Selection of the optimal personalized treatment from multiple treatments with right-censored multivariate outcome measures

Chathura Siriwardhana a, K.B. Kulasekera b and Somnath Datta c

aDepartment of Quantitative Health Sciences, University of Hawaii John A. Burns School of Medicine, Honolulu, HI, USA; bDepartment of Bioinformatics & Biostatistics, University of Louisville, Louisville, KY, USA; cDepartment of Biostatistics, University of Florida, Gainesville, FL, USA

ABSTRACT
We propose a novel personalized concept for the optimal treatment selection for a situation where the response is a multivariate vector that could contain right-censored variables such as survival time. The proposed method can be applied with any number of treatments and outcome variables, under a broad set of models. Following a working semiparametric Single Index Model that relates covariates and responses, we first define a patient-specific composite score, constructed from individual covariates. We then estimate conditional means of each response, given the patient score, correspond to each treatment, using a nonparametric smooth estimator. Next, a rank aggregation technique is applied to estimate an ordering of treatments based on ranked lists of treatment performance measures given by conditional means. We handle the right-censored data by incorporating the inverse probability of censoring weighting to the corresponding estimators. An empirical study illustrates the performance of the proposed method in finite sample problems. To show the applicability of the proposed procedure for real data, we also present a data analysis using HIV clinical trial data, that contained a right-censored survival event as one of the endpoints.

1. Introduction
The goal of personalized medicine is to use data to improve decision making in health care to provide the ‘best’ outcome for a patient based on his/her individualized features. The principle of personalized medicine is centuries old, but the idea advanced dramatically after the introduction of randomized controlled clinical trials. The primary aim in a majority of clinical trials is to make only a population-level decision but not an individualized decision that accounts for patient heterogeneity. But the increasing availability of data from such studies has increased the awareness of heterogeneity in both patient

CONTACT Chathura Siriwardhana cksiri@hawaii.edu.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02664763.2022.2164759.

© 2023 Informa UK Limited, trading as Taylor & Francis Group
characteristics and outcomes and lead to new evidence-based medicine concepts. Over the last two decades, the statistical methodology in personalized medicine has led to new methodologies and insights, owing to advancement in computational power, bioinformatics discoveries, and access to electronic health data [16,20,28,36,37]. The basic premise of many existing personalized strategies is to make data-driven decisions to optimize a targeted patient outcome while looking at the patient profile. In real-life situations, the success of treatment does not fully reflect through a single outcome as a variety of factors may compel both patients and clinicians to consider recovery in a rather broad view. For example, in the treatment for type-2 diabetes, control in HbA1c, systolic blood pressure, low-density lipoproteins, cholesterol levels, and prevention from hypoglycemia and weight gain have been suggested as therapeutic goals to address the net clinical utility of a treatment [35]. In cancer studies, although the overall survival is considered the most important in this setting, a variety of other factors such as reduction in tumor size or eradication of cancerous cells, considered to be meaningful outcomes [14]. Similarly, in other situations where the disease is a life-threatening condition, time-to-event outcomes (e.g. overall survival) are commonly considered the best outcome, but there could be other factors with secondary importance, such as those relating to the quality of life and economic impact. Also, it is common to use a collection of surrogate outcomes during the early development of treatments [7].

This work is a culmination of our previous contributions in the area of personalized medicine, more specifically, optimal treatment selection based on patient characteristics measured through covariates. We consider a situation where randomized trial data are available on patients where a number of responses (or a multi-dimensional response) were observed per patients. In addition, one or more of these responses are time to event and therefore subject to right censoring. This is the main novel feature of the methodology presented here and extends the papers by Siriwardhana et al. [28,29], where only completely observed responses were observed.

This paper was motivated in part by a clinical trial data set on HIV patients which was only partially analyzed by the methodology available thus far in earlier articles by Siriwardhana et al. [28,29]. Amongst other clinical characteristics such as the CD4 and CD8 counts, disease-free survival times were also observed which were subject to right censoring. However, the survival information was not directly implementable in the optimal treatment selection procedures because they were right-censored. We show in this work how to adjust the procedures in Siriwardhana et al. [29] to incorporate right-censored training data. The HIV data set is reanalyzed based on a trivariate response leading to better predictions of optimal treatment selection.

The remainder of this paper is organized as follows: in Section 2, we discuss the proposed methodology. Section 3 includes simulation results, followed by real-life data illustration in Section 4. Finally, the main body of this paper ends with a discussion in Section 5.

2. Treatment selection

In this section, we describe the proposed personalized treatment selection procedure. Consider a situation where we observe $J$ number of response variables for a patient undergoing
a treatment selected from \( K \) possible treatment options and, with no loss of generality, consider that larger values of individual responses are indicative of better outcomes. We assume that there is at least one response among \( J \) responses is a survival type outcome that can be subjected to right censoring. Suppose \( \mathbf{Y}^*_k = (Y^*_1, \ldots, Y^*_J)' \) provides the \( J \) dimensional vector of responses under the \( k \)th treatment option, for \( r \)-dimensional covariate vector \( \mathbf{X} \). In this work, we assume that we have data from a randomized clinical trial (RCT) study that provides responses and covariate information of a set of patients randomized into \( K \) arms. It is important to note that a dataset resulting from an RCT trial does not have independent pairs of observations \( (\mathbf{Y}_k, \mathbf{X}_k) \) to obtain an estimator of the conditional mean. Following Siriwardhana \cite{21,22}, we use the semi-parametric Single Index Model (SIM) as the working model for relating outcomes and covariates. This model provides a great flexibility in handling unknown nonlinearity between a response and a set of covariates. However, as SIM models could be different from actual mean models in real-life problems, we use the SIM model to obtain the first approximation of the conditional mean given a covariate. Then we produce a working model-based low dimensional score to use at the next level of smoothing via a fully nonparametric approach, to obtain an estimator for the conditional mean. Following Siriwardhana et al. \cite{28,29}, we use a patient’s covariate value \( \mathbf{X} \) to obtain a lower dimensional composite patient score \( U(\mathbf{X}) \) that summarizes each patient’s characteristics.

Following the ideas given in previous works, we use a patient’s covariate value \( \mathbf{X} \) to obtain a lower dimensional composite patient score \( U(\mathbf{X}) \) that summarizes each patient’s characteristics \cite{3,28,29}.

Here, we consider pairs of independent observations \( (\mathbf{Y}_k, \mathbf{X}_k) \) from the marginal distribution of \( (\mathbf{Y}^*_k, \mathbf{X}) \), \( k = 1, \ldots, K \) to select the optimum treatment for \( K \) treatments using the vectors of smoothed conditional means for each treatment. We define

\[
\mu_{jk}(u_j) = E[Y_{jk}|U_j(X_k) = u_j]; \quad j = 1, \ldots, J; \quad k = 1, \ldots, K
\]  

and vectors \( \mu_k(\mathbf{u}) = (\mu_{1k}(u_1), \ldots, \mu_{Jk}(u_J))' \) for \( \mathbf{u} = (u_1, \ldots, u_J)' \) where components of these vectors correspond to each response.

We rank the \( K \) values for each component of \( \mu_k(\mathbf{u}) = (\mu_{1k}(u_1), \ldots, \mu_{Jk}(u_J))' \) vectors \( (k = 1, \ldots, K) \) to get size \( K \) vectors \( \mathbf{v}_j(\mathbf{u}) = (v_{j1}(\mathbf{u}), \ldots, v_{jK}(\mathbf{u}))' \), where \( v_{jk}(\mathbf{u}) \) is the rank of \( \mu_{jk} \) \( k = 1, \ldots, K \) among \( \mu_j(\mathbf{u}) = (\mu_{j1}(u_1), \ldots, \mu_{jk}(u_j))' \) for each \( j \) (here \( j = 1, \ldots, J \)) with the largest \( \mu_{jk}(u_j) \) value given the rank 1. Then we use an aggregation method given by Pihur et al. \cite{21,22} to combine these rank vectors to get an overall ranking of treatments \( \mathbf{v}^* = (v^*_1, \ldots, v^*_K) \).

For a set of weights \( w_j, j = 1, \ldots, J \), selected appropriately, and for a distance measure \( \gamma \) \cite{21}, we minimize a quantity

\[
\psi(\mathbf{v}) = \sum_{j=1}^{J} w_j \gamma(\mathbf{v}, \mathbf{v}_j(\mathbf{u}))
\]
over $v \in P_K$, the set of all permutations of $\{1, \ldots, K\}$. Among possible distance measures for $\gamma$, we used the weighted Spearman’s Footrule distance \[22\], given by

$$
\sum_{l=1}^{I} |v_{1l} - v_{l1}| |M_{vl} - M_{vrl}|^\rho,
$$

where $M_1, \ldots, M_I$ is a list of real values, $r_1, \ldots, r_I$ are the ranks of $M_l, l = 1, \ldots, I$ and $v_1, \ldots, v_I$ is a permutation of integers $1, \ldots, I$ and $\rho$ is a positive number.

The above minimization yields a vector $v^* = (v^*_1, \ldots, v^*_K)'$ where

$$
v^* = \arg\min_{v \in P_K} \psi(v).
$$

In our approach, for a covariate $X_0$ with a score $U_0 = U(X_0)$, we define the optimal treatment as

$$
k^*(U_0) = \arg\min_{1 \leq k \leq K} \{v^*_k\}.
$$

Note that, in the unlikely event of observing multiple minimizers in 3 or 4, we randomly select one among those with ties.

We link the $j$th component $Y_{jk}$ of the response vector $Y_k$ for the $k$th treatment and covariates $X_k$ via a SIM model,

$$
Y_{jk} = g_{jk}(\beta_{jk}'X_k) + \epsilon_{jk}
$$

for $j = 1, \ldots, J$ and $k = 1, \ldots, K$, where each $\beta_{jk}$ is a $r$-vector of parameters, $g_{jk}$ is an unknown link function for which we assume some reasonable smoothness conditions to hold, and $\epsilon_{jk}$ are error terms with $E[\epsilon_{jk}|X] = 0$. Furthermore we assume independence of $\epsilon_{jk}$s across $k = 1, \ldots, K$ for a fixed $j$ where these terms are correlated across $js$ for any given $k$. The SIM formulation provides flexibility and reasonable efficiency in modeling many types of data.

As we described in the beginning of this section, we may observe right-censored outcomes for time-to-event type responses while using data sources such as historical clinical trial records. In our discussion below, we initially focus on explaining some aspects of the proposed strategy for completely observed responses, summarizing the procedure by Siriwardhana et al. [29]. Later, we generalize these ideas for problems with right-censored outcomes.

2.1. Treatment selection with the complete (uncensored) data

When complete response (i.e. uncensored) data available, observations are of the form $(Y_{ki}, X_{ki})$ where $Y_{ki} = (Y_{1ki}, \ldots, Y_{Jki})'$ and $Y_{jki}$ indicates the $j$th component of the $i$th individual under treatment $k$ with associated covariate values $X_{ki}, i = 1, \ldots, n_k$. Then, for this
data, relationship (5) can be written as
\[ Y_{jki} = g_{jk}(\beta'_{jk}X_{ki}) + \epsilon_{jki}, \quad i = 1, \ldots, n_k. \]  
(6)

Siriwardhana et al. [29] define a score vector \( U(X) \) for a patient with \( X \) covariate as follows. First, we define
\[ S_{jk}(X) = g_{jk}(\beta'_{jk}X) - \max_{l \neq k} \{ g_{jl}(\beta'_{jl}X) \}. \]

Next, define the \( j \)th components of the combined overall score vectors as
\[ S_j(X) = \max_k \{ S_{jk}(X) \} \]
\[ \delta_j(X) = \arg\max_k \{ S_{jk}(X) \}. \]  
(7)

The overall score is given as \( U(X) = (U_1(X), \ldots, U_J(X))' \) where \( U_j(X) = (S_j(X), \delta_j(X))' \) for \( j = 1, \ldots, J \). Note that, \( \delta \) will be selected randomly among ties, in the improbable event that multiple treatments produce the largest \( S_{jk}(X) \).

However, in reality, error distributions and model functions for models defined in (5) are unknown. Therefore standard function estimation method as described in Siriwardhana et al. [28,29] should be applied to estimate components of these score vectors.

In particular, for any given vector \( x \), let
\[ \hat{S}_{jk}(x) = \hat{g}_{jk}(\hat{\beta}'_{jk}x) - \max_{l \neq k} \{ \hat{g}_{jl}(\hat{\beta}'_{jl}x) \} \]
\[ \hat{S}_j(x) = \max_k \{ \hat{S}_{jk}(x) \} \]
\[ \hat{\delta}_j(x) = \arg\max_k \{ \hat{S}_{jk}(x) \} \]
and
\[ \hat{U}_j(x) = (\hat{S}_j(x), \hat{\delta}_j(x))'; \quad j = 1, \ldots, J. \]  
(8)

Siriwardhana et al. [29] proposed estimators for \( \mu_{jk}(u_j), \quad k = 1, \ldots, K \) at a given \( u_j = (s_j, d_j)' \) as follows. Here \( \omega \) is a kernel function with \( \omega \geq 0 \) and \( \int \omega(t) \, dt = 1 \), and \( h_k, \quad k = 1, \ldots, K \) are a set of smoothing parameters.

\[ \hat{\mu}_{jk}(u_j) = \frac{\sum_{i=1}^{n_k} Y_{jki} \omega \left( (s_j - \hat{S}_j(X_{ki}))/h_{jk} \right) I \left( \hat{\delta}_j(X_{ki}) = d_j \right)}{\sum_{i=1}^{n_k} \omega \left( (s_j - \hat{S}_j(X_{ki}))/h_{jk} \right) I \left( \hat{\delta}_j(X_{ki}) = d_j \right)}, \]  
(9)

where \( I(A) \) is the indicator of \( A \). The bandwidth selection for estimating \( \mu_{jk}s \) is a challenging issue and we do not investigate the optimal bandwidth selection for this problem. However, as Siriwardhana et al. [29] suggested, the method given in Wand and Jones [38] for kernel smoothing provides a reasonable solution for this estimation problem. We adopt this method for the current work.

For a realization \( x_0 \) of the covariate \( X \), if one could find the corresponding realizations of the scores, \( u_{j0} = (S_j(x_0), \delta_j(x_0))' \), this allows estimating \( \mu_{jk}(u_{j0}) \) by \( \hat{\mu}_{jk}(u_{j0}) \). However, due
to aforementioned reasons, one may only find an estimate \( \hat{u}_{j0} = (\hat{s}_{j0}, \hat{\delta}_{j0})' \) of \( u_{j0} \) using (8) above. Thus in practice one may use \( \hat{\mu}_{jk}(\hat{u}_{j0}) \) as the estimate of \( \mu_{jk}(u_{j0}) \) for \( j = 1, \ldots, J; k = 1, \ldots, K \), with \( \hat{u}_0 = (\hat{u}_{10}, \ldots, \hat{u}_{J0})' \).

### 2.2. Treatment selection with right-censored data

In this section, we extend the treatment selection concept by Siriwardhana et al. [29] problems that are involved with the right-censored survival type data. When complete (i.e. uncensored) data are available, we estimate each component of the overall score \( U(X) = (U_1(X), \ldots, U_J(X))' \) via a set of estimated single index models or partial linear models, given by (5). Subsequently, we obtain an estimator for \( \mu_{jk}(\hat{u}_j), j = 1, \ldots, J, k = 1, \ldots, K \), by (9). Suppose there exists at least one response among \( J \) that subject to right-censoring. Let \( j, j = 1, \ldots, J, \) be such a time-to-event type right-censored outcome. Since some \( Y_{jki}'s, i = 1, \ldots, n_k \), are now unobservable, the direct application of the previous technique is no longer possible.

The use of data weighting schemes for the purpose of bias reduction is well known in the statistical literature. In statistical literature, data weighting schemes are popularly applied for bias reduction in estimates. A common strategy to handle censored observations in regression setting is to introduce a re-weighing scheme such as the Inverse Probability Censoring Weighting (IPCW) to the original estimator developed for complete data, in a way that the bias caused due to censoring fades away asymptotically. The basics of this concept was first introduced by Koul et al. [17] for the randomly right censored data in linear regression and later extended by many other authors for various problems with censoring. For example, Satten and Datta [23] described estimating the marginal survival time in the presence of time-dependent covariates, using a re-weighted Kaplan–Meier estimate based on the IPCW weights calculated from Aalen’s additive hazard model [1,2]. Recently, Siriwardhana et al. [27] described estimating the binary choice single index model incorporating IPCW for censored data. Following a similar approach, we introduce a reweighing scheme to estimators of the SIM model and \( \mu_{jk}(u_j) \).

Let \( C_{jki} \) be the right-censoring time correspond to the \( j \)th time-to-event response, \( j = 1 \cdots, J, \) for the subject \( i, i = 1, \ldots, n_k, \) who treated from the \( k \)th treatment group, \( k = 1, \ldots, K, \) and let \( Y_{jki} \) is the true response of the individual. Thus we observe the censored response \( T_{jki} = \min\{Y_{jki}, C_{jki}\} \) with the censoring indicator \( \delta_{jki}' = I[Y_{jki} \leq C_{jki}] \). We denote the survival function of the censoring distribution by \( K_{jki}^c(t) = E[I[C_{jki} > t]] \), with hazard function \( \lambda_{jki}^c(t) \) and cumulative hazard function \( \Lambda_{jki}^c(t) \), at a time point \( t \).

### 2.3. IPCW re-weighted single index estimator

We use an IPCW re-weighted Ichimura et al. [11] SIM model to estimate the model (5), which allows one to calculate the score \( \hat{U}_j(X) = (\hat{S}_j(X), \hat{\delta}_j(X)) \) with respect to a right-censored outcome \( j, j = 1, \ldots, J, \) by \( k = 1, \ldots, K \) [25]. The covariate response relationship for the \( j \)th outcome by \( k \)th treatment can be represented as

\[
Y_{jki} = f_{jk}(\theta_{jk}'X_{ki}) + \epsilon_{jki}, \quad i = 1, \ldots, n_k,
\]
for some function $f_{jk}(\cdot)$ and parameter vector $\theta_{jk}$ ($\theta_{jk} \in \mathcal{R}^r$), where $\epsilon_{jki}$'s are independent errors with $E(\epsilon_{jki}|X_{ki}) = 0$ and bounded common variance $\sigma$. For the purpose of identifiability, we replace $\theta_{jk}$ by a unit vector,

$$\beta_{jk} = \theta_{jk} \| \theta_{jk} \|^{-1},$$

where $\| \cdot \|$ is the Euclidean norm. Thus an equivalent model can be written as

$$Y_{jki} = g_{jk}(\beta_{jki}'X_{ki}) + \epsilon_{jki},$$

which has the same form of model (5), where $g_{jk}(\cdot)$ is a unknown univariate smooth link function. For the notational convenience, we suppress the treatment indicator $k$ and the response indicator $j$ to describe the IPCW re-weighted SIM estimation procedure.

Assuming all $Y$s are completely observed, Ichimura et al. [11] proposed an estimator to estimate the above SIM model that estimate the unknown function $g(\cdot)$ at point $\nu$, by leave-one-out cross validation method, omitting the pair of $(Y_i, X_i)$,

$$\hat{g}_{-i}(\nu|\beta) = \frac{\sum_{l \neq i} Y_l \phi_{\tilde{h}}^i (\nu - \beta' X_i)}{\sum_{l \neq i} \phi_{\tilde{h}}^i (\nu - \beta' X_i)},$$

where $\tilde{h}$ is a smoothing parameter, $\phi_{\tilde{h}}^i (\cdot) = \phi(\cdot/\tilde{h})$, and $\phi(\cdot)$ is a fixed kernel function with $\phi(\cdot) \geq 0$ and $\int \phi(t) \, dt = 1$. Ichimura et al. [11] showed estimates of $\beta$ and $\tilde{h}$ can be achieved by simultaneously minimizing the following objective function with respect to $\beta$ and $\tilde{h}$.

$$\hat{S}(\beta, \tilde{h}) = \sum_i (Y_i - \hat{g}_{-i}(\beta' X_i|\beta))^2$$

As the above SIM estimator is no longer valid when the data are subject to right-censoring, we suggest an alternative SIM estimator that is capable of handling right-censored data under survival-type outcomes. This new estimator is primarily based on the method proposed by Ichimura et al. [11] but it is re-weighted by an IPCW weighting scheme. We define leave-one-out re-weighed estimator of $g(\cdot)$ as

$$\hat{g}_{-i}(\nu|\beta) = \frac{\sum_{l \neq i} \delta_{i}^{T_{i-l}} K_{i}^c(T_{i-l}) \phi_{\hat{h}}^i (\nu - \beta' X_i)}{\sum_{l \neq i} \delta_{i}^{T_{i-l}} K_{i}^c(T_{i-l}) \phi_{\tilde{h}}^i (\nu - \beta' X_i)},$$

Here, $K_{i}^c(T_{\cdot})$ is the survival probability of an individual not being censored just before time $T$ with $K_{i}^c(t) = \prod_{s \geq t} [1 - \lambda_{i}^{c}(t)\tilde{Z}_{i}(s)] ds$ and $\tilde{Z}_{i}(t)$ is a generalized covariate defined for the $i$th individual, which we will explain in detail in the sequel. Note that $K_{i}^c(t)$ does not have a survival function interpretation unless $\tilde{Z}_{i}(t)$ is formed with non-time-varying covariates. In reality, we replace $K_{i}^c(\cdot)$ by its corresponding estimator $\hat{K}_{i}^c(\cdot)$ and later we will introduce a flexible model to estimate $K_{i}^c(t)$ for any time $t$. 
We estimate $\beta$ and $\tilde{h}$ by minimizing the following weighted objective function denoted by $(\hat{S}'(\beta, \tilde{h}))$ simultaneously with respect to both $\beta$ and $h$.

$$\hat{S}'(\beta, \tilde{h}) = \sum_i \frac{\delta_i}{K_i(T_i-)} (T_i - \hat{g}_{-i}(\beta'X_i|\beta))^2.$$ 

For the estimator of $\beta$ ($\hat{\beta}$) and optimal $\tilde{h}$ ($\tilde{h}_0$), $g(.)$ function at a new point $v_0 = \beta'x_0$ can be estimated as

$$\hat{g}(v_0|\hat{\beta}) = \frac{\sum_i \frac{\delta_i}{K_i(T_i-)} \phi_{\tilde{h}_0}(v_0 - \hat{\beta}'X_i)}{\sum_i \frac{\delta_i}{K_i(T_i-)} \phi_{\tilde{h}_0}(v_0 - \hat{\beta}'X_i)}.$$

### 2.4. IPCW re-weighted estimator for $\mu_{jk}(u_j)$

As described before, the estimator given by (9) is intend to use for estimating $\mu_{jk}(u_j)$, $k = 1, \ldots, K$, when the outcome $j$, $j = 1, \ldots, J$, is uncensored. Now, we provide an IPCW re-weighted smooth mean estimator for estimating $\mu_{jk}(u_j)$ to handle the right-censored survival outcomes, where the estimator (9) is no longer valid. The new re-weighted estimator was obtained similarly as the approach followed to adjust the SIM estimator for censored case. This is given by

$$\hat{\mu}_{jk}(u_j) = \frac{\sum_{i=1}^{n_k} \frac{\delta_{jk}^i}{K_{jk}(T_{jk}^-)} T_{jki} \omega \left( (s_j - \hat{S}(X_{ki})) / h_{jk} \right) I \left( \hat{S}(X_{ki}) = d_j \right)}{\sum_{i=1}^{n_k} \frac{\delta_{jk}^i}{K_{jk}(T_{jk}^-)} \omega \left( (s_j - \hat{S}(X_{ki})) / h_{jk} \right) I \left( \hat{S}(X_{ki}) = d_j \right)},$$

for a score value $u_j = (s_j, d_j)$ correspond to a covariate $x_0$. The estimation of weights can be performed using the Aalen’s additive model as described in the next section.

Following the estimation of IPCW re-weighted SIM models for a right-censored outcome $Y_j$ and covariate $X_k$ pairs, we subsequently estimate scores $\hat{U}_j(X) = (\hat{S}_j(X), \hat{h}_j(X))$ at each covariate point $X$, including the covariate of new patient $x_0$. Finally, $\hat{\mu}_{jk}(u_j)$ is obtained using (11).

Now, if we have a set of $J$ responses that contain some right-censored responses, we will be able to calculate the estimated score $\hat{u}_0 = (\hat{u}_{10}, \ldots, \hat{u}_{J0})'$ and its corresponding $\hat{\mu}_k(\hat{u}_0) = (\hat{\mu}_{1k}(\hat{u}_{10}), \ldots, \hat{\mu}_{Jk}(\hat{u}_{J0}))'$ for a given new patient covariate value $x_0$. The estimated best treatment $k^*(u)$ is obtained following steps given in (2) to (3).

### 2.5. Estimation of IPCW weights

We use the Aalen’s nonparametric additive model [1,2] to calculate IPCW weights, which provides a flexible structure to estimate the censoring hazards by allowing covariates to be varied over time. Let $Z_i(t)$ be a generalized covariate defined for individual $i$, $1 \leq i \leq n_k$, at time $t$, which may contain both baseline and additional covariates (could be time varying) than covariates of primary interest $X$ that affect the censoring hazards. Suppose $\tilde{Z}_i(t) = \sigma \{Z_i(s) : 0 \leq s < t \}$ is the observed covariate history prior to $t$. The additive censoring hazard of $i$th individual at time $t$ is given by the linear form $\lambda_i(t|\tilde{Z}_i(t)) = \sum_{m=0}^{M} \eta_m(t) W_{im}(t)$,
Table 1. Smooth mean functions: Model Set-1. Here, the first response ($j = 1$) is considered to be right censored. Index vectors $\beta_{jk}, j = 1, \ldots, 4, k = 1, \ldots, 3$ are provided in Table 5.

| Response ($j$) | Treatment Group ($k$) |
|---------------|----------------------|
|               | $k = 1$             | $k = 2$             | $k = 3$             |
| $j = 1$ (cen.)| $1 + \exp[0.5(\beta_{11}X)^3]$ | $1 + \exp(-0.5 + 0.5(-0.8 + \beta_{22}X)^2)$ | $1 + \exp(1.0 - 5(\beta_{13}X)^2)$ |
| $j = 2$       | $\exp(0.5 - 2.5(-1.0 + \beta_{23}X)^2)$ | $\exp(0.5 - 2.5(1.0 + \beta_{22}X)^2)$ | $\exp(0.5 - 2.5(\beta_{23}X)^2)$ |
| $j = 3$       | $\exp(0.5 - 3(-1.0 + \beta_{31}X)^2)$ | $\exp(0.5 - 3(1.0 + \beta_{32}X)^2)$ | $\exp(0.5 - 3(\beta_{33}X)^2)$ |
| $j = 4$       | $1 + \exp(-1.0 + \beta_{41}X)$ | $1 + \exp(-1.0 - \beta_{42}X)$ | $\exp(1.0 - (\beta_{43}X)^2)$ |

where $W_{i0}(t) = 1$ and $W_{im}(t) = f_m(\tilde{Z}_i(t)), m = 1, \ldots, M$, are possibly time-dependent functions of the past history of the covariate process for subject $i$ and $\eta_m(t)$ are unknown regression functions that measure the effect of corresponding covariate functions on the censoring hazard. Define $W_i(t) = (W_{i1}(t), \ldots, W_{ip}(t))$. Then the Aalen’s estimator of cumulative censoring hazard for the $j$th individual is given by

$$\Lambda_j^c(t|\tilde{Z}_i(t)) = \int_0^t \hat{\lambda}_j^c(u|\tilde{Z}_i(t)) \, du = \sum_{i=1}^{nk} I(T_i \leq t) (1 - \delta_i) W_i(T_i) R^{-1}(T_i) W_i(T_i),$$

where $R(t) = \sum_{i=1}^{nk} I(T_i \geq t) W_i(t) W_i(t)$. The estimated IPCW weight for $i$th individual can be expressed as $\hat{\Lambda}_j^c(t|\tilde{Z}_i(t)) = \exp(-\hat{\Lambda}_j^c(t|\tilde{Z}_i(t)))$, where $\Lambda_j^c(t|\tilde{Z}_i(t)) = \int_0^t \hat{\lambda}_j^c(u|\tilde{Z}_i(t)) \, du$.

3. Empirical studies

In this section, we present a simulation study that investigates the properties of the proposed procedure in finite sample cases.

In this investigation, we primarily focused on the accuracy of treatment assignment of a new (test) observation using the proposed technique based on simulated samples for $K$ treatment groups and $J$ responses. We fixed one of the responses among $J$ set to be a right-censored time-to-event type response. The censoring time for the right-censored variable was generated under both random and covariate dependent settings with rates ranged between 25% and 50%. To give a relevance for a personalized treatment scenario, we selected our model sets such that each model in a set dominates other competing models for some combination of covariate values. In other words, none of considered models fully dominate other models within the whole covariate space that makes the optimal treatment to be a function of patients’ covariate profiles. We used two types of nonlinear model functions, given by Model Set-1 (Table 1) and Mode Set-2 (Table 2), for the generation of mean response per each $k = 1, \ldots, K$ and $j = 1, \ldots, J$. The kernel function for all smoothing was taken to be a Normal probability density function ($N(0, 1)$). The bandwidth selection for smooth the mean estimators was by the algorithm given by Wand and Jones [38].

Primary steps of the simulation study are given below:

1. Covariate $X$: Simulate $K$ independent $r = 5$ dimensional multivariate random samples of size $n$, letting distribution of each component of $X$ as $U(-1, 1)$. Select $n$ from the set $\{100, 200, 400\}$. 
Table 2. Smooth mean functions: Model Set-2. Here, the first response \( (j = 1) \) is considered to be right-censored. We choose the common vector \( C \) to be a unit vector; \( C = (\frac{1}{\sqrt{r}}, \ldots, \frac{1}{\sqrt{r}})^T \), for all combinations of \( j \) and \( k \).

| Treatment Group \( (k) \) | \( k = 1 \) | \( k = 2 \) | \( k = 3 \) |
|---------------------------|---------|---------|---------|
| Response \( (j) \) | \( j = 1 \) (cen.) | \( 2 + \sin(\frac{2\pi}{5} + \pi/2(C'X)) \) | \( 2 + \sin(\frac{6\pi}{5} + \pi/2(C'X)) \) | \( 2 + \sin(\frac{-2\pi}{5} + \pi/2(C'X)) \) |
| | \( j = 2 \) | \( \cos(\frac{\pi}{2}(C'X)) \) | \( \cos(\frac{4\pi}{5} + \pi/2(C'X)) \) | \( \cos(\frac{-4\pi}{5} + \pi/2(C'X)) \) |
| | \( j = 3 \) | \( \sin(\frac{\pi}{3} + \frac{\pi}{3}(C'X)) \) | \( \sin(\frac{\pi}{3} + \frac{\pi}{3}(C'X)) \) | \( \sin(\frac{-3\pi}{3} + \pi/3(C'X)) \) |
| | \( j = 4 \) | \( \cos(\frac{\pi}{3}(C'X)) \) | \( \cos(\frac{2\pi}{3} + \frac{\pi}{3}(C'X)) \) | \( \cos(\frac{-2\pi}{3} + \frac{\pi}{3}(C'X)) \) |

Table 3. Parameters used for generating censoring times for Model Set-1, with random and covariate dependent censoring settings.

| Group | Random \( (\zeta) \) | \( 25\% \) | \( 50\% \) | \( \zeta_1 \) | \( \zeta_2 \) | \( \zeta_1 \) | \( \zeta_2 \) |
|-------|-----------------|---------|---------|---------|---------|---------|---------|
| \( k = 1 \) | 0.14 | 0.35 | 0.10 | 0.20 | 0.20 | 0.50 |
| \( k = 2 \) | 0.14 | 0.35 | 0.10 | 0.20 | 0.20 | 0.50 |
| \( k = 3 \) | 0.12 | 0.32 | 0.07 | 0.18 | 0.18 | 0.47 |

Table 4. Parameters used for generating censoring times for Model Set-2, with random and covariate dependent censoring settings.

| Group | Random cen. \( (\zeta) \) | \( 25\% \) | \( 50\% \) | \( \zeta_1 \) | \( \zeta_2 \) | \( \zeta_1 \) | \( \zeta_2 \) |
|-------|-----------------|---------|---------|---------|---------|---------|---------|
| \( k = 1 \) | 0.11 | 0.26 | 0.05 | 0.20 | 0.15 | 0.40 |
| \( k = 2 \) | 0.18 | 0.45 | 0.10 | 0.26 | 0.25 | 0.62 |
| \( k = 3 \) | 0.23 | 0.55 | 0.15 | 0.30 | 0.35 | 0.70 |

(2) For the functions \( g_{j1}(.), \ldots, g_{jK}(.) \) and index vectors \( \beta_1, \ldots, \beta_K \) generate treatment responses from model (5) for \( K \) groups.

(3) For each \( k, k = 1, \ldots, K \), generate \( n \) errors from either a \( J \)-dimensional multivariate normal distribution or a multivariate double exponential distribution with zero mean and a correlation matrix with off-diagonal elements given by \( \rho \) and dispersion parameter \( \sigma \). Select \( \sigma \) from the set \{0.3, 0.5, 1.0\}. Use the R package \texttt{mvtnorm} [9] for multivariate normal case and use the package \texttt{LaplacesDemon} [30] for generating double exponential random variables.

(4) Fix the first component of the response vector \( (j = 1) \) to be a right-censored survival time type response and generate censoring times \( (C) \) as follows. For censoring rate 25% and 50% cases, select parameters from Tables 3 and 4 for Model Sets 1 and 2, respectively.

(a) Random censoring time: Use the Exponential distributions with a scale parameter \( \zeta \).
Table 5. Index vectors $\beta_{jk}$, $j = 1, \ldots, J$, $k = 1, \ldots, K$, selected for the simulation study with Model Set-1. We specified each $\beta_{jk}' = (\beta_{jk1}, \ldots, \beta_{jkr})_{1 \times r}$ with $|\beta_{jk}| = 1$.

| $\beta_{11}$ | 0.74 | -0.37 | -0.37 | 0.37 | -0.19 |
| $\beta_{12}$ | 0.80 | 0.00 | 0.20 | -0.40 | -0.40 |
| $\beta_{13}$ | 0.07 | 0.15 | 0.30 | 0.45 | -0.82 |
| $\beta_{21}$ | 0.23 | 0.15 | 0.45 | -0.83 | -0.15 |
| $\beta_{22}$ | -0.69 | 0.51 | 0.34 | 0.34 | -0.17 |
| $\beta_{23}$ | 0.63 | 0.21 | 0.32 | -0.53 | -0.42 |
| $\beta_{31}$ | 0.15 | 0.30 | 0.07 | -0.82 | 0.45 |
| $\beta_{32}$ | 0.24 | 0.16 | 0.65 | -0.24 | -0.65 |
| $\beta_{33}$ | -0.18 | 0.36 | 0.54 | 0.18 | -0.72 |
| $\beta_{41}$ | -0.40 | 0.00 | 0.80 | -0.40 | 0.20 |
| $\beta_{42}$ | 0.48 | 0.27 | -0.55 | 0.41 | -0.48 |
| $\beta_{43}$ | -0.75 | 0.34 | -0.14 | 0.14 | 0.54 |

(b) Covariate-dependent censoring: Use the Exponential distribution with the scaler parameter given by the indicator function,

\[ C \sim I(\rho'X > w) \exp(\xi_1) + I(\rho'X \leq w) \exp(\xi_2), \]

with $w = 0$ and an arbitrarily selected vector $\rho = (0.7, 0.3, 0, 0.5, -0.5)'$.

(5) Estimate corresponding SIMs for $k = 1, \ldots, K$ and $j = 1, \ldots, J$, using observed response covariate pairs: Use the IPCW re-weighted SIM estimator for the right censored responses and use the regular SIM estimation for the completely observed cases $j \neq 1$.

(6) Estimate individual scores $U_j(X) = (S_j(X), \delta_j(X))'$ at each covariate value $X$ for $j = 1, \ldots, J$, following (8). This gives the estimated overall score $\hat{U}(X) = (\hat{U}_1(X), \ldots, \hat{U}_J(X))$ for $X$.

(7) Generate a new covariate value $x_0$ using the parameters in Step 1. Then estimate the score $\hat{u}(x_0) = (\hat{u}_{10}, \ldots, \hat{u}_{J0})$.

(8) Calculate $\hat{\mu}_{jk}(\hat{u}_{j0})$ for $j = 1, \ldots, J, k = 1, \ldots, K$: Use the IPCW re-weighted estimator (11) for the right censored response $j = 1$ and the regular estimator (9) for complete cases $j \neq 1$.

(9) Estimate the corresponding $\hat{k}^*$ for weights vector $w$.

(10) Generate $K$ response vectors $y_{0k} = (g_{1k}(\beta_{1k}'x_0), \ldots, g_{jk}(\beta_{jk}'x_0))'$ + $\epsilon_{0k}$. Obtain $\epsilon_{0k}$ using the same $J$-dimensional multivariate distribution as in Step 3.

(11) Obtain rank vectors $\tilde{v}_j, j = 1, \ldots, J$, for each row of the mean matrix $(y_{01}, \ldots, y_{0K})$, and minimize

\[ \psi(v) = \sum_{j=1}^{J} w_j \gamma(v, \tilde{v}_k) \] (12)

over $P_K$ for same weights $\omega$ above to get the corresponding aggregated vector $(\hat{v}_1^*, \ldots, \hat{v}_K^*)'$ and define the treatment assignment to be correct if

\[ \hat{k}^* = \arg \min_{1 \leq k \leq K} \{ \hat{v}_k^* \} \]

for the $\hat{k}^*$ corresponding to $\hat{\mu}_{jk}$.s.

(12) Repeat steps 1–11 1000 times.
We conducted our simulation study for \( K = 3 \) with \( J = 3 \) and \( J = 4 \) cases by fixing one of the responses to be a right-censored response (i.e. \( j = 1 \) in each case). Tables 6–9 summarize the frequencies of accurate treatment selection in 1000 test cases. These results are stratified for two choices of weights: equal and unequal cases. In the unequal weight scenario, we assigned a large weight to the right-censored outcome, to closely examine the impact of censoring on the selection performance. As we found, especially when a large weight was assigned to the censored-response, the accuracy dropped at a high censoring rate, but continued to improve as the per group size \( n \) increased. In general, a performance drop should be expected for the high censoring rate and small sample combination, as the IPCW weights correct the censoring bias in an asymptotic fashion. For a given censoring rate, comparing random and covariate-dependent censoring types, there were no noticeable differences in the selection performance. As to be expected, the selection accuracy drops when the error distribution has high variability.

In conclusion, we observed reasonable performance by the proposed technique while considering the complexity of overall the treatment selection problem. For instance, in Model Set-2, we used highly nonlinear mean functions based on sine and cosine functions, which creates a complex treatment selection scenario, however, the selection accuracy remained reasonably high for those cases, showing the potential of the proposed technique in real-life applications.

In addition to those primary simulations, we also conducted a few additional studies on two different aspects. In one of those studies, we investigated the performance for the single response case \( (J = 1) \), with only a censored outcome under two treatment option \( (K = 2) \) scenario. Results of this study are provided in the Supplementary Material (see Supplementary Tables S1 and S2). Similar to previous results, we observed increased accuracy in the optimal selection as the training set size increases. In the other study, we investigated the performance of the method under a set of perturbed SIM models (see Supplementary Table S4). Demonstrating the robustness under the departure from SIM structure, results of this study showed comparable performance similar to its counterpart, the Model-2 case (see Supplementary Table S3).

4. Illustration using data from the ACTG-175 HIV randomized controlled trial

In this section, we provide an illustration of our proposed method using the data resulted from the ACTG 175 Clinical Trial [10]. This clinical trial was a randomized, double-blinded, placebo-controlled clinical trial that was conducted to compare single nucleoside or two nucleoside antiviral medications in adults infected with human immunodeficiency (HIV-1) whose T-cell CD4 counts were in the range of 200–500 per cubic millimeter. The study randomized HIV-1 infected patients to one of four daily regimens: 600 mg of zidovudine (arm-0), 600 mg of zidovudine plus 400 mg of didanosine (arm-1), 600 mg of zidovudine plus 2.25 mg of zalcitabine (arm-2), or 400 mg of didanosine (arm-3). The data set contains information on 2136 HIV-1-infected subjects. Arms 0, 1, 2, and 3 contain 532, 519, 524, and 561 patients, respectively [15]. The primary end-point of this trial was a survival event which we describe in the sequel.

Since the original treatment assignment in this study was based on random selection, there was no consideration for the optimality. Our intention in this data analysis is simply to
Table 6. Accuracies of treatment selection in 1000 test cases using the proposed technique for the case of three treatments and four responses with Model Set-1.

| Weights ($w_1, w_2, w_3, w_4$) | Error dist. | Error disp. para. ($\sigma$) | Size ($n$) | No censoring | 25% | 50% | 25% | 50% | No censoring | 25% | 50% | 25% | 50% |
|-------------------------------|-------------|------------------------------|-----------|--------------|-----|-----|-----|-----|--------------|-----|-----|-----|-----|
| (0.7,0.1,0.1, 0.1)           | Normal      | 0.1                          | 100       | 853          | 809 | 760 | 804 | 765 | 838          | 808 | 765 | 801 | 764 |
|                               |             |                              | 200       | 867          | 841 | 836 | 840 | 832 | 852          | 830 | 812 | 826 | 810 |
|                               |             |                              | 400       | 886          | 860 | 850 | 860 | 846 | 878          | 860 | 850 | 851 | 851 |
|                               | D.E.        | 0.1                          | 100       | 758          | 707 | 686 | 700 | 683 | 718          | 686 | 619 | 687 | 615 |
|                               |             |                              | 200       | 787          | 753 | 735 | 780 | 762 | 757          | 740 | 710 | 755 | 715 |
|                               |             |                              | 400       | 794          | 768 | 748 | 765 | 759 | 775          | 772 | 749 | 772 | 769 |
|                               |             |                              | 0.5       | 618          | 584 | 507 | 570 | 501 | 629          | 598 | 501 | 595 | 509 |
|                               |             |                              | 100       | 677          | 642 | 611 | 659 | 602 | 664          | 645 | 595 | 621 | 582 |
|                               |             |                              | 200       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               |             |                              | 400       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               |             |                              | 0.3       | 618          | 584 | 507 | 570 | 501 | 629          | 598 | 501 | 595 | 509 |
|                               |             |                              | 100       | 677          | 642 | 611 | 659 | 602 | 664          | 645 | 595 | 621 | 582 |
|                               |             |                              | 200       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               |             |                              | 400       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               |             |                              | 0.5       | 618          | 584 | 507 | 570 | 501 | 629          | 598 | 501 | 595 | 509 |
|                               |             |                              | 100       | 677          | 642 | 611 | 659 | 602 | 664          | 645 | 595 | 621 | 582 |
|                               |             |                              | 200       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               |             |                              | 400       | 690          | 651 | 671 | 676 | 651 | 686          | 642 | 622 | 636 | 625 |
|                               | (1,0,1,1, 0,1.0) | Normal      | 0.1       | 100          | 872          | 852 | 829 | 856 | 831 | 859          | 837 | 814 | 821 | 812 |
|                               |             |                              | 200       | 887          | 875 | 861 | 874 | 864 | 868          | 845 | 837 | 849 | 836 |
|                               |             |                              | 400       | 902          | 888 | 879 | 889 | 876 | 879          | 860 | 851 | 861 | 848 |
|                               |             |                              | 0.3       | 831          | 810 | 798 | 807 | 779 | 730          | 709 | 698 | 714 | 692 |
|                               |             |                              | 100       | 838          | 827 | 812 | 826 | 804 | 782          | 770 | 747 | 773 | 741 |
|                               |             |                              | 200       | 848          | 826 | 818 | 824 | 817 | 794          | 761 | 755 | 761 | 752 |
|                               |             |                              | 400       | 697          | 683 | 645 | 689 | 665 | 653          | 630 | 607 | 627 | 606 |
|                               |             |                              | 0.5       | 730          | 715 | 707 | 717 | 708 | 667          | 638 | 633 | 637 | 626 |
|                               |             |                              | 100       | 894          | 879 | 859 | 872 | 855 | 847          | 829 | 814 | 830 | 814 |
|                               |             |                              | 200       | 896          | 889 | 876 | 886 | 874 | 852          | 844 | 838 | 840 | 833 |
|                               |             |                              | 400       | 910          | 903 | 891 | 904 | 889 | 877          | 879 | 847 | 879 | 866 |
|                               |             |                              | 0.3       | 843          | 820 | 794 | 812 | 793 | 783          | 767 | 740 | 762 | 741 |
|                               |             |                              | 100       | 850          | 822 | 804 | 814 | 804 | 790          | 776 | 770 | 778 | 767 |
|                               |             |                              | 200       | 861          | 846 | 836 | 845 | 834 | 807          | 797 | 784 | 798 | 788 |
|                               |             |                              | 400       | 746          | 737 | 706 | 735 | 707 | 699          | 672 | 644 | 671 | 645 |
|                               |             |                              | 0.5       | 746          | 737 | 706 | 735 | 707 | 699          | 672 | 644 | 690 | 644 |
|                               |             |                              | 200       | 773          | 759 | 741 | 764 | 740 | 710          | 698 | 677 | 690 | 664 |
|                               |             |                              | 400       | 789          | 760 | 754 | 759 | 749 | 722          | 705 | 706 | 719 | 707 |
Table 7. Accuracies of treatment selection in 1000 test cases using the proposed technique for the case of three treatments and three responses with Model Set-1.

| Weights $(w_1, w_2, w_3)$ | Error dist. | Error disp. para. $(\sigma)$ | Size $(n)$ | No censoring | Random | Cov. dep. | Random | Cov. dep. |
|---------------------------|-------------|-----------------------------|----------|-------------|--------|----------|--------|----------|
|                           |             |                             |          | 25%         | 50%    | 25%      | 50%    | 25%      | 50%      |
| (0.8,0.1,0.1)             | Normal      | 0.1                         | 100      | 844         | 811    | 778      | 809    | 776      | 841      |
|                           |             |                              | 200      | 829         | 826    | 815      | 830    | 808      | 857      |
|                           |             |                              | 400      | 853         | 852    | 833      | 857    | 835      | 874      |
|                           | D.E.        | 0.1                         | 100      | 722         | 690    | 647      | 686    | 642      | 745      |
|                           |             |                              | 200      | 756         | 720    | 699      | 737    | 704      | 760      |
|                           |             |                              | 400      | 785         | 751    | 740      | 755    | 729      | 780      |
|                           | Normal      | 0.5                         | 100      | 619         | 578    | 538      | 580    | 539      | 641      |
|                           |             |                              | 200      | 660         | 624    | 599      | 615    | 602      | 655      |
|                           | D.E.        | 0.5                         | 100      | 676         | 634    | 625      | 649    | 627      | 679      |
| (1.0,1.0,1.0)             | Normal      | 0.1                         | 100      | 843         | 803    | 769      | 801    | 772      | 829      |
|                           |             |                              | 200      | 849         | 820    | 798      | 820    | 808      | 851      |
|                           |             |                              | 400      | 861         | 833    | 822      | 835    | 821      | 864      |
|                           | D.E.        | 0.1                         | 100      | 757         | 704    | 674      | 701    | 668      | 758      |
|                           |             |                              | 200      | 771         | 761    | 763      | 774    | 765      | 763      |
|                           |             |                              | 400      | 795         | 792    | 773      | 785    | 780      | 789      |
|                           | Normal      | 0.5                         | 100      | 659         | 608    | 540      | 609    | 545      | 632      |
|                           |             |                              | 200      | 686         | 677    | 620      | 674    | 619      | 703      |
|                           | D.E.        | 0.5                         | 100      | 711         | 698    | 671      | 702    | 681      | 718      |
| (0.3,0.1,0.1)             | Normal      | 0.1                         | 100      | 881         | 838    | 797      | 832    | 800      | 848      |
|                           |             |                              | 200      | 885         | 861    | 845      | 869    | 843      | 865      |
|                           |             |                              | 400      | 898         | 870    | 847      | 872    | 861      | 868      |
|                           | D.E.        | 0.1                         | 100      | 807         | 764    | 714      | 756    | 720      | 793      |
|                           |             |                              | 200      | 811         | 797    | 773      | 789    | 781      | 801      |
|                           |             |                              | 400      | 842         | 838    | 817      | 838    | 818      | 805      |
|                           | Normal      | 0.5                         | 100      | 694         | 663    | 614      | 654    | 618      | 682      |
|                           |             |                              | 200      | 731         | 709    | 671      | 694    | 677      | 696      |
|                           | D.E.        | 0.1                         | 100      | 881         | 859    | 845      | 857    | 844      | 855      |
|                           |             |                              | 200      | 885         | 868    | 857      | 870    | 861      | 870      |
|                           |             |                              | 400      | 909         | 872    | 857      | 869    | 860      | 884      |
|                           | Normal      | 0.3                         | 100      | 826         | 792    | 780      | 797    | 788      | 788      |
|                           |             |                              | 200      | 832         | 819    | 804      | 820    | 817      | 796      |
|                           | D.E.        | 0.1                         | 100      | 826         | 792    | 780      | 797    | 788      | 788      |
|                           |             |                              | 200      | 832         | 819    | 804      | 820    | 817      | 796      |
|                           | Normal      | 0.5                         | 100      | 815         | 692    | 681      | 685    | 671      | 698      |
|                           |             |                              | 200      | 753         | 741    | 727      | 745    | 729      | 706      |
|                           | D.E.        | 0.1                         | 100      | 826         | 792    | 780      | 797    | 788      | 788      |
|                           |             |                              | 200      | 832         | 819    | 804      | 820    | 817      | 796      |
Table 8. Accuracies of treatment selection in 1000 test cases using the proposed technique for the case of three treatments and four responses with Model Set-2.

| Weights ($w_1, w_2, w_3, w_4$) | Error dist. | Error disp. para. ($\sigma$) | Size ($n$) | No censoring | Random | Cov. dep. | No censoring | Random | Cov. dep. |
|---------------------------------|-------------|-------------------------------|------------|--------------|--------|----------|--------------|--------|----------|
| 0.7,0.1,0.1,0.1                  | Normal      | 0.1                           | 100        | 921          | 890    | 835      | 919          | 891    | 834      |
|                                 |             |                               | 200        | 938          | 911    | 895      | 919          | 891    | 834      |
|                                 |             |                               | 400        | 946          | 924    | 918      | 933          | 916    | 867      |
| 0.3                             | Normal      | 0.3                           | 100        | 857          | 789    | 756      | 819          | 897    | 878      |
|                                 |             |                               | 200        | 881          | 821    | 788      | 819          | 897    | 878      |
|                                 |             |                               | 400        | 896          | 857    | 817      | 861          | 828    | 891      |
| 0.5                             | Normal      | 0.5                           | 100        | 799          | 733    | 682      | 741          | 691    | 788      |
|                                 |             |                               | 200        | 836          | 798    | 724      | 783          | 722    | 795      |
|                                 |             |                               | 400        | 842          | 818    | 763      | 815          | 752    | 812      |
| D.E.                            | Normal      | 0.1                           | 100        | 934          | 895    | 832      | 889          | 825    | 921      |
|                                 |             |                               | 200        | 943          | 914    | 886      | 918          | 879    | 933      |
|                                 |             |                               | 400        | 959          | 933    | 905      | 929          | 902    | 945      |
| 0.3                             | Normal      | 0.3                           | 100        | 862          | 796    | 733      | 803          | 724    | 842      |
|                                 |             |                               | 200        | 890          | 837    | 787      | 832          | 891    | 864      |
|                                 |             |                               | 400        | 896          | 856    | 833      | 865          | 821    | 878      |
| 0.5                             | Normal      | 0.5                           | 100        | 811          | 767    | 710      | 754          | 704    | 787      |
|                                 |             |                               | 200        | 830          | 798    | 782      | 800          | 769    | 796      |
|                                 |             |                               | 400        | 852          | 835    | 816      | 831          | 807    | 822      |
| (1.0,1.0,1.0,1.0)               | Normal      | 0.1                           | 100        | 924          | 906    | 896      | 912          | 896    | 926      |
|                                 |             |                               | 200        | 930          | 910    | 902      | 918          | 905    | 939      |
|                                 |             |                               | 400        | 942          | 929    | 917      | 934          | 917    | 951      |
| 0.3                             | Normal      | 0.3                           | 100        | 889          | 867    | 855      | 864          | 856    | 864      |
|                                 |             |                               | 200        | 896          | 874    | 861      | 877          | 863    | 878      |
|                                 |             |                               | 400        | 904          | 878    | 878      | 876          | 869    | 884      |
| 0.5                             | Normal      | 0.5                           | 100        | 848          | 828    | 805      | 829          | 805    | 844      |
|                                 |             |                               | 200        | 858          | 834    | 825      | 835          | 822    | 844      |
|                                 |             |                               | 400        | 867          | 854    | 842      | 857          | 843    | 861      |
| D.E.                            | Normal      | 0.1                           | 100        | 951          | 927    | 904      | 921          | 889    | 903      |
|                                 |             |                               | 200        | 959          | 921    | 904      | 932          | 909    | 925      |
|                                 |             |                               | 400        | 967          | 947    | 935      | 944          | 926    | 938      |
| 0.3                             | Normal      | 0.3                           | 100        | 893          | 873    | 844      | 873          | 842    | 849      |
|                                 |             |                               | 200        | 897          | 866    | 866      | 884          | 866    | 878      |
|                                 |             |                               | 400        | 910          | 890    | 870      | 891          | 880    | 854      |
| 0.5                             | Normal      | 0.5                           | 100        | 834          | 805    | 792      | 816          | 792    | 814      |
|                                 |             |                               | 200        | 858          | 842    | 827      | 844          | 827    | 837      |
|                                 |             |                               | 400        | 881          | 869    | 852      | 860          | 850    | 853      |
Table 9. Accuracies of treatment selection in 1000 test cases using the proposed technique for the case of three treatments and three responses with Model Set-2.

| Weights \( (w_1, w_2, w_3) \) | Error dist. | Error disp. para. \( (\sigma) \) | Size \( (n) \) | No censoring | 25% | 50% | 25% | 50% | No censoring | 25% | 50% | 25% | 50% |
|-----------------------------|-------------|----------------------------------|--------------|---------------|-----|-----|-----|-----|---------------|-----|-----|-----|-----|
| \((0.8,0.1,0.1)\)          | Normal      | 0.1                              | 100          | 926           | 862 | 826 | 858 | 818 | 927           | 880 | 818 | 865 | 816 |
|                             |             |                                  | 200          | 937           | 901 | 867 | 900 | 851 | 937           | 905 | 885 | 901 | 877 |
|                             |             |                                  | 400          | 947           | 917 | 899 | 916 | 897 | 946           | 919 | 904 | 923 | 897 |
| \(0.3\)                    |             |                                  | 100          | 844           | 799 | 733 | 786 | 721 | 854           | 780 | 717 | 775 | 712 |
|                             | Random      |                                  | 200          | 873           | 834 | 789 | 810 | 785 | 876           | 829 | 791 | 821 | 780 |
|                             | Cov. dep.   |                                  | 400          | 890           | 860 | 830 | 854 | 806 | 897           | 859 | 816 | 861 | 818 |
| \(0.5\)                    |             |                                  | 100          | 774           | 706 | 667 | 708 | 657 | 780           | 708 | 663 | 709 | 650 |
|                             | Random      |                                  | 200          | 791           | 739 | 709 | 732 | 702 | 790           | 753 | 702 | 747 | 702 |
|                             | Cov. dep.   |                                  | 400          | 832           | 801 | 756 | 802 | 749 | 802           | 788 | 735 | 788 | 741 |
| D.E.                       |             | 0.1                              | 100          | 908           | 856 | 807 | 862 | 805 | 923           | 855 | 808 | 857 | 796 |
|                             | Normal      |                                  | 200          | 922           | 889 | 868 | 903 | 863 | 935           | 902 | 861 | 902 | 858 |
|                             |             |                                  | 400          | 939           | 915 | 898 | 920 | 894 | 945           | 931 | 903 | 924 | 889 |
| \(0.3\)                    |             |                                  | 100          | 861           | 784 | 731 | 785 | 734 | 869           | 787 | 737 | 781 | 729 |
|                             | Random      |                                  | 200          | 883           | 820 | 781 | 818 | 772 | 842           | 814 | 787 | 818 | 783 |
|                             | Cov. dep.   |                                  | 400          | 891           | 859 | 810 | 847 | 812 | 860           | 827 | 815 | 830 | 811 |
| \(0.5\)                    |             |                                  | 100          | 799           | 702 | 661 | 706 | 661 | 817           | 702 | 679 | 716 | 683 |
|                             | Random      |                                  | 200          | 816           | 745 | 706 | 746 | 705 | 787           | 738 | 708 | 757 | 707 |
|                             | Cov. dep.   |                                  | 400          | 832           | 798 | 747 | 804 | 758 | 806           | 785 | 742 | 789 | 727 |
| \((1.0,1.0,1.0)\)          | Normal      | 0.1                              | 100          | 929           | 905 | 887 | 905 | 885 | 932           | 917 | 881 | 915 | 893 |
|                             |             |                                  | 200          | 930           | 916 | 900 | 912 | 908 | 942           | 929 | 904 | 926 | 910 |
|                             | D.E.        |                                  | 400          | 941           | 922 | 907 | 914 | 911 | 957           | 928 | 917 | 928 | 919 |
| \(0.3\)                    |             |                                  | 100          | 869           | 847 | 814 | 843 | 809 | 867           | 855 | 821 | 852 | 822 |
|                             | Random      |                                  | 200          | 891           | 879 | 855 | 879 | 846 | 878           | 851 | 847 | 862 | 845 |
|                             | Cov. dep.   |                                  | 400          | 902           | 882 | 869 | 878 | 871 | 894           | 865 | 850 | 869 | 856 |
| \(0.5\)                    |             |                                  | 100          | 830           | 799 | 785 | 805 | 789 | 799           | 772 | 759 | 769 | 751 |
|                             | Random      |                                  | 200          | 839           | 819 | 807 | 823 | 812 | 810           | 783 | 779 | 787 | 770 |
|                             | Cov. dep.   |                                  | 400          | 849           | 825 | 825 | 832 | 830 | 821           | 793 | 788 | 793 | 774 |
| D.E.                       |             | 0.1                              | 100          | 924           | 909 | 886 | 905 | 887 | 920           | 904 | 873 | 892 | 878 |
|                             |             |                                  | 200          | 936           | 916 | 905 | 916 | 901 | 931           | 912 | 897 | 907 | 900 |
| \(0.3\)                    |             |                                  | 400          | 940           | 917 | 911 | 918 | 911 | 949           | 927 | 910 | 928 | 917 |
|                             | Random      |                                  | 100          | 889           | 860 | 845 | 858 | 843 | 864           | 848 | 819 | 847 | 818 |
|                             | Cov. dep.   |                                  | 200          | 898           | 879 | 854 | 874 | 860 | 871           | 837 | 828 | 838 | 846 |
| \(0.5\)                    |             |                                  | 400          | 905           | 899 | 887 | 895 | 873 | 887           | 857 | 853 | 857 | 850 |
|                             | Random      |                                  | 100          | 836           | 810 | 785 | 812 | 789 | 812           | 789 | 770 | 793 | 763 |
|                             | Cov. dep.   |                                  | 200          | 840           | 823 | 805 | 820 | 810 | 824           | 799 | 790 | 807 | 795 |
|                             |             |                                  | 400          | 861           | 859 | 842 | 857 | 848 | 830           | 823 | 815 | 824 | 812 |
demonstrate what would be the optimal treatment for a new subject based on his/her characteristics, if training data from an RCT trial were available in advance. In this trial, study participants were periodically examined to capture T-cell counts (i.e. CD4 T-helper cells and CD8 cytotoxic T cells), that are critical components in the human immune system. Between these two types of T cells, CD4 helper cells considered to be the most important component on HIV/AIDS immune response as it involves in suppressing the HIV cell replication via signaling through various other cell types. The main role of the CD8 cell is typically referred to as the antibody reaction against cancers and various types of other viruses. However, studies suggest the important role of the CD8 during the early stages of HIV progression [31].

An individual’s survival endpoint in this study was defined based on three types of events: (i) individual’s CD4 count dropping less than 50% of the pre-treatment count; (ii) an event indicating progression of AIDS; (iii) death. Thus the term ‘survival time’ would denote an event time in that sense, that was subjected to right-censoring.

In previous studies, Sriwardhana et al. [28,29] have used this dataset to demonstrate personalized treatment strategies with respect to single and dual outcomes. However, the outcomes used in those analyses were completely observed (i.e. uncensored). In the current work, we considered three outcome types, including right-censored survival time and two uncensored clinical parameters given by CD4 and CD8 counts of a patient observed at week 20. All these outcomes were transformed via log transformation. As covariates, we used log-CD4 and log-CD8 counts at baseline, age, weight, and the number of months a patient received the pre-antiviral therapy.

In our analysis, we selected 200 patients at random from each treatment arm and considered their trial outcomes and baseline covariates to produce a training dataset. Based on this training data, we estimated the corresponding SIMs for three different outcome types, log-survival (log-T), log-CD4, and log-CD8, coupled with the above covariates. We utilized the IPCW re-weighted SIM for the right-censored log-T outcome. Next, assuming that the remaining 1,336 cases are ‘new’ patients, we applied the proposed treatment selection for those cases using combinations of weights ranging from 0 to 1 for each response.

We summarized the resulted assignments for test patients in Table 10. For example, when we chose log-T : log-CD4 : log-CD8 weights to be 0.7:0.2:0.1, using the proposed treatment selection, 16, 602, 352, and 366 test patients are proposed to be assigned to arms 0, 1, 2, and 3, respectively, whereas the corresponding assignment is 0, 544, 331, and 461 for corresponding weights 0.3:0.6:0.1. The overall pattern of these assignments indicates that only a few patients are proposed to be assigned to zidovudine alone, which used as the control arm in the study (arm-0). The marginal analysis of ACTG-175 data concludes the treatment with Zidovudine and Didanosine combination (arm-1), Zidovudine and Zalcitabine combination (arm-2), or Didanosine alone (arm-3) slows the progression of HIV disease and are superior to treatment with Zidovudine alone [10]. Although the proposed assignment in our approach was based on the patient-specific features, the overall assignment agrees with the above observation.

5. Discussion

In this article, we extend the personalized treatment plan suggested by Sriwardhana et al. [29] for a situation where the response data are subject to right censoring. The method
Table 10. Treatment assignment summary for ACTG-175 Clinical Trial data, by the proposed method based on three types of outcomes: log transformed survival time (log-T), and two clinical parameters CD4 (log-CD4) and CD8 (log-CD8) counts at week 20, using weights $w_{\text{log-T}}, w_{\text{log-CD4}},$ and $w_{\text{log-CD8}}$ for CD4 and CD8 counts, respectively. All outcomes were log transformed for the analysis.

| Weights | Assignments |
|---------|-------------|
| $w_{\text{log-T}}$ | $w_{\text{log-CD4}}$ | $w_{\text{log-CD8}}$ | Arm-0 | Arm-1 | Arm-2 | Arm-3 |
| 0.33 | 0.33 | 0.33 | 16 (1.20%) | 602 (45.06%) | 352 (26.35%) | 366 (27.40%) |
| 0.70 | 0.20 | 0.10 | 87 (6.51%) | 258 (19.31%) | 640 (47.90%) | 351 (26.27%) |
| 0.50 | 0.40 | 0.10 | 43 (3.22%) | 366 (27.40%) | 533 (39.90%) | 394 (29.49%) |
| 0.10 | 0.60 | 0.30 | 0 (0%) | 544 (40.72%) | 331 (24.78%) | 461 (34.51%) |

by Siriwardhana et al. [29] intends to select the optimal treatment among multiple treatment options ($K \geq 2$) when the outcome is multivariate. However, this method uses only the complete data. Our empirical studies show that the modified method performs very satisfactorily in selecting the optimal treatment in a multiple treatment setting and with different censoring scenarios. We demonstrated the proposed method using ACTG-175 trial data analysis, taking three different response variables, including the right-censored survival outcome as one of the outcomes.

We introduce an IPCW re-weighting scheme to adjust estimators used for the treatment selection and SIM estimation to handle right-censored data. As the sample mean is not calculable for censored survival data, we have used an IPCW weighted estimator (i.e. a localized version) which is defined for survival data under right censoring. As pointed out by Datta [6], this estimator is equivalent to a restricted mean survival time estimator where the Kaplan–Meier curve is integrated up to the largest event time (i.e. failure or censoring) in the data set. Further theoretical properties of this estimator can be derived, if needed following techniques in Satten et al. [24]. The estimation of IPCW weights is performed using the Aalen’s nonparametric additive model that gives the flexibility of using time-varying covariates. The proposed technique is based on semiparametric SIM models that add great flexibility in modeling real-life situations. This method could also be applied using quantile regression SIMs providing additional model flexibility, compared with the methods based on conditional expectations. However, we did not investigate this option in current research. The problem of selecting the scores summarizing the patient covariates could be an important one. Indeed, we have selected the scores mostly on intuitive grounds. More research on this issue needs to take place.

In the context of personalized medicine, there have been several previous works focused on the optimization of conditional restricted mean survival time given individualized details (for example, see Tian et al. [33]; Geng et al. [8]; Jiang et al. [12]; Cui et al. [4]). However, the main goal of these studies was to select the optimal therapy only based on the survival outcome, but not based on multiple responses.

Similar to our work, there have been many previous studies utilized the data set from ACTG-175 trial in various aspects, for illustrating statistical techniques developed on the optimal treatment selection. For example, following a semiparametric concept, Tsiatis et al. [34] analyzed the differences in mean CD4 count at week 20, to demonstrate covariate adjustments in RCT data when estimating the mean treatment effects. In their analysis, arm 0 (zidovudine alone) was compared with combined arms 1–3 (arm 1: zidovudine +
didanosine; arm 2: zidovudine + zalcitabine; arm 4: didanosine alone). In a similar analysis setting, Tan et al. [19] used a two-sample empirical likelihood weighting approach to estimate average treatment effect in 20-week CD4 count, while accounting for baseline covariates. In another study, Liang et al. [13] used a data-adaptive empirical likelihood-based approach under high-dimensional setting, adopting penalized regression to model the covariate-outcome relationship. They considered many forms of baseline variables, including linear and quadratic forms together with interactions to produce a high dimensional baseline predictor set. In summary, these three studies suggested the marginal superiority of the combined arm (i.e. arms 1–3), compared to arm 0 in terms of 20-week CD4 count response. To improve the efficiency of the log-rank test, Lu and Tsiatis [5] proposed a semi-parametric approach to recover information from auxiliary covariates that are correlating with the survival time. While using ACTG-175 data with multiple predictors of the survival time, they concluded marginal superiority in arm 1, compared to arm 0. Jiang et al. [32] illustrated a nonparametric personalized treatment strategy that maximizes the t-year survival probability with this dataset. Using arm 0 and arm 1 data, their analysis suggested large assignment proportions to arm 1, in the cases of $t = 600$ days and $t = 800$ days cases. Also, Cui et al. [18] analyzed the ACTG-175 survival data to demonstrate a method that based on causal survival forests assess heterogeneous treatment effects. Among many findings, their analysis suggested better outcomes from arm 1, compared to arm 3 for elderly patients. Addressing the multi response problem in personalized treatment case, Siriwardhana and Kualsekera [26] used ACTG-175 data considering two outcomes: CD4 and CD8 at week 20. They showed 2%, 75%, 11%, and 12% assignment rates for arms 0–3, while equally prioritizing two responses. Their analysis showed changes in group allocation rates under different response weights. Similar to the proposed method, Siriwardhana and Kualsekera [26] method uses a semiparametric framework together with conditional means, however, their method is limited only for fully observed (uncensored data).

The proposed methodology for handling multiple responses through rank aggregation is extremely general. Clearly, one could posit different regression models and estimation techniques for different components of the response vector. In addition, different optimality criteria can be used for ranking the treatments for different components of the response vector. The choices of weights are subjective selections based on patient preferences and advice from clinical experts or consultation from suitable support groups. We believe the response weight specification is more of a subjective matter than a quantitative issue where one can guide the choice of weights based on discussions with the patient and advocates, together with advice by clinicians. For responses that are related to the recovery or the progression of the disease, a group of clinical experts may be better suited for deciding the importance of each response in selecting a treatment. Whereas in cases where the outcome measures are related to indicators such as quality of life, economic impact, patient behavior, and so on, a clinician with consultation from various support groups may be better suited to determine what weights should be used. While it is useful to find data-driven guidelines, we have not explored those aspects in here. Also, in the current work, we do not objectively account for outcome associations, which is one of the limitations of this work. Even though prioritizing outcomes based on weights could provide a basic solution to the multiple response problem, it could be useful to investigate strategies for prioritizing outcomes while accounting for the natural correlation among responses. It is
a worthwhile endeavor to find data-driven guidelines on the weight specification in future research.

**Acknowledgments**

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The research work by Chathura Siriwardhana was partially supported by grant U54MD007601 from the National Institutes of Health.

**ORCID**

Chathura Siriwardhana [http://orcid.org/0000-0002-5424-003X](http://orcid.org/0000-0002-5424-003X)

K.B. Kulasekera [http://orcid.org/0000-0001-7938-7217](http://orcid.org/0000-0001-7938-7217)

Somnath Datta [http://orcid.org/0000-0003-4381-1842](http://orcid.org/0000-0003-4381-1842)

**References**

[1] O.O. Aalen, *A linear regression model for the analysis of lifetimes*, Stat. Med. 8 (1980), pp. 907–925.

[2] O.O. Aalen, *A model for non-parametric regression analysis of counting processes*, in: Lecture notes on mathematical statistics and probability, W. Klonecki, A. Kozek, and J. Rosiski, eds., Vol. 2, Springer, New York, 1980, pp. 1–25.

[3] T. Cai, L. Tian, P.H. Wong, and L.J. Wei, *Analysis of randomized comparative clinical trial data for personalized treatment selections*, Biostatistics 12 (2011), pp. 270–282.

[4] Y. Cui, R. Zhu, and M. Kosorok, *Tree based weighted learning for estimating individualized treatment rules with censored data*, Electron. J. Stat. 11 (2017), pp. 3927–3953.

[5] Y. Cui, M. Kosorok, E. Sverdrup, S. Wager, and R. Zhu, *Estimating heterogeneous treatment effects with right-censored data via causal survival forests*. 2021. Available at arXiv:2001.09887.

[6] S. Datta, *Estimating the mean life time using right censored data*, Stat. Methodol. 2 (2005), pp. 65–69.

[7] T.R. Fleming and J.H. Powers, *Biomarkers and surrogate endpoints in clinical trials*, Stat. Med. 31 (2012), pp. 2973–2984. doi:10.1002/sim.5403

[8] Y. Geng, H.H. Zhang, and W. Lu, *On optimal treatment regimes selection for mean survival time*, Stat. Med. 34 (2015), pp. 1169–1184.

[9] A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, and T. Hothorn, *mvtnorm: Multivariate Normal & t Distributions*. R package version 1.0–3. 2017. September 29 2020. Available at [https://cran.r-project.org/web/packages/mvtnorm/](https://cran.r-project.org/web/packages/mvtnorm/).

[10] S.M. Hammer, D.A. Katzenstein, M.D. Hughes, H. Gundacker, R.T. Schooley, R.H. Haubrich, W.K. Henry, M.M. Lederman, J.P. Phair, M. Niu, and M.S. Hirsch, *A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter*, N. Engl. J. Med. 335 (1996), pp. 1081–1090.

[11] H. Ichimura, P. Hall, and W. Hardle, *Optimal smoothing in single index models*, Ann. Stat. 21 (1993), pp. 157–178.

[12] R. Jiang, W. Lu, R. Song, M.G. Hudgens, and S. Napravnik, *Doubly robust estimation of optimal treatment regimes for survival data with application to an hiv/aids study*, Ann. Appl. Stat. 11 (2017), pp. 1763–1786.
[13] R. Jiang, W. Lu, R. Song, and M. Davidian, On estimation of optimal treatment regimes for maximizing t-year survival probability, J. R. Stat. Soc. B 79 (2017), pp. 1165–1185.

[14] P. Johnson, W. Greiner, I. Al-Dakak, and S. Wagner, Which metrics are appropriate to describe the value of new cancer therapies? Biomed. Res. Int. 2015 (2015), p. 9. doi:10.1155/2015/865101

[15] M. Juraska, P.B. Gilbert, X. Lu, M. Zhang, D. Davidianet, and A.A. Tsiatis, SPE 2trial: Semi parametric efficient estimation for a two-sample treatment effect. R package version 1.0.4 2012. September 29 2017. Available at http://cran.r-project.org/package=spe2trial.

[16] M.R. Kosorok and E.E. Moodie (eds.), Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA, 2016. Available at https://epubs.siam.org/doi/pdf/10.1137/1.9781611974188.

[17] H. Koul, V. Susarla, and J.V. Ryzin, Regression analysis with randomly right-censored data, Ann. Statist. 9 (1981), pp. 1276–1288.

[18] W. Liang and Y. Yan, Empirical likelihood inference in randomized controlled trials with high-dimensional covariates. 2020. Available at arXiv:2010.01772.

[19] X Lu and A.A. Tsiatis, Improving the efficiency of the logrank test using auxiliary covariates, Biometrika 95 (2008), pp. 679–694.

[20] S.A. Murphy, Optimal dynamic treatment regimes, J. R. Stat. Soc. Series B Stat. Methodol. 65 (2003), pp. 331–355. doi:10.1111/1467-9868.00389

[21] V. Pihur, S. Datta, and S. Datta, Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach, Bioinformatics 23 (2007), pp. 1607–1615.

[22] V. Pihur, S. Datta, and S. Datta, RankAggreg, an R package for weighted rank aggregation, BMC Bioinformatics. 10 (2009), pp. 62.

[23] G.A. Satten and S. Datta, The Kaplan–Meier estimator as an inverse-probability-of-censoring weighted average, Am Stat. 55 (2001), pp. 207–210.

[24] G. Satten, S. Datta, and J. Robins, Estimating the marginal survival function in the presence of time dependent covariates, Stat. Probab. Lett. 54 (2001), pp. 397–403.

[25] C. Sirirwardhana, Semi-parametric methods for personalized treatment selection and multi-state models, Doctoral dissertation. University of Louisville, KY USA. 2016. September 29 2020. Available at https://ir.library.louisville.edu/cgi/preview.cgi?article=3393&context=etd.

[26] C. Sirirwardhana and K.B. Kulasekera, Personalized treatment plans with multivariate outcomes, Biometrical Journal 62 (2020), pp. 1973–1985.

[27] C. Sirirwardhana, K.B. Kulasekera, and S. Datta, Flexible semi-parametric regression of state occupational probabilities in a multistate model with right-censored data, Lifetime Data Anal. 24 (2018), pp. 464–491. doi:10.1007/s10985-017-9403-6

[28] C. Sirirwardhana, M. Zhao, S. Datta, and K. Kulasekera, A probability based method for selecting the optimal personalized treatment from multiple treatments, Stat. Methods. Med. Res. 28 (2019), pp. 749–760.

[29] C. Sirirwardhana, S. Datta, and K.B. Kulasekera, Selection of the optimal personalized treatment from multiple treatments with multivariate outcome measures, J. Biopharm. Stat. 30 (2020), pp. 462–480. doi:10.1080/10543406.2019.1684304

[30] L.L.C Statisticat, LaplacesDemon: Complete Environment for Bayesian Inference. Bayesian-Inference.com. R package version 16.0.1. 2017. September 29 2020. Available at https://cran.r-project.org/web/packages/LaplacesDemon/.

[31] H. Streeck and D.F. Nixon, T cell immunity in acute HIV-1 infection, J Infect Dis. 202 (2010), pp. 302–308.

[32] Y. Tan, X. Wen, W. Liang, and Y. Yan, Empirical likelihood weighted estimation of average treatment effects. 2020. Available at arXiv:2008.12989.

[33] L. Tian, L. Zhao, and L. Wei, Predicting the restricted mean event time with the subject’s baseline covariates in survival analysis, Biostatistics 15 (2014), pp. 222–233.

[34] A.A. Tsiatis, M. Davidian, M. Zhang, and X. Lu, Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach, Stat. Med. 27 (2008), pp. 4658–4677.
[35] A.G. Unnikrishnan, A. Bhattacharyya, M.P. Baruah, B. Sinha, M. Dharmalingam, and P.V. Rao, *Importance of achieving the composite endpoints in diabetes*, Indian J. Endocrinol. Metab. 17 (2013), pp. 835–843. doi: 10.4103/2230-8210.117225
[36] A. Vazquez, *Optimization of personalized therapies for anticancer treatment*, BMC Syst. Biol. 7 (2013), p. 31. doi: 10.1186/1752-0509-7-31
[37] L.J. van’t Veer and R. Bernards, *Enabling personalized cancer medicine through analysis of gene-expression patterns*, Nature 452 (2008), pp. 564–570.
[38] M.P. Wand and M.C. Jones, *Kernel Smoothing*, Chapman & Hall London, 1995.