beta^{opt}) - U(\beta^{opt})] \rightsquigarrow N(0, \Sigma_n^{opt}(\beta)) \). Since \( \beta^{opt} \) is the maximizer of \( M_n \), it can be obtained from Taylor expansion that \( d_n^2(M_n(\beta^{opt}) - M_n(\beta^{opt})) \) converges to a weighted chi-square distribution. Then by
rewriting

\[ M_n(\hat{\beta}_{opt}) - M(\beta_{opt}) = \{M_n(\hat{\beta}_{opt}) - M(\beta_{opt})\} + \{M_n(\beta_{opt}) - M(\beta_{opt})\}, \] (11)

we can conclude that if \( d_n^2/c_n \to 0 \), \( c_n[M_n(\hat{\beta}_{opt}) - M(\beta_{opt})] \) converges to a normal distribution; while if \( d_n^2/c_n \to \infty \), \( d_n^2[M_n(\hat{\beta}_{opt}) - M(\beta_{opt})] \) converges to a weighted chi-square distribution. In our estimation, since treatment initiation time \( A \) follows a continuous distribution and \( M_n(\hat{\beta}) \) contains kernel term \( K_{h_1}[g(A_t) - g(\phi(x^T\beta))] \), this leads to \( d_n = (n|h_1|h_2)^{1/2} \), \( c_n = (n|h_1|h_2)^{1/2} \), and \( d_n^2/c_n = (n|h_1|h_2)^{1/2} \to 0 \) as \( n \to \infty \) under condition 5. Thus, the asymptotic distribution of \( M_n(\hat{\beta}_{opt}) \) is a weighted chi-square distribution. In contrast, if \( A \) follows a discrete distribution, \( M_n(\hat{\beta}) \) will not include the kernel term \( K_{h_1}[g(A_t) - g(\phi(x^T\beta))] \) and \( c_n = d_n \). In that case, \( M_n(\hat{\beta}_{opt}) \) would converge to a normal distribution as \( d_n^2/c_n = d_n \to \infty \), as widely studied in the literature for the value search estimators (Zhang et al., 2012; Fan et al., 2017; Jiang et al., 2017).

In addition, since the analytic forms of the asymptotic variances of the parameter and value estimators are too complicated due to various kernel estimations, direct estimation of these asymptotic variances in finite samples is difficult. Thus, in this article, we use the bootstrap method to obtain the variance estimators. In particular, the confidence interval for the value function is constructed based on the empirical distribution of bootstrapped estimators, while for \( \hat{\beta}_{opt} \), a normal-based confidence interval is adopted.

5 SIMULATIONS

Now we conduct simulations to assess the finite-sample performance of the proposed method. Let covariate \( X = (X_1, X_2)^T \) be a two-dimensional random vector with \( X_1 \) generated from a discrete Bernoulli distribution with mean 0.5 and \( X_2 \) generated from a continuous normal distribution with mean 0 and variance 1. Let \( A^* \in [0, a_0] \) be the assigned treatment initiation time which may depend on \( X \). Given covariate value \( x = (x_1, x_2)^T \) and \( A^* = a_1 \), the failure event \( T \) is generated by one of the following hazard models:

- (m1) \( \lambda(t; a, x) = \lambda_0 \exp[I(t \geq a)Q(1 - \Phi(\phi(x^T\beta_0))] \).
- (m2) \( \lambda(t; a, x) = \lambda_0 \exp[I(t \geq a)Q(1 + x_1^2)/2 + I(t \geq a)Q(1 - \Phi(\phi(x^T\beta_0))] \).
- (m3) \( \lambda(t; a, x) = \lambda_0 \exp[I(t \geq a)Q(1 + x_1^2)/2 + I(t \geq a)Q(1 - \Phi(\phi(x^T\beta_0))] \).

The properties of these hazard models have been discussed in the Web Appendix A1 in the Supporting Information. Here, model (m1) is the basic case. Model (m2) allows the effect of the optimal treatment initiation \( Q(0)(1 + x_1) \) to depend on covariates, which indicates that patients are heterogeneous even when the optimal treatment initiation time is adopted. Under model (m3), the hazard rate before initiating the treatment is also a function of covariates \( X \).

For all these three models, we set \( Q(u) = 2(u^2 - 1) \), \( \beta_0 = (0.5, 0.5)^T \), \( \lambda_0 = 0.2 \) or 0.3, and \( a_0 = 2 \) or 3. Let \( U(a, b) \) denote a uniform distribution on \([a, b]\) and let \( B(a, b) \) denote a beta distribution with mean 1/(a + b) and variance \( ab/((a + b)^2(a + b + 1)) \). We consider the following two scenarios for the treatment initiation time distribution:

- (a1) independent case: \( A^* \sim U(0, a_0) \);
- (a2) dependent case: \( A^* = a_0B_1I(X_1 + X_2 < 0) + a_0B_2I(X_1 + X_2 \geq 0) \) with \( B_1 \sim B(1, 2) \) and \( B_2 \sim B(2, 1) \).

Lastly, let \( C = \min(C^*, \tau) \) be the censoring time for \( T \), where \( \tau = 30 \) is the endpoint of the study and \( C^* \) is generated from uniform distribution \( U(0, 100) \). Two sample sizes \( n = 600 \) and 1000 are considered.

In total, we generated the data in 36 scenarios. Among these scenarios, the censoring rate of event time \( T \) ranges from 10.5% to 44.2%, and the observable rate of treatment initiation time \( A^* \) (i.e., \( A^* \leq T \) and \( A^* \leq C \)) ranges from 61.7% to 83.4%. For each scenario, we applied the proposed method to estimate \( \hat{\beta}_{opt} = (\hat{\beta}_1, \hat{\beta}_2) \) and \( V_0 = V(d_{opt}) = M(\beta_{opt}) \), and calculated the variance of \( \hat{\beta}_{opt} \) and \( M_{sd}(\beta_{opt}) \) by bootstrapped samples. In our implementation, we take monotonic transform function \( g(u) \) = \( -\Phi^{-1}(u/a_0) \), where \( \Phi \) denotes the cumulative distribution function of the standard normal distribution. Since \( X_1 \) is a binary covariate, we use the indicator function instead of a kernel for stratification. For continuous variables, \( X_2 \) and \( A \), a Gaussian kernel is used, and the bandwidth is selected as \( h_1 = \gamma_1 n^{-1/3} \text{sd}(X_2) \) and \( h_2 = \gamma_2 n^{-1/3} \text{sd}(A) \), where \( \text{sd}(X_2) \) is the sample standard deviation of \( X_2 \) and \( \text{sd}(A) \) is the sample standard deviation of the observed treatment initiation time. In our numerical studies, \( \gamma_1 = \gamma_2 = 1 \) generally gives good results for all scenarios. For a better performance, \( (\gamma_1, \gamma_2) \) can also be selected by a cross-validation procedure. For example, we can divide the data into \( K \) equal sized subsamples and consider a finite set of candidate values for \( (\gamma_1, \gamma_2) \). For each pair \( (\gamma_1, \gamma_2) \), let \( d_{opt}^{k} \) (\( 1 \leq k \leq K \)) be the estimated OTIR obtained under \( K - 1 \) subsamples excluding the \( k \)th one, and let \( \hat{V}_k(d_{opt}^{k}) \) be the estimated value of \( d_{opt}^{k} \) where \( \hat{V}_k \) is obtained under the \( k \)th subsample. Then by searching over the candidate values of \( (\gamma_1, \gamma_2) \), we can select the tuning parameter as the
pair maximizing $K^{-1} \sum_{k=1}^{K} \tilde{V}_k(d_k^{opt})$. To do the optimization, we adopted the Nelder–Mead algorithm (Nelder and Mead, 1965), which can be implemented by the R function `optim`. All the initial values are set as zero. Since kernel estimation and bootstrapping could be computationally expensive, the simulations were carried out with an AMD EPYC 7452 32-Core processor, and processing of one dataset with $n = 600$ samples takes about 115–155 s.

Table 1 summarizes the simulation results for $\hat{\beta}^{opt}$ and $\hat{M}_i(\beta^{opt})$ based on 500 replications. For saving space, we only report the results for $n = 600$ with $a_0 = 2$, and those for $n = 1000$ with $a_0 = 3$, and $n = 600$ with $a_0 = 2$ are

| Model | $\lambda_0$ | Estimate | $A^i$ independent with $X$ | $A^i$ dependent with $X$ |
|-------|--------------|----------|----------------|----------------|
|       |              | Bias     | SD   | SE  | CP  | Bias     | SD   | SE  | CP  |
| $m_1$ | 0.2          | $\hat{\beta}_1$ | 0.021 | 0.232 | 0.215 | 0.934 | 0.033 | 0.218 | 0.216 | 0.928 |
|       |              | $\hat{\beta}_2$ | 0.064 | 0.345 | 0.342 | 0.960 | 0.020 | 0.317 | 0.325 | 0.966 |
|       |              | $\hat{\beta}_3$ | 0.014 | 0.180 | 0.186 | 0.970 | $-0.001$ | 0.171 | 0.176 | 0.958 |
|       |              | $V_{\delta}(19.155)$ | $-0.465$ | 0.974 | 0.971 | 0.966 | $-0.667$ | 0.820 | 0.841 | 0.968 |
|       |              | CR: 0.172 | PL: 0.881 | OR: 0.741 | PI: 0.789 | CR: 0.208 | PI: 0.856 | OR: 0.715 | PI: 0.737 |
| $m_2$ | 0.3          | $\hat{\beta}_1$ | 0.033 | 0.238 | 0.223 | 0.922 | 0.033 | 0.221 | 0.221 | 0.932 |
|       |              | $\hat{\beta}_2$ | 0.048 | 0.347 | 0.372 | 0.954 | 0.039 | 0.324 | 0.347 | 0.964 |
|       |              | $\hat{\beta}_3$ | 0.003 | 0.185 | 0.193 | 0.964 | $-0.026$ | 0.168 | 0.186 | 0.968 |
|       |              | $V_{\delta}(16.400)$ | $-0.476$ | 1.060 | 1.094 | 0.948 | $-0.688$ | 0.896 | 0.948 | 0.966 |
|       |              | CR: 0.105 | PI: 0.883 | OR: 0.650 | PI: 0.792 | CR: 0.126 | PI: 0.857 | OR: 0.617 | PI: 0.744 |
| $m_3$ | 0.2          | $\hat{\beta}_1$ | 0.027 | 0.234 | 0.217 | 0.926 | 0.035 | 0.216 | 0.214 | 0.922 |
|       |              | $\hat{\beta}_2$ | 0.057 | 0.367 | 0.345 | 0.938 | 0.036 | 0.313 | 0.331 | 0.950 |
|       |              | $\hat{\beta}_3$ | 0.022 | 0.176 | 0.183 | 0.970 | $-0.006$ | 0.166 | 0.173 | 0.972 |
|       |              | $V_{\delta}(22.432)$ | $-0.228$ | 0.857 | 0.861 | 0.968 | $-0.453$ | 0.725 | 0.745 | 0.964 |
|       |              | CR: 0.245 | PI: 0.880 | OR: 0.741 | PI: 0.787 | CR: 0.296 | PI: 0.856 | OR: 0.715 | PI: 0.734 |
| $m_3$ | 0.3          | $\hat{\beta}_1$ | 0.038 | 0.230 | 0.221 | 0.920 | 0.036 | 0.228 | 0.220 | 0.930 |
|       |              | $\hat{\beta}_2$ | 0.056 | 0.355 | 0.340 | 0.936 | 0.044 | 0.310 | 0.328 | 0.960 |
|       |              | $\hat{\beta}_3$ | 0.015 | 0.180 | 0.184 | 0.966 | $-0.024$ | 0.159 | 0.173 | 0.960 |
|       |              | $V_{\delta}(20.747)$ | $-0.287$ | 0.970 | 0.994 | 0.970 | $-0.540$ | 0.845 | 0.862 | 0.972 |
|       |              | CR: 0.170 | PI: 0.884 | OR: 0.650 | PI: 0.791 | CR: 0.207 | PI: 0.860 | OR: 0.617 | PI: 0.743 |
| $m_3$ | 0.2          | $\hat{\beta}_1$ | 0.013 | 0.241 | 0.233 | 0.948 | 0.025 | 0.231 | 0.231 | 0.956 |
|       |              | $\hat{\beta}_2$ | 0.048 | 0.366 | 0.368 | 0.962 | 0.022 | 0.349 | 0.357 | 0.966 |
|       |              | $\hat{\beta}_3$ | $-0.005$ | 0.187 | 0.208 | 0.970 | $-0.037$ | 0.193 | 0.198 | 0.946 |
|       |              | $V_{\delta}(19.697)$ | $-0.516$ | 0.925 | 0.928 | 0.950 | $-0.682$ | 0.792 | 0.814 | 0.960 |
|       |              | CR: 0.217 | PI: 0.870 | OR: 0.793 | PI: 0.789 | CR: 0.254 | PI: 0.849 | OR: 0.775 | PI: 0.744 |
| $m_3$ | 0.3          | $\hat{\beta}_1$ | 0.022 | 0.243 | 0.229 | 0.924 | 0.030 | 0.233 | 0.233 | 0.952 |
|       |              | $\hat{\beta}_2$ | 0.047 | 0.376 | 0.359 | 0.960 | 0.014 | 0.350 | 0.356 | 0.950 |
|       |              | $\hat{\beta}_3$ | $-0.032$ | 0.214 | 0.217 | 0.960 | $-0.079$ | 0.194 | 0.214 | 0.942 |
|       |              | $V_{\delta}(17.268)$ | $-0.398$ | 1.033 | 1.022 | 0.960 | $-0.624$ | 0.890 | 0.898 | 0.974 |
|       |              | CR: 0.146 | PI: 0.871 | OR: 0.717 | PI: 0.793 | CR: 0.170 | PI: 0.851 | OR: 0.692 | PI: 0.754 |
given the Web Appendix C in the Supporting Information. For each scenario, we report the censoring rate of event time (CR), the observable rate of treatment initiation time (OR), the bias of the estimators (Bias), the standard deviation of the estimators (SD), the mean of estimated standard errors (SE), and the empirical coverage probability of 95% confidence intervals (CP). We also reported the true values of \( V_0 \) for each scenario in the parentheses. From the results, we can see that under all cases, the proposed estimators for \( \hat{\beta}^{opt} \) and \( V_0 \) are nearly unbiased, and the estimated standard errors are close to the standard deviation of the estimators. Moreover, the empirical coverage probabilities of 95% confidence intervals are close to the nominal level for both parameters and value estimators. Both bias and standard deviation of the estimators get smaller when the sample size increases from \( n = 600 \) to \( n = 1000 \) as expected.

To demonstrate the effect of the proposed treatment initiation regimes on individual level, Table 1 also presents the percent of individuals with improved counterfactual outcomes under the OTIR (Pl), and under the optimal constant treatment initiation regime (Pl). The details of the calculation procedure are given in the Web Appendix A2 of Supporting Information. From the results, it can be concluded that around 78.5–89.6% of the individuals would achieve better outcomes if \( \hat{\beta}^{opt} \) had been followed by the entire population, and the OTIR \( \hat{\beta}^{opt} \) does perform better than the optimal constant regime.

6 | Application

Now we apply our method to the breast cancer dataset linked between the SCCCR and the South Carolina RFA and aim to choose the optimal initiation regime of adjuvant therapy for these breast cancer patients. Let \( A \) be the initiation time (in days) of the adjuvant chemotherapy or radiotherapy since surgery and let \( T \) be the patients’ survival time (in days). For the selection of \( a_0 \), since patients are usually recommended to start adjuvant therapy within 24 weeks from the surgery (Lohrisch et al., 2006), we set \( a_0 = 168 \) (in days), and exclude patients who started the adjuvant therapy after 24 weeks from the surgery. As shown in Figure 1, the majority of patients did start adjuvant therapy within 168 days since surgery. Also, since \( pr(C \geq \tau) > 0 \) is required to ensure the uniform convergence of \( \hat{S}_C(t) \) (as discussed in Section 4), we choose \( \tau = 3720 \) such that about 85% of patients’ event times (censoring or failure) are less than \( \tau \). In our analysis, we consider two covariates \( X = (X_1, X_2) \), where \( X_1 \) is the age of the patient at surgery standardized to mean 0 and variance 1, and \( X_2 \) is an indicator of breast cancer stage. Specifically, we define \( X_2 = 0 \) for patients with breast cancer stage 0 or stage 1 (localized only); and define \( X_2 = 1 \) for patients with breast cancer stages 2, 3, 4, or 7. In medical practice, stages 0 and 1 are earlier stages of breast cancer. Thus, we refer \( X_2 = 0 \) as earlier stage and \( X_2 = 1 \) as advanced stage. Since \( X_1 \) and \( A \) are continuous, a Gaussian kernel is used in the estimation procedure, and the kernel bandwidths are selected as \( h_1 = n^{-1/5} sd(X_1) \) and \( h_2 = n^{-1/5} sd(A) \) separately. For a better performance, it would also be worthwhile to develop some data-driven bandwidth selection algorithms based on the empirical bias bandwidth selection (EBBS) method. More details are given in the Web Appendix A3 in the Supporting Information.

Figure 3 ab presents the estimated optimal treatment initiation regime among two classes of decision rules:

\[
D_{\text{logistic}} = \left\{ d_\beta : d_\beta(x) = \frac{a_0 \exp(\hat{x}^T \beta)}{1 + \exp(\hat{x}^T \beta)}, \quad x \in \mathbb{R} \times \{0, 1\}, \beta \in \mathbb{R}^3 \right\},
\]

\[
D_{\text{normal}} = \left\{ d_\beta : d_\beta(x) = a_0 \Phi(\hat{x}^T \beta), \quad x \in \mathbb{R} \times \{0, 1\}, \beta \in \mathbb{R}^3 \right\}.
\]

(12)

It can be seen from the plots that the optimal treatment initiation regimes obtained based on the logistic link function and the normal link function recommend similar treatment initiation times for patients. This shows certain robustness of the proposed method to the choice of link functions in the considered treatment initiation regimes.

For simplicity, from now on we only present analysis results under the logistic link function. By calculation, the estimate of \( \beta \) is \( \hat{\beta}^{opt} = (0.669, -0.154, -1.643) \) with the standard error \( SE = (0.288, 0.163, 0.476) \), and the \( p \)-values are 0.020, 0.348, and 0.001, respectively. Further illustration is presented in Figure 3. Specifically, Figure 3a,b plots the observed treatment initiation times \( A_i \) and the estimated optimal treatment initiation times \( \hat{d}^{opt}(x_i) \) over patients’ age. As shown in the plots, although there is no significant difference in the observed treatment initiation times among patients with different ages, the estimated optimal treatment initiation regime suggests moderate delay in treatment initiation for younger patients. Figure 3c,d compares the distribution of the observed treatment initiation times and the distribution of the estimated optimal treatment initiation times for patients with earlier stages and advanced stages. It can be concluded from the plots that, according to the estimated optimal treatment initiation regime, patients with advanced breast cancer stages \( X_2 = 1 \) should initiate adjuvant therapy earlier than those with earlier breast cancer stages. All these results are consistent with the \( p \)-values.
We have also compared the performance of the estimated optimal treatment initiation regime $\hat{d}_o^{opt}$ with that of the observed treatment initiation regime $d_A(X_i) = A_i$. On the one hand, we let $\hat{V}(d)$ denote the estimated value function under a given regime $d$. Using the proposed kernel estimation method, we can calculate $\hat{V}(\hat{d}_o^{opt}) = 4134.528$ and $\hat{V}(d_A) = 3408.576$. Therefore, the increase in the value function comparing the estimated optimal treatment initiation regime with the observed treatment initiation time is $\hat{V}_{\text{diff}} = \hat{V}(\hat{d}_o^{opt}) - \hat{V}(d_A) = 725.952$, which suggests a nearly 2-year improvement in expected overall restricted survival time for breast cancer patients when patients follow $\hat{d}_o^{opt}$. We can further obtain the empirical distribution of $\hat{V}_{\text{diff}}$ by bootstrapping samples for 500 times. Specifically, for each bootstrapped sample, we calculate $\hat{V}_*^{\hat{d}_o^{opt}}$, and let $\hat{V}_{\text{diff}}^* = \hat{V}_*^{\hat{d}_o^{opt}} - \hat{V}_*^{d_A}$. Then based on the 500 bootstrapped $\hat{V}_{\text{diff}}^*$, we can obtain a quantile-based 95% confidence interval of $\hat{V}_{\text{diff}}$ as $(308.128, 1489.948)$. This again suggests a significant improvement in the overall restricted mean survival time. On the other hand, to compare the performances of $\hat{d}_o^{opt}$ and $d_A$ on individual level, we calculate the percent of individuals with improved counterfactual outcomes under $\hat{d}_o^{opt}$ compared to $d_A$. The obtained PI is 0.798, which indicates that 79.8% of the breast cancer patients would achieve larger restricted mean survival time if $\hat{d}_o^{opt}$ had been followed by all the breast cancer patients.

Lastly, we compare the derived optimal treatment initiation regime with some fixed treatment initiation regimes $d_a(x) \equiv a$ where $a$ is a constant taking value in $[0, a_0]$. Specifically, given a regime $d$, we calculate the value functions $\hat{V}^\gamma(d)$ based on 500 bootstrapped samples, and compare the empirical distributions of $\hat{V}^\gamma(d)$ under $\hat{d}_o^{opt}$, $d_{28}(x) \equiv 28$, $d_{56}(x) \equiv 56$, $d_{84}(x) \equiv 84$, $d_{112}(x) \equiv 112$, and $d_{140}(x) \equiv 140$. For completeness, we also include the results for $d_A$. The results are plotted in Figure 4. Based
on the plot, the values of the estimated optimal treatment initiation regime $d^{opt}$ usually are much larger than those under $d_a$ and $d_A$, which indicates that patients following $d^{opt}$ tend to have better treatment effect than those following constant regime $d_a$ for $a \in [0, a_0]$.

7 | DISCUSSION

In this article, we presented a formulation of treatment initiation time decision problem and proposed a new value search approach to find the optimal individualized treatment initiation time regime for censored time-to-event data. Different from existing value search methods, our value function is constructed on the restricted mean residual lifetime at the endpoint of treatment initiation interval $a_0$. The proposed value function can be estimated consistently even when the treatment initiation times are not completely observable, and their distribution is unknown. As a matter of fact, if the value function is constructed based on the restricted mean survival time or restricted mean residual lifetime at some time point $t < a_0$, the estimation of the value function requires additional assumptions. More details are given in the Web Appendix A4 in the Supporting Information.

In the estimation procedure, we focus on patients whose survival times are longer than the maximum treatment initiation time $a_0$. Such a value function may cause some selection bias when there is a proportion of patients in critical conditions with potentially short life expectancy. To deal with this issue, we propose a refined two-step estimation procedure, by first identifying a subset of patients who may be in critical conditions and need to start the treatment early on based on their estimated optimal treatment initiation time obtained in the first step. More details are given in the Web Appendix A5 in the Supporting Information.

For simplicity, this paper only considered the covariate-independent censoring case (A4). For the case with the conditionally independent censoring given in assumption (A3), we may construct a Cox model on the censoring time $C$ conditional on $X$, or we can estimate the conditional survival function of censoring times by kernel conditional Kaplan–Meier estimator (Dabrowska, 1989). The associated asymptotic properties of the proposed estimators can also be derived but will be more involved. Also, since the kernel conditional Kaplan–Meier estimator needs $n_{h_i}^{p+4} \rightarrow 0$ to obtain the convergence rate $\{nh_i^{p+4}/(\log h_i)\}^{1/2}$ (Dabrowska, 1989), the convergence rate of the OTIR estimator could be slower especially when the dimension of covariates $p$ is high. Moreover, it is worthwhile considering the cases where the censoring time is also affected by the assigned treatment initiation time. For such cases, if the dependence between the censoring time and the assignment treatment initiation time can be fully captured by the observed covariates, (i.e., $C \perp A^*|X$), then our method is still valid. Otherwise, finding the optimal treatment initiation regime would be an open problem that warrants future research.

ACKNOWLEDGMENTS

The authors would like to thank the Co-Editor, an associate editor, and two referees for their insightful comments and suggestions that greatly improved the article. This work was partly supported by grant NIH P01 CA142538, NSF-DMS-1555244, the National Natural Science Foundation of China (11771431, 11690015 and 12001371), the Key Laboratory of RCSDS, CAS (No. 2008DP173182), and Shanghai Sailing Program (20YF1434200).

DATA AVAILABILITY STATEMENT

The data that support the findings in this paper are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Xin Chen https://orcid.org/0000-0002-3815-1221
Jiajia Zhang https://orcid.org/0000-0003-4566-0822
Liuquan Sun https://orcid.org/0000-0002-8816-942X
Wenbin Lu https://orcid.org/0000-0002-7320-4755

REFERENCES

Blatt, D., Murphy, S. and Zhu, J. (2004) A-learning for approximate planning. Ann Arbor, 1001, 48109–2122.
Dabrowska, D.M. (1989) Uniform consistency of the kernel conditional Kaplan–Meier estimate. Annals of Statistics, 17, 1157–1167.
Fan, C., Lu, W., Song, R. and Zhou, Y. (2017) Concordance-assisted learning for estimating optimal individualized treatment regimes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79, 1565–1582.
Havlir, D.V., Kendall, M.A., Ive, P., Kumwenda, J., Swindells, S., Qasba, S.S. et al. (2011) Timing of antiretroviral therapy for HIV-1 infection and tuberculosis. *New England Journal of Medicine*, 365, 1482–1491.

Hu, L., Hogan, J.W., Mwangi, A.W. and Siika, A. (2018) Modeling the causal effect of treatment initiation time on survival: Application to HIV/TB co-infection. *Biometrics*, 74, 703–713.

Jiang, R., Lu, W., Song, R. and Davidian, M. (2017) On estimation of optimal treatment regimes for maximizing t-year survival probability. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79, 1165–1185.

Jung, H., Cho, J.Y. and Lee, Y. (2019) Progression of aortic regurgitation after subarterial ventricular septal defect repair: optimal timing of the operation. *Pediatric Cardiology*, 40, 1696–1702.

Karim, S.S.A., Naidoo, K., Grobler, A., Padayatchi, N., Baxter, C., Gray, A. et al. (2010) Timing of initiation of antiretroviral drugs during tuberculosis therapy. *New England Journal of Medicine*, 362, 697–706.

Lohrisch, C., Paltiel, C., Gelmon, K., Speers, C., Taylor, S., Barnett, J. and Olivotto, I.A. (2006) Impact on survival of time from definitive surgery to initiation of adjuvant chemotherapy for early-stage breast cancer. *Journal of Clinical Oncology*, 24, 4888–4894.

Lok, J.J. and DeGruttola, V. (2012) Impact of time to start treatment following infection with application to initiating HAART in HIV-positive patients. *Biometrics*, 68, 745–754.

Murphy, S.A. (2003) Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65, 331–355.

Nelder, J.A. and Mead, R. (1965) A simplex method for function minimization. *The Computer Journal*, 7, 308–313.

Qian, M. and Murphy, S.A. (2011) Performance guarantees for individualized treatment rules. *Annals of Statistics*, 39, 1180–1210.

Riba, L.A., Gruner, R.A., Fleishman, A. and James, T.A. (2018) Surgical risk factors for the delayed initiation of adjuvant chemotherapy in breast cancer. *Annals of Surgical Oncology*, 25, 1904–1911.

Sugano, H. and Arai, H. (2015) Epilepsy surgery for pediatric epilepsy: optimal timing of surgical intervention. *Neurologia Medico-Chirurgica*, 55, 399–406.

Watkins, C.J. and Dayan, P. (1992) Q-learning. *Machine Learning*, 8, 279–292.

Yang, C.H., Chen, K.J., Tsai, J.J., Lin, Y.H., Cheng, S.H., Wang, K.F. and Chiou, H.Y. (2014) The impact of HAART initiation timing on HIV-TB co-infected patients, a retrospective cohort study. *BMC Infectious Diseases*, 14, 304.

Yu, K.D., Fan, L., Qiu, L.X., Ling, H., Jiang, Y.Z. and Shao, Z.M. (2017) Influence of delayed initiation of adjuvant chemotherapy on breast cancer survival is subtype-dependent. *Oncotarget*, 8, 46549–46556.

Zhang, B., Tsiatis, A.A., Laber, E.B. and Davidian, M. (2012) A robust method for estimating optimal treatment regimes. *Biometrics*, 68, 1010–1018.

Zhao, Y., Zeng, D., Rush, A.J. and Kosorok, M.R. (2012) Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107, 1106–1118.

Zhao, Y., Zeng, D., Socinski, M.A. and Kosorok, M.R. (2011) Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, 67, 1422–1433.

**SUPPORTING INFORMATION**

Web Appendices referenced in Sections 3-7 are available with this paper at the Biometrics website on Wiley Online Library. The R code implementing our method is available from this website as well.

**How to cite this article:** Chen, X., Song, R., Zhang, J., Adams, S.A., Sun, L., Lu, W. On estimating optimal regime for treatment initiation time based on restricted mean residual lifetime. *Biometrics*. 2022;78:1377–1389. https://doi.org/10.1111/biom.13530