Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
2024, Vol. 33(5) 794–806

DOI: 10.1177/09622802241236954
journals.sagepub.com/home/smm
Sage

A matching-based machine learning approach to estimating optimal dynamic treatment regimes with time-to-event outcomes

Xuechen Wang[1], Hyejung Lee[1], Benjamin Haaland[1], Kathleen Kerrigan[2], Sonam Puri[2], Wallace Akerley[2] and Jincheng Shen[1]

Abstract

Observational data (e.g. electronic health records) has become increasingly important in evidence-based research on dynamic treatment regimes, which tailor treatments over time to patients based on their characteristics and evolving clinical history. It is of great interest for clinicians and statisticians to identify an optimal dynamic treatment regime that can produce the best expected clinical outcome for each individual and thus maximize the treatment benefit over the population. Observational data impose various challenges for using statistical tools to estimate optimal dynamic treatment regimes. Notably, the task becomes more sophisticated when the clinical outcome of primary interest is time-to-event. Here, we propose a matching-based machine learning method to identify the optimal dynamic treatment regime with time-to-event outcomes subject to right-censoring using electronic health record data. In contrast to the established inverse probability weighting-based dynamic treatment regime methods, our proposed approach provides better protection against model misspecification and extreme weights in the context of treatment sequences, effectively addressing a prevalent challenge in the longitudinal analysis of electronic health record data. In simulations, the proposed method demonstrates robust performance across a range of scenarios. In addition, we illustrate the method with an application to estimate optimal dynamic treatment regimes for patients with advanced non-small cell lung cancer using a real-world, nationwide electronic health record database from Flatiron Health.

Keywords

Dynamic treatment regime, time-to-event outcomes, censored data, matching, machine learning, electronic health record data, non-small cell lung cancer

1 Introduction

In the past decade, with the rapid development of technology and the increasing use of computers and mobile devices, a substantial volume of health-related data has been collected and stored. Among the various data sources, Electronic health records (EHRs) have become increasingly important in healthcare research, providing real-time, patient-centered records. While EHR data has evolved into a powerful tool for researchers, its inherent complexity introduces significant challenges to them. This complexity has inspired growing interest in developing appropriate statistical methods to facilitate evidence-based research harnessing EHRs. Patients with certain types of cancer, for example, advanced non-small cell lung

[1]Department of Population Health Sciences, Division of Biostatistics, University of Utah, Salt Lake City, UT, USA
[2]Department of Internal Medicine, Division of Oncology, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA

Corresponding author:
Xuechen Wang, Department of Population Health Sciences, Division of Biostatistics, University of Utah, Salt Lake City, UT 84108, USA.
Email: xuechen.wang@hsc.utah.edu

cancer (NSCLC), have a poor prognosis. A crucial clinical challenge is to customize the treatments according to patients' disease progress in order to elongate overall survival (OS). This problem is related to the concept of a dynamic treatment regime (DTR), which is a sequence of decision rules determining how to individualize treatments to patients based on their evolving clinical characteristics and treatment history.[1] It is usually of great interest to understand the optimal DTR that produces the best clinical results when patients adhere to tailored treatment decisions.

Generally, two primary directions can be considered to identify optimal DTRs. The first involves the design and execution of sequential multiple assignment randomized trials (SMARTs).[2,3] However, these trials usually require large sample sizes and extensive resources and may remain underpowered for analyses beyond the first stage comparisons. Particularly, SMARTs aiming to compare various complex guideline-concordance treatment sequences for advanced cancers are difficult to fund and implement. Generalization of results from SMARTs is questionable because of the stringent inclusion criteria, for example, disease stage and comorbidity restrictions. The other direction involves leveraging observational data (e.g. EHR data), which provides comprehensive clinical information about the general population, to inform future decisions. Using such a rich and cost-effective resource to conduct evidence-based research on personalized treatment strategies has gained growing popularity. The paramount challenge in using EHR data lies in the necessity for sophisticated statistical approaches to account for complex confounding mechanisms. Time-to-event outcomes (e.g. OS), which are commonly of primary interest in health research, are subject to right-censoring, and this feature further complicates the estimation procedure.

Statistical methods have been developed to identify optimal DTRs from observational data under a causal inference framework. Common approaches include Q-, A-, and O-learning.[4–8] To properly handle the confounding problem in observational data, a common practice is to employ inverse propensity score weighting (IPW).[9] In addition, special handling is usually needed to extend these approaches to accommodate time-to-event outcomes, inverse probability of censoring weights (IPCWs) are often used to account for censoring of time-to-event outcomes.[10–12] Theoretically, both Q- and A-learning depend heavily on the correct specification of parametric models. Although O-learning approaches based on so-called direct optimization methods do not require an explicit outcome model, they can still be vulnerable to the misspecification of the propensity score model. Recent developments in employing machine learning methods for propensity score estimation have partially relieved concerns of model misspecification,[9,13] but those approaches can lead to extreme estimates with high variability. Especially, when considering multiple treatment decision stages, we need to be more careful about extreme weights as weights will be multiplied across stages and grow exponentially. To summarize, the IPW-based approaches are generally useful and effective, but they could face difficulties and lead to biased results in certain situations. As an alternative to IPW, matching-based approaches have been successfully applied to handle confounding in average treatment effect estimation and have shown effectiveness when exist extreme weights.[15,14] Recently, a matching-based machine learning algorithm, known as M-learning, has demonstrated feasibility in individualized treatment rule estimation for a single decision stage with continuous or binary outcomes with nice asymptotic properties.[16]

In this article, we propose the Matched Learning on DTR for Survival algorithm (MLSurv), which extends the M-learning methodology to estimate optimal two-stage DTRs for time-to-event endpoints using EHR data. The proposed approach utilizes matching methods to address both confounding and censoring. We adopt a backward induction framework[19,18,17] and implement a weighted support vector machine (SVM) to estimate DTRs that optimize the time-to-event outcomes. Our algorithm, MLSurv, should relieve concerns of model misspecification and be efficient in situations where existing methods may suffer from extreme weights. The remainder of this article is structured as follows. In Section 2, we introduce the proposed MLSurv algorithm. Section 3 presents the results of simulation studies evaluating the performance of MLSurv and comparing it with another robust algorithm. Section 4 illustrates the proposed algorithm in an application to patients with advanced NSCLC from the nationwide multi-institutional Flatiron Health de-identified EHR-derived database. In Section 5, we engage in a thorough discussion covering the limitations of MLSurv and future directions.

2 Methodology

In this section, we will elaborate on the proposed MLSurv, which targets time-to-event outcomes and aims to estimate the optimal DTR that can lead to the best outcome for a given patient. For the explanation purpose, we focus on the OS outcome in this article. Similar principles will still hold when the goal is to minimize a harmful outcome.

2.1 Notation and assumptions

Consider an observational study with a cohort of N subjects, in which the maximum number of stages an individual might receive is K. As almost all variables are individual-specific, the subscript specifying individuals is dropped when

no confusion arises. For simplicity, we consider binary treatment decisions at each stage. For $k = 1, \ldots, K$, let $A_k \in \{-1, 1\}$ be the observed treatment assignment at the beginning of the kth stage, $\mathbf{X}_k$ be a vector of covariates measured at stage k prior to the assignment of A_k. We use an overbar to denote the variable history up to the corresponding time, e.g. $\overline{A}_k = (A_1, A_2, \ldots, A_k)$. Let η_{ik} denote an indicator variable with value 1 if an individual i ($i \in \{1, \ldots, N\}$) entered stage k ($k \in \{1, \ldots, K\}$) and 0 otherwise. Clearly, all subjects should have $\eta_{i1} = 1$. This indicator is introduced by Simoneau et al.,[20] and adopted in this article. Let $\tilde{T}_k$ denote the survival time during the interval of treatment A_k (within stage k), with $\tilde{T}_k = 0$ when $\eta_k = 0$. Then, based on the indicator of η_k, the OS is the accumulated survival time in all stages and can be expressed as $\tilde{T} = \sum_{k=1}^{K} \eta_k \tilde{T}_k$. We employ the counterfactual framework and use a superscript to denote counterfactual outcomes, e.g. $\tilde{T}^{\overline{a}_K} = \sum_{k=1}^{K} \eta_k \tilde{T}_k^{\overline{a}_k}$ represents the counterfactual survival time had an individual followed the treatment sequence of $\overline{a}_K = (a_1, a_2, \ldots, a_K)$, where $\tilde{T}_k^{\overline{a}_k}$ is the counterfactual survival time within the kth stage. Let C denote the censoring time, $Y = min(\tilde{T}, C)$ denote the observed OS time, and $\delta = I\{\tilde{T} \leq C\}$ denote the event indicator. Meanwhile, let Y_k denote the observed survival time within stage k. Notably, $Y_k = \tilde{T}_k$ if a subject is not censored within stage k, and $Y_k < \tilde{T}_k$ if a subject is censored within stage k. We use H_k to denote an individual's clinical history just prior to the kth treatment decision, $H_k = (\eta_{i1}, \mathbf{X}_{i1}, A_{i1}, \tilde{T}_{i1}, \ldots, \eta_{ik}, \mathbf{X}_{ik})$, with its realization denoted as h_k. Meanwhile, when $\eta_{ik} = 0$, A_{ik} and $\mathbf{X}_{ik}$ are not available. A DTR is a set of treatment decision rules, $\boldsymbol{g} = \{g_1(h_1), \ldots, g_K(h_K)\} \in \mathcal{G}$, where $\mathcal{G}$ represents the set of possible treatment regimes. The decision rule at a specific stage k, $g_k(h_k)$, is a map from the current history to treatments $\{-1, 1\}$.

To estimate the optimal DTR, we want to connect observational data to the counterfactual quantities. Following related literature, we require several identification assumptions: (1) Stable unit treatment value assumption (SUTVA),[21] which requires the outcome of an individual is not influenced by the treatment applied to other individuals. (2) Consistency,[22] that the counterfactual outcome under the observed treatment is the observed outcome. (3) Sequential ignorability,[22] which extends the ignorability assumption for point treatments to longitudinal settings, further requiring that the treatment assignment at a given stage k is independent with future covariates. It can be expressed as $\{\sum_{j \geq l}^{K} \tilde{T}_j^{\overline{a}_j} : l = k, \ldots, K\} \perp A_k | H_k, \eta_1, \ldots, \eta_k$. (4) Censoring at random (Gill 1997), which assumes that at the beginning of each stage, given the current history, the future probability of censoring does not depend on future outcomes, $\{\sum_{j \geq l}^{K} \tilde{T}_j^{\overline{a}_j} : l = k, \ldots, K\} \perp \delta | H_k, \eta_1, \ldots, \eta_k$.

2.2 Definition of the Optimal DTR

The value function that assesses the overall benefit of a given DTR ($\boldsymbol{g}$) is defined as the expected survival time when $\boldsymbol{g}$ is followed by the population of interest. Due to right censoring, accurate estimation of the tail distribution may be challenging, particularly for patients with advanced cancer and having a long follow-up period (e.g. 5 years) from the EHR database, where observed events are sparsely distributed. Hence, we choose to use restricted survival time hereafter (denoted as T) as an alternative outcome in this article to enable fair comparisons of outcomes,[23] and follow convention to choose τ to denote the study end. Therefore, $T = min(\tilde{T}, \tau)$, where $\tilde{T}$ is the true survival time.[24] Then, we define the value function as: $V(\boldsymbol{g}) = E(T^{\boldsymbol{g}})$. The optimal DTR ($\boldsymbol{g}^{\text{opt}} = \{g_1^{\text{opt}}(h_1), \ldots, g_K^{\text{opt}}(h_K)\}$) is the regime that maximizes the value function:

$$\boldsymbol{g}^{\text{opt}} = \arg\max_{\boldsymbol{g} \in \mathcal{G}} E(T^{\boldsymbol{g}})$$

2.3 Matching based procedure

Here, we propose a matching-based value function to estimate the optimal DTR by solving a classification problem, which is amenable to a variety of off-the-shelf machine learning algorithms (e.g. SVM, random forest).

The matched set for a subject i who entered stage k can be denoted as: $\mathcal{M}_{ik} = \{j : \eta_{jk} = 1, A_{jk} = -A_{ik}, d(H_{jk}, H_{ik}) \leq \epsilon_{ik}\}$, which contains subjects entering stage k, receiving the alternative treatment, and with similar history as subject i. The similarity is controlled by a proper distance metric ($d(\cdot, \cdot)$), which is defined on the history space. There are a variety of ways to identify the $\mathcal{M}_{ik}$, for example, we can specify a threshold ϵ_{ik} first and include all the qualified matches for every index patient to determine the size of $\mathcal{M}_{ik}$. Throughout this study, we choose to use another method—the nearest neighbor with a one-to-one match to find $\mathcal{M}_{ik}$, thus ϵ_{ik} is the minimal distance between subject i and any other subject who entered stage k and received the alternative treatment. The underlying idea of estimating the optimal treatment rule is that when two subjects with matched clinical history receive different treatments at stage k, the subject receiving the optimal treatment is more likely to have a better clinical outcome. This rationale establishes that the penalty for misclassification depends on the loss of benefit in clinical outcome if not receiving the optimal treatment.

When the endpoint is survival time, right censoring complicates the optimization of the value function. IPCW has been widely used to handle censoring. However, IPCW is recognized to be vulnerable to model misspecification. Here, we employ a matching approach to generate a pseudo-population without right censoring by imputing the censored survival time with the restricted survival time when the two subjects share similarities in the covariate and treatment history. Specifically, for a subject i who is censored in stage k and at Y_{ik} (Note: $Y_{ik} < \tilde{T}_{ik}$, when $\delta_i = 0$), we can find a matched set including subjects who also enter stage k and are still at risk at Y_{ik}, with observed events and similar history to subject i up until Y_{ik}, to get a potential survival time in stage k for subject i. Denote this matched set as $\mathcal{MC}_i = \{j : Y_{jk} > Y_{ik}, \delta_j = 1, d(H_i, H_j) \leq \epsilon_i, A_{ik} = A_{jk}\}$, where k is the last stage for subject i in this notation. The nearest neighbor with the one-to-one match method is applied to find the $\mathcal{MC}_i$. Notably, this matching step for imputing censored survival will be repeated at each stage. Compared to the matching step of identifying $\mathcal{M}_{ik}$, the major differences here include the utilization of treatment at the current stage and the requirement that patients must still be at risk at time Y_{ik}.

2.4 Matching based classification procedure

Estimation of the optimal DTRs in our approach relies on backward induction (Sutton and Barto[25]), which is often utilized in a sequential decision-making process. Without loss of generality, we consider two stages of treatment as an example here, and let $a_1^{\text{opt}} = g_1^{\text{opt}}(h_1)$, $a_2^{\text{opt}} = g_2^{\text{opt}}(h_2)$ denote the optimal treatment decision rules for the first and second stage. Extension to more than two stages is conceptually straightforward, but it introduces a significant increase in the complexity of notations and mathematical expressions. In practice, it may become infeasible due to a rapid decline in sample sizes as it goes into later stages.

For a two-stage study, the first step is to estimate the optimal second stage treatment rule depending on a value function $V_2(g_2(h_2)) = V_2(a_2) = E[T_2^{\bar{a}_2}]$ of the expected counterfactual restricted survival time from the second stage forward. The next step after having a_2^{opt} is to construct a counterfactual restricted survival time had all subjects who entered the second stage received the optimal second stage treatment. Then, the last step is to estimate the optimal first stage treatment rule depending on a value function $V_1(g_1(h_1), g_2^{opt}(h_2)) = V_1(a_1, a_2^{\text{opt}}) = E[T^{a_1, a_2^{\text{opt}}}]$ of the expected counterfactual restricted OS from the first stage onward with the optimal treatment in the second stage following a_2^{opt}. In matching steps, a variety of techniques can be considered, such as one-to-one matching or one-to-many matching. We chose to use the one-to-one nearest neighbor matching method with replacement in this article. In detail, the estimation procedure of the proposed MLSurv method can be described by the following algorithm:

1. Within the subcohort of patients who entered the second stage ($S_2 = \{i : \eta_{i2} = 1\}$), for a subject i who had censored restricted survival time, find a matched set containing subjects who are similar to subject i on covariates and treatment history, $\mathcal{MC}_{i2} = \{j : Y_{j2} > Y_{i2}, \delta_j = 1, d(H_{i2}, H_{j2}) \leq \epsilon_{i2}, A_{i2} = A_{j2}\}$. Then, $T_{i2} = \frac{1}{|\mathcal{MC}_{i2}|} \sum_{j \in \mathcal{MC}_{i2}} T_{j2}$.
2. Based on the matched set of $\mathcal{M}_{i2} = \{j : \eta_{j2} = 1, A_{j2} = -A_{i2}, d(H_{j2}, H_{i2}) \leq \epsilon_{i2}\}$ for a subject i who entered the second stage, specify the matching based value function for stage two, which is the expected restricted survival time given that the second stage treatment assignments follow regimen $g_2(h_2)$:

$$
\begin{aligned}
V_2(g_2(h_2)) &= \frac{1}{|S_2|} \sum_{i \in S_2} \left\{ I(g_2(h_{i2}) = a_{i2}) * T_{i2} + I(g_2(h_{i2}) \neq a_{i2}) * \frac{1}{|\mathcal{M}_{i2}|} \sum_{j \in \mathcal{M}_{i2}} T_{j2} \right\} \\
&= \frac{1}{|S_2|} \sum_{i \in S_2} \left\{ \left[\left(\frac{1}{|\mathcal{M}_{i2}|} \sum_{j \in \mathcal{M}_{i2}} T_{j2} \right) - T_{i2} \right] * I(g_2(h_{i2}) \neq a_{i2}) + T_{i2} \right\} \\
&= \frac{1}{|S_2|} \sum_{i \in S_2} \{ I(f_2(h_{i2}) * a_{i2} * \text{sign}(B_{i2}) \leq 0) * |B_{i2}| + T_{i2} \}
\end{aligned}
$$

where $B_{i2} = (\frac{1}{|\mathcal{M}_{i2}|} \sum_{j \in \mathcal{M}_{i2}} T_{j2}) - T_{i2}$.

The optimal second stage treatment rule should be the one that maximizes this value function. From the classification perspective, the maximization problem can be reformulated as a minimization of misclassification errors. Let $g_2(h_2) = \text{sign}(f_2(h_2))$ for a second stage decision function f_2. Equivalently, to obtain $g_2^{\text{opt}}(h_2)$, we can minimize this objective function:

$$
\frac{1}{|S_2|} \sum_{i \in S_2} I \left\{ f_2(h_{i2}) * a_{i2} * \text{sign} \left(T_{i2} - \frac{1}{|\mathcal{M}_{i2}|} \sum_{j \in \mathcal{M}_{i2}} T_{j2} \right) \leq 0 \right\} * |T_{i2} - \frac{1}{|\mathcal{M}_{i2}|} \sum_{j \in \mathcal{M}_{i2}} T_{j2}|
$$

Specifically, it has the form of loss function for a weighted classification problem where the weights are constructed using matched pairs. By estimating the selected classifier, we can derive the estimated decision rule for stage two, $\hat{g}_2^{\text{opt}}(H_2)$. In this article, we chose to use the SVM method employing the hinge loss function to create a soft margin (see details in Zhao et al.[26]; Wu et al.[16]).

3. With the estimated optimal second stage treatment $\hat{a}_2^{\text{opt}}$, construct the pseudo-outcome for stage 1 as:

$$T^{a_1,\hat{a}_2^{opt}} = T^* = T_1 + \eta_2 * T_2^{\hat{a}_2^{opt}}$$

4. For a subject i who did not enter the second stage and had censored restricted survival time, find a matched set containing subjects who are similar to subject i on covariates and treatment history, $\mathcal{MC}_{i1} = \{j : Y_{j1} > Y_{i1}, \delta_j = 1, d(H_{i1}, H_{j1}) \leq \epsilon_i, A_{i1} = A_{j1}\}$. Then, $T_i^* = \frac{1}{|\mathcal{MC}_{i1}|} \sum_{j \in \mathcal{MC}_{i1}} T_j^*$. Note patients in the matched set can enter the second stage. For a subject i who did not enter the second stage and had the event observed, $T_i^* = Y_{i1} = T_{i1}$.
5. Then, define the value function for the first stage using the constructed pseudo-outcome T^*:

$$V_1(g_1(h_1), g_2^{opt}(h_2)) = \frac{1}{N} \sum_i \left\{ \left[\left(\frac{1}{|\mathcal{M}_{i1}|} \sum_{j \in \mathcal{M}_{i1}} T_j^* \right) - T_i^* \right] * I(g_1(h_{i1}) \neq a_{i1}) + T_i^* \right\}$$

Similar to step 2, the weighted classification objective function can be used to estimate the optimal first stage decision rule, $\hat{g}_1^{\text{opt}}(h_1)$.

The underlying idea of the proposed method is to utilize matching based procedures to obtain the counterfactual restricted survival time, and then apply weighted classification machine learning methods to estimate the optimal DTRs, where weights depend on counterfactual outcomes.

3 Simulations

We conduct simulation studies to evaluate the performance of the proposed MLSurv method. We adopt similar data-generating mechanisms as those outlined in the study by Simoneau, Moodie et al.[20] to facilitate comparison with their method (i.e. DWSurv). In contrast to the matching-based method and classification approach, DWSurv models the censoring probability and treatment probability. Subsequently, it estimates the DTR based on weighted regression models. When applying the DWSurv, we correctly specify the treatment assignment model and the censoring model. Meanwhile, the outcome models include all the involved variables in a linear form. We compare the performance of different methods via two metrics: Correct classification rate (CCR), which refers to the proportion of individuals that are accurately classified to the true optimal treatment group by the estimated DTRs, and the estimated mean restricted OS time $E(T^{\hat{a}_1^{\text{opt}},\hat{a}_2^{\text{opt}}})$.

In the matching steps, we choose to use a one-to-one matching with replacement based on the shortest Euclidean distance. The misclassification penalty and kernel bandwidth tuning parameters are selected through a 5-fold cross-validation process. A training data set with a sample size of 1000 is used to estimate the optimal DTRs. Then, an independent testing data set of 1000 is used to evaluate the estimation accuracy. We repeat this process 300 times. For computational efficiency, the testing dataset is randomly drawn from a separate validation dataset with a sample size of 10,000, specifically generated and reserved for testing purposes.

3.1 Simulation setting

In alignment with the paper of DWSurv, we consider one covariate at the beginning of stage one $\{X_1\}$ following a uniform distribution on $(0.2, 1)$, and two more independent covariates at the beginning of stage two $\{X_2, X_3\}$ following uniform distributions on $(0.5, 2)$ and $(0.1, 1.3)$. This study explores two scenarios: a nonlinear setting and a linear setting.

3.1.1 Nonlinear setting

For an individual i, the first stage treatment assignment A_{i1} is simulated through a Bernoulli distribution with $P(A_{i1}) = \text{expit}(-1 + 2X_{i1})$ where expit(x) is defined as $\text{expit}(x) = \exp(x)/(1 + \exp(x)), x \in R$. The second stage treatment assignment A_{i2} is simulated through another Bernoulli distribution with $P(A_{i2}) = \text{expit}(3 - 2X_{i2})$. The allocation ratio for both treatment choices in both stages is approximately 1:1. The indicator of whether individual i enters the second stage (η_i) and the indicator of whether the death event is observed (δ_i) are independently generated from Bernoulli distributions with probability 0.8 and 0.9, respectively.

For individuals who enter the second stage ($\eta_2 = 1$) and experience the event ($\delta = 1$), the survival time within the second stage is simulated using an accelerated failure time (AFT) model similar to the DWSurv paper.[20] Fundamentally, DWSurv requires parametric modeling of the outcome. Despite the robustness provided through the estimation procedure, the performance of DWSurv still depends on the correct specification of the outcome model. So the simulation setting may favor DWSurv when the regression models are correctly specified, as we will see in the linear setting case. Even though, the simulation settings still can be used to facilitate comparisons between the two methods to illustrate the performance of MLSurv.

With the covariates at the beginning of stage 2, the survival time within the second stage is:

$$\begin{aligned}\log(\tilde{T}_{i2}) = {} & 0.6 - 0.5X_{i1} + 0.2X_{i2} + 0.1X_{i3} \\ & + A_{i2}[-0.55 + 0.45X_{i1} + (X_{i2} - 1.2)^2 + (X_{i3} - 0.7)^2] + \epsilon_{i2}\end{aligned}$$

where ϵ_{i2} follows a normal distribution $N(0, 0.1^2)$. Then, the restricted survival time in stage two is $T_{i2} = \min(\tilde{T}_{i2}, \tau)$ and the true optimal stage two treatment A_{i2}^{opt} is determined by $\text{sign}(-0.55+0.45X_{i1}+(X_{i2}-1.2)^2+(X_{i3}-0.7)^2)$. The counterfactual survival time within stage two had an individual i who enters the second stage and receives the optimal stage two treatment can be expressed as

$$\log(\tilde{T}_{i2}^{\text{opt}}) = \log(\tilde{T}_{i2}) + (A_{i2}^{\text{opt}} - A_{i2})(-0.55 + 0.45X_{i1} + (X_{i2} - 1.2)^2 + (X_{i3} - 0.7)^2)$$

Then, $T_{i2}^{\text{opt}} = \min(\tilde{T}_{i2}^{\text{opt}}, \tau)$.

For individuals with the event observed, $\delta = 1$, the counterfactual OS with optimal stage two treatment is generated from another AFT model as:

$$\log(\tilde{T}_i) = 2.25 - 0.7X_{i1} + 0.6X_{i1}^2 + A_{i1}[-1.63 - 3.85X_{i1}^2 + 5.26X_{i1}] + \epsilon_{i1}$$

where ϵ_{i1} follows a normal distribution $N(0, 0.08^2)$. The restricted OS time is $T_i = \min(\tilde{T}_i, \tau)$.

3.1.2 Linear setting

In this setting, the first stage treatment assignment A_{i1} is simulated via a Bernoulli distribution with $P(A_{i1}) = \text{expit}(-0.6 + X_{i1})$, and the second stage treatment assignment A_{i2} is simulated via another Bernoulli distribution with $P(A_{i2}) = \text{expit}(2.6 - 1.8X_{i2})$. The allocation ratio is approximately 1:1 as well. η_2 and δ are simulated in the same manner as in the non-linear setting.

For individuals with $\eta_2 = 1$ and $\delta = 1$, an AFT model is utilized to generate the survival time within the second stage as:

$$\log(\tilde{T}_{i2}) = 0.8 + 0.6X_{i1} - 0.45X_{i2} + 0.15X_{i2}^2 + 0.55X_{i3} + A_{i2}(1.2 + 0.6X_{i1} - X_{i2} - 0.4X_{i3}) + \epsilon_{i2}$$

Clearly, the true optimal stage two treatment A_{i2}^{opt} is determined by $\text{sign}(1.2 + 0.6X_{i1} - X_{i2} - 0.4X_{i3})$. The restricted survival time on stage two is $T_{i2} = \min(\tilde{T}_{i2}, \tau)$. The counterfactual survival time within stage two had an individual i with $\eta_2 = 1$ receives the optimal stage two treatment can be computed as:

$$\log(\tilde{T}_{i2}^{\text{opt}}) = \log(\tilde{T}_{i2}) + (A_{i2}^{\text{opt}} - A_{i2})(1.2 + 0.6X_{i1} - X_{i2} - 0.4X_{i3})$$

Then, for those with $\delta = 1$, the counterfactual OS with optimal stage two treatment is simulated from the AFT model as:

$$\log(\tilde{T}_i) = 2.2 + 0.45X_{i1} + 0.2X_{i1}^2 + A_{i1}(-0.75 + 1.2X_{i1}) + \epsilon_{i1}$$

Then, $T_i = \min(\tilde{T}_i, \tau)$

Specifically, for individuals who do not enter the second stage ($\eta_2 = 0$), their restricted OS time is T_i. For individuals who enter the second stage ($\eta_2 = 1$), their restricted OS time is calculated as $T_i = T_{i1} + T_{i2}$, where $T_{i1} = T_i - T_{i2}^{\text{opt}}$. For individuals who are censored, the censoring times are generated from the uniform distribution of $U(0, T_i)$.

3.2 Simulation results

As the overarching goal of the estimated DTR is to optimize the restricted OS time, we expect to observe an improvement in the average restricted OS, assuming all patients adhere to treatments assigned by a better DTR. The true optimal restricted

Table 1. Comparison of CCR and estimated restricted survival time ($\hat{T}$) of estimated optimal DTRs. Standard errors are reported in the parentheses.

Method	First-line CCR	Second-line CCR	Overall CCR	$\hat{T}$
Nonlinear				
MLSurv	0.957 (0.028)	0.903 (0.018)	0.865 (0.030)	9.721 (0.067)
DWSurv	0.603 (0.030)	0.651 (0.016)	0.394 (0.021)	9.056 (0.100)
Linear				
MLSurv	0.956 (0.013)	0.970 (0.014)	0.927 (0.019)	16.64 (0.114)
DWSurv	0.982 (0.014)	0.996 (0.003)	0.977 (0.014)	16.66 (0.114)

* The average of true optimal restricted survival time is 9.75 in the nonlinear setting and 16.67 in the linear setting. CCR: correct classification rate; DTR: dynamic treatment regime.

survival time is 9.75 in the nonlinear setting and 16.67 in the linear setting. Table 1 shows the average CCRs of first-line (stage one) treatment, second-line (stage two) treatment, overall (both stages), and the average estimated restricted OS. The results are based on 300 simulated data sets in each setting. The MLSurv algorithm employs an SVM classifier with a radial basis function (RBF) kernel in the nonlinear setting and a linear kernel in the linear setting. Meanwhile, in both settings, the DWSurv method incorporates all variables that are used in simulating the data into the outcome models. Additionally, it correctly specifies the treatment and censoring models. When the true DTRs are nonlinear, the proposed matching-based machine learning method (MLSurv) outperforms the DWSurv. As indicated in Table 1, the MLSurv method provides an average estimated restricted survival time in the nonlinear setting that is almost the same as the true value. In contrast, the corresponding estimate from DWSurv is shorter than the true value. Specifically, the average estimated restricted OS is 9.721 (standard error (SE) = 0.067) for MLSurv and 9.056 (SE = 0.100) for DWSurv. In addition, the overall CCR is 0.865 (SE = 0.030) for MLSurv and 0.394 (SE = 0.021) for DWSurv. Although the CCRs of MLSurv are not perfect, the estimated OS of MLSurv almost equals the truth. In the linear setting, the average estimated restricted survival times from both methods closely approximate the true value. The estimated OS is 16.64 (SE = 0.114) for MLSurv and 16.66 (SE = 0.114) for DWSurv. The overall CCR is 0.927 (SE = 0.019) for MLSurv and 0.977 (SE = 0.014) for DWSurv. Obviously, as all the required models are correctly specified, the DTRs estimated by DWSurv are very close to the truth. Compared to the DWSurv, the CCRs of MLSurv are slightly lower, but the estimated OS of MLSurv is almost the same.

Figure 1 displays the actual 3D plots using data from the nonlinear and linear settings. Figures 2 and 3 show the true decision boundaries, and boundaries of estimated DTRs given by MLSurv and DWSurv in both settings. The top panel shows the decision boundaries at a fixed X_1. This plot is interpreted to provide optimal treatment suggestions for different combinations of X_2 and X_3 when X_1 is at the specified fixed value. Meanwhile, the bottom panel shows the decision boundaries at a fixed X_2, with a similar interpretation. In situations where the actual boundaries are nonlinear, and the nonlinear forms of the confounders are unclear, MLSurv with an RBF kernel yields similar boundaries as the truth, while DWSurv fails to produce such nonlinear boundaries due to misspecification in the outcome models. Furthermore, the estimated average restricted OS time from MLSurv is nearly optimal. Then, in scenarios where the actual boundaries are linear, DWSurv generates boundaries almost identical to the truth, and MLSurv also yields comparable estimated decision boundaries. Therefore, the proposed MLSurv approach can fulfill the overarching optimization goal and demonstrate robust performance in both simulation settings. Additionally, as suggested by one of the reviewers, we have investigated the scenarios in which covariates are simulated from normal distributions. Similar results are observed (data not shown), and consistent conclusions are drawn.

4 Application

Lung cancer is a leading cause of cancer mortality in the United States. Approximately 85% of the diagnosed lung cancer is NSCLC, and more than half of the NSCLC patients are already in advanced stages when they are first diagnosed and have a very poor prognosis. In recent years, advances in treatments have provided better options for a variety of patient subgroups, while making the treatment and biomarker landscape more complex. Although National Comprehensive Cancer Network (NCCN) guidelines indicate possible treatment choices for specific patients' profiling,[27] clinicians and patients still need to make a specific choice within these guideline compatible options. Notably, little data is available on optimal treatment regimes, especially optimal DTRs in patients with advanced NSCLC, and most of the time, the sequence of treatment decisions would depend on the judgment and preference of specific clinicians and patients involved. As such, the outcomes can vary significantly from case to case. Given the short duration of survival, it is essential to have a clear regime assigning treatment to a patient that produces the optimal outcome.

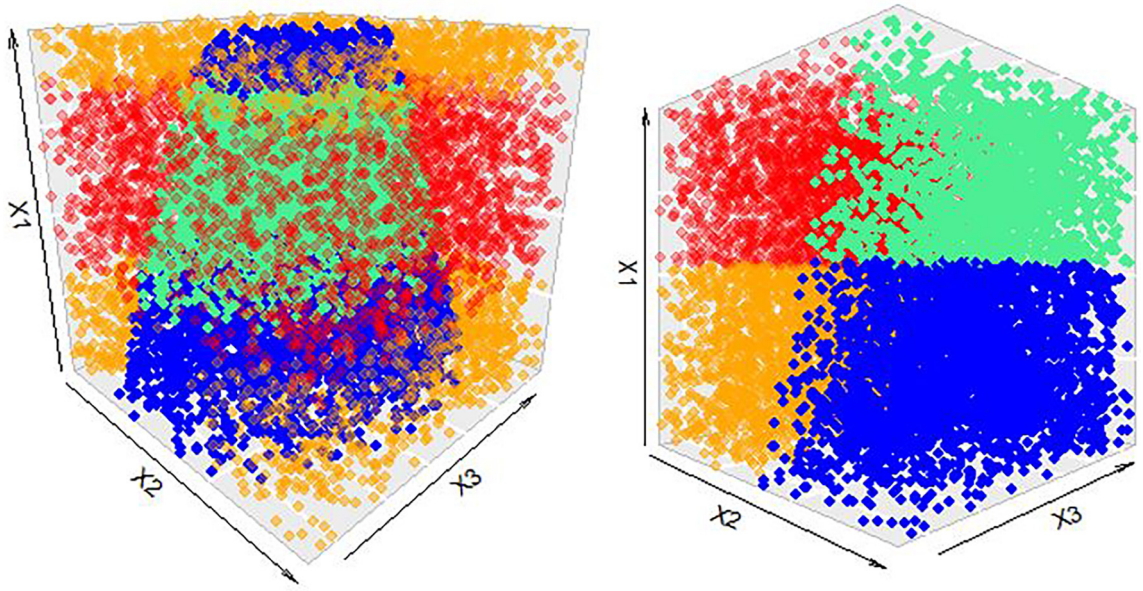

Figure 1. The three-dimensional (3D) plot of the actual simulated data in nonlinear setting (left) and linear setting (right). Blue represents $A_1 = -1, A_2 = -1$; green represents $A_1 = 1, A_2 = -1$; orange represents $A_1 = -1, A_2 = 1$; red represents $A_1 = 1, A_2 = 1$.

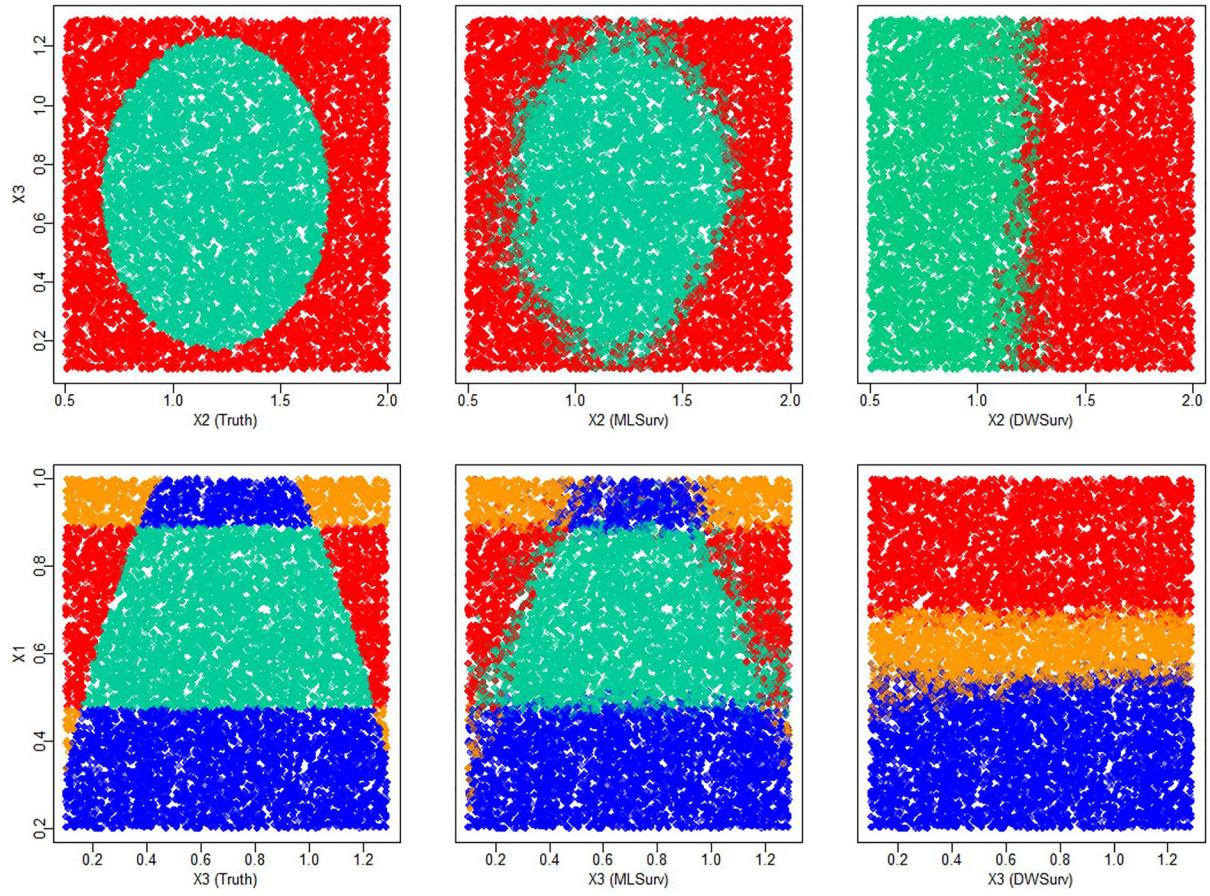

Figure 2. In nonlinear simulation setting, the true decision boundaries (left panel), the estimated decision boundaries from MLSurv method (middle panel), and from DWSurv method (right panel). X_1 was fixed for the 3 plots in the top panel, and X_2 was fixed for the 3 plots in the bottom panel. Blue represents $A_1 = -1, A_2 = -1$; green represents $A_1 = 1, A_2 = -1$; orange represents $A_1 = -1, A_2 = 1$; red represents $A_1 = 1, A_2 = 1$.

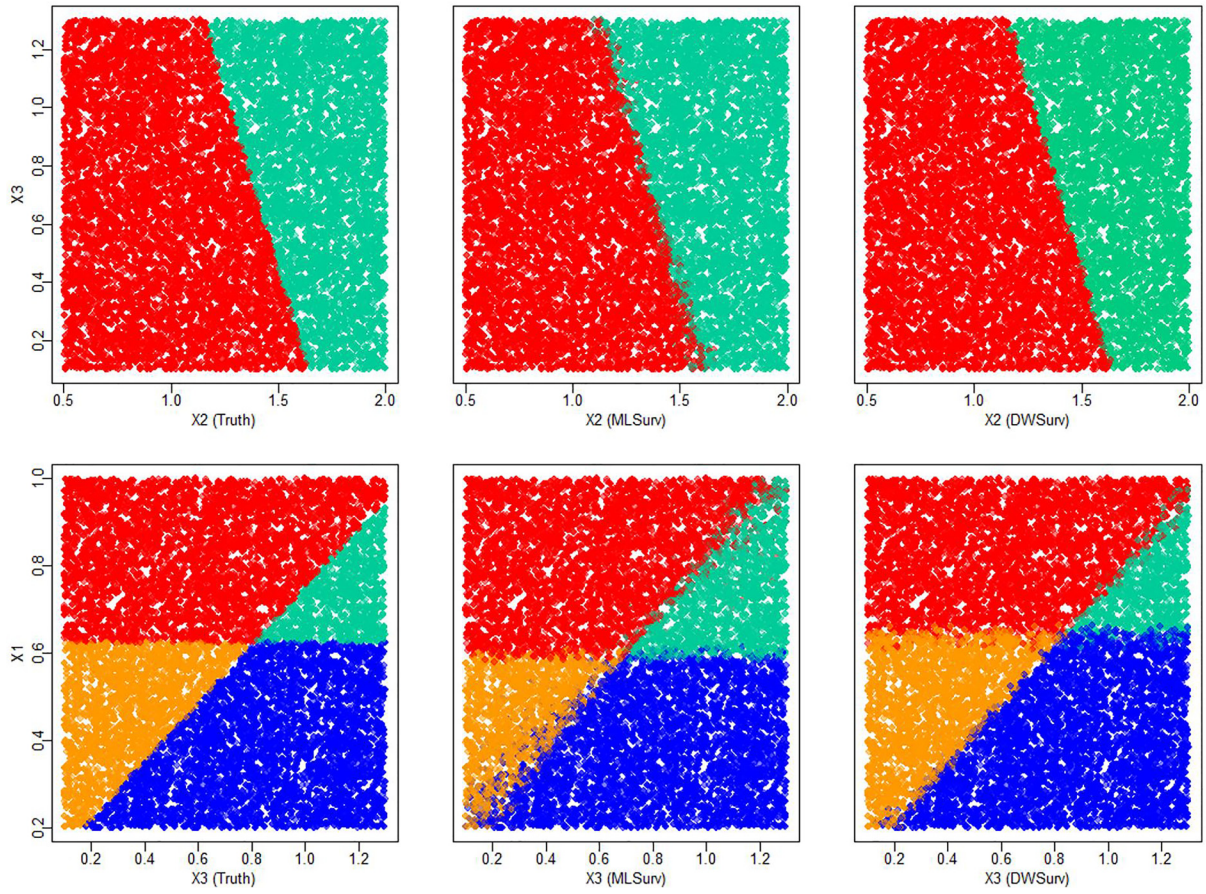

Figure 3. In linear simulation setting, the true decision boundaries (left panel), the estimated decision boundaries from MLSurv method (middle panel), and from DWSurv method (right panel). X_1 was fixed for the 3 plots in the top panel, and X_2 was fixed for the 3 plots in the bottom panel. Blue represents $A_1 = -1, A_2 = -1$; green represents $A_1 = 1, A_2 = -1$; orange represents $A_1 = -1, A_2 = 1$; red represents $A_1 = 1, A_2 = 1$.

In this study, analyses were focused on patients with advanced NSCLC lacking targetable mutations (*EGFR*, *ALK*, *BRAF*, and *ROS1* wild-type) from the USA Flatiron Health nationwide EHR, de-identified longitudinal database comprising patient-level structured and unstructured data curated via technology-enabled abstraction.[29,28] At the time we requested the data, the de-identified data originated from approximately 280 USA cancer clinics (~800 sites of care). We further restricted our study sample to patients who initiated their first-line therapy from 2017 onward and whose first-line therapy was concordant with current NCCN guidelines applicable to the target patients without contraindications for immunotherapy, i.e. first-line immunotherapy or combination chemoimmunotherapy. Normally, during a follow-up visit after the initiation of first-line therapy, if a patient presents disease progression based on imaging or experiences unacceptable toxicity, the recommendation is to consider providing him/her with a second-line therapy. Current NCCN guidelines suggest that the preferred second-line therapies are either chemotherapy or a combination of chemotherapy and monoclonal antibody. However, in this EHR database, a small subset of patients (approximately 2.5%) received clinical trial drugs or targeted therapies as the second-line treatment. For simplicity, we excluded them from our study. This may introduce a slight selection bias as the sample selection at the initiation of first-line therapy (baseline) depends on future information. Subsequently, patients were frequently monitored to assess their disease status and the effectiveness of treatment. Given the short median survival time, we focused on a restricted OS time with a pre-defined study follow-up duration of 24 months. OS endpoints were based on a composite mortality variable that aggregates EHR-derived structured and unstructured information, as well as third-party death surveillance sources.[30] A total of 3952 patients were included in this study with a median OS of 11.6 months. Among them, approximately 39% ($N = 1543$) were censored and 14.5% ($N = 572$) had second-line treatment observed. The ultimate goal was to estimate two-stage treatment decision rules to optimize the restricted OS time from the initiation of first-line therapy. To evaluate the efficacy of the estimated DTRs, we made a comparison between the average restricted OS given by the estimated DTRs, that given by the uniform regimes, and the observed average restricted overall survival.

Table 2. Comparison of estimated restricted overall survival time ($\hat{T}$) and estimated restricted survival time since the initiation of second-line ($\hat{T}_2$) among different regimes.

Regimens	$\hat{T}$	$\hat{T}_2$
MLSurv	12.47 (0.08)	8.10 (0.13)
Observed	12.03 (0.14)	7.46 (0.21)
$A_1 = 1, A_2 = 1$	12.18 (0.15)	7.94 (0.26)
$A_1 = 1, A_2 = -1$	12.10 (0.15)	7.54 (0.24)
$A_1 = -1, A_2 = 1$	11.91 (0.14)	7.94 (0.26)
$A_1 = -1, A_2 = -1$	11.86 (0.15)	7.54 (0.24)

We used "MLSurv" to represent the estimated regimens given by the proposed method and "Observed" to represent the regimens concordant with patients' actual treatments. First-line treatment of chemoimmunotherapy and immunotherapy were denoted as $A_1 = -1$ and $A_1 = 1$, respectively. Second-line treatment of chemotherapy and a combination of chemotherapy and monoclonal antibody were denoted as $A_2 = -1$ and $A_2 = 1$, respectively.

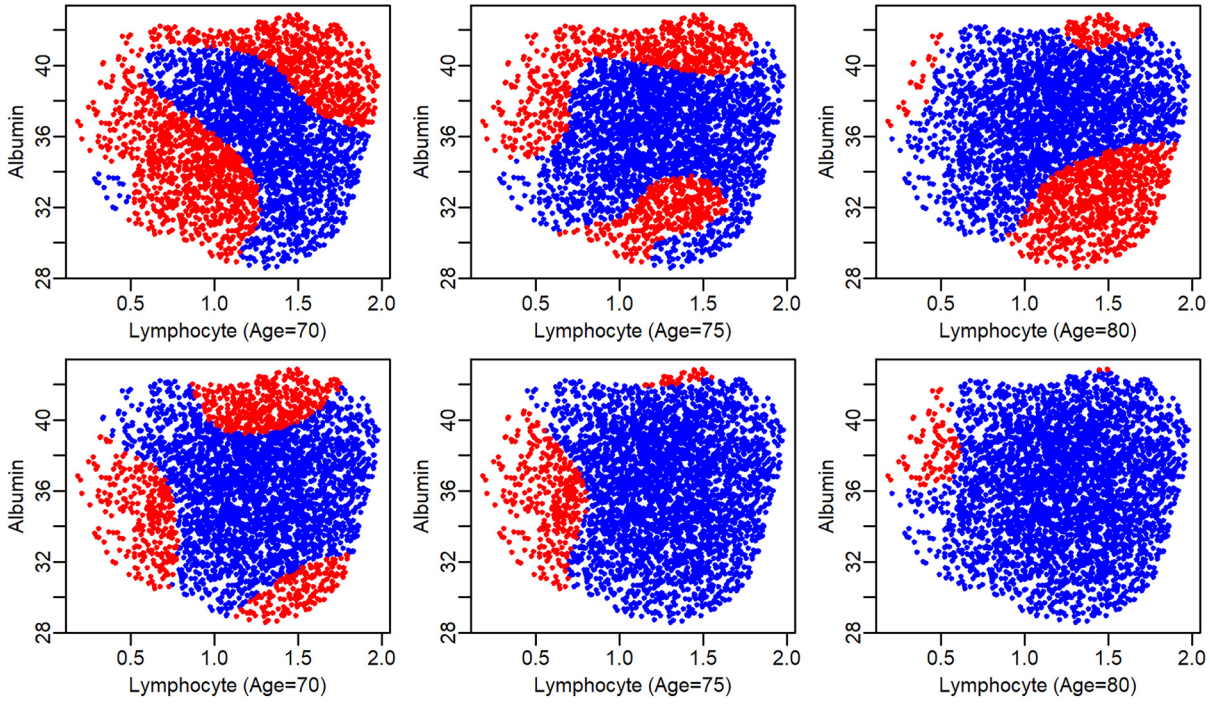

Figure 4. Decision boundaries of first-line treatment for males with ECOG PS of 0 or 1 are displayed in the top panel. Decision boundaries of first-line treatment for males with ECOG PS of 2 or 3 are displayed in the bottom panel. From left to right, patients age are 70, 75, and 80. Blue and red represent first-line Immunotherapy and Chemoimmunotherapy, respectively. The *x*-axis is the lymphocyte, and the *y*-axis is the albumin.

The uniform regimes refer to the four potential combinations of first-line and second-line treatments. The two first-line options consist of chemoimmunotherapy and immunotherapy, while the two second-line treatments include chemotherapy and a combination of chemotherapy with a monoclonal antibody. For a detailed breakdown, please see Table 2. To ensure consistency, the estimated restricted OS had all patients received uniform regimes was obtained using matching as well. We were interested in a treatment decision model capable of accommodating non-linearity and interactions. Therefore, we used an SVM with the RBF kernel in the MLSurv algorithm for estimating the DTRs.

To assess the generalizability of the proposed approach (MLSurv), we utilized a five-fold cross-validation strategy. The study sample was split into five folds, with each serving as a test set once while the DTRs were estimated using the data from the other four folds combined. Then, we obtained the estimated restricted OS for the whole study sample. We repeated this process 200 times, and the results are summarized in Table 2. The estimated restricted OS ($\hat{T} = 12.47$) and the estimated restricted survival time from second-line initiation ($\hat{T}_2 = 8.10$) had all patients followed the treatment regimens given by the MLSurv method were better than those estimates given by other regimens. Considering the short survival time for

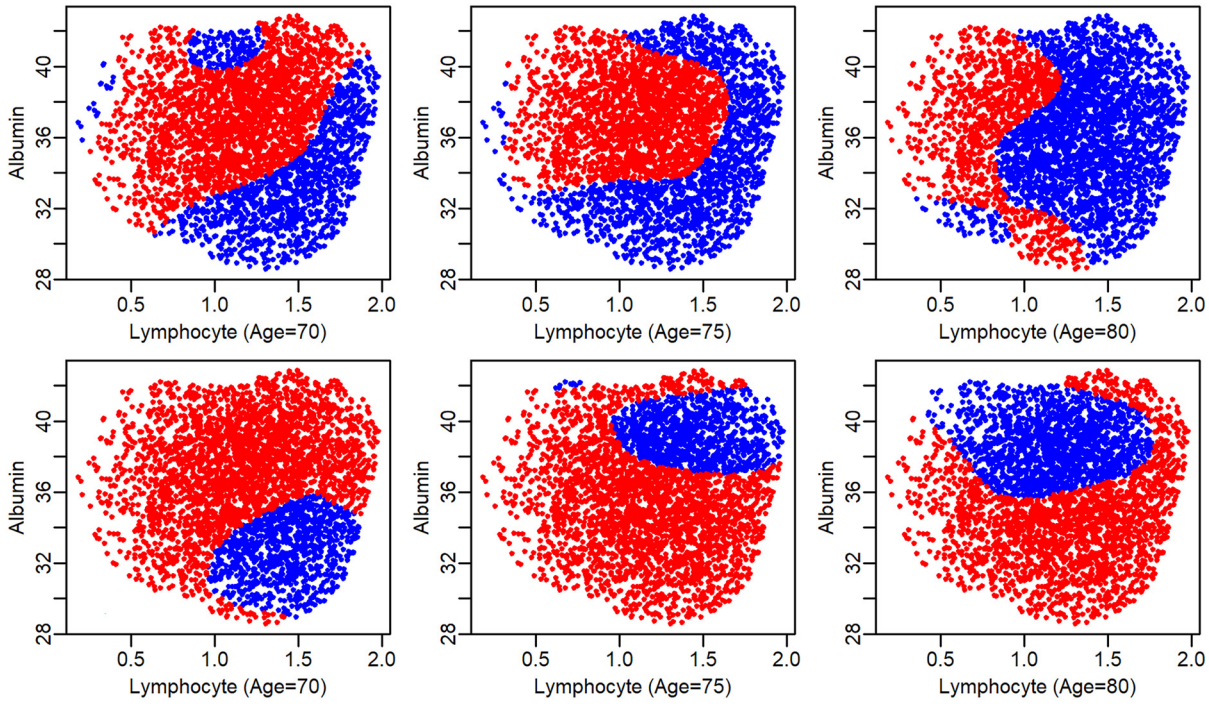

Figure 5. Decision boundaries of first-line treatment for females with ECOG PS of 0 or 1 are displayed in the top panel. Decision boundaries of first-line treatment for females with ECOG PS of 2 or 3 are displayed in the bottom panel. From left to right, patients age are 70, 75 and 80. Blue and red represent first-line Immunotherapy and Chemoimmunotherapy, respectively. The x-axis is the lymphocyte, and the y-axis is the albumin.

these patients, even a modest extension of one or two weeks in their lives may be considered a meaningful improvement. In this application, we chose to use an SVM with minimal tuning and context-specific adjustment. Perhaps exploring a broader range of tuning parameter options could enhance the estimated DTRs and OS time from MLSurv. Nevertheless, we demonstrated the overall promise and value of matching-based machine learning approaches to DTR estimation with survival endpoints.

To facilitate the visualization of the first-line treatment decision rule, we selected patients with particular baseline characteristics to illustrate the decision boundaries. According to Figures 4 and 5, it is obvious that older patients, males, and patients with severe clinical conditions tend to be assigned the less aggressive treatment, which is immunotherapy. However, it is not quite clear how the decision boundaries are associated with the baseline albumin and lymphocyte. Hence, a more involved medical discussion on the implementation strategy for such a DTR, as well as additional trial and observational studies, are needed.

5 Discussion

We have extended a matching-based direct optimization algorithm to estimate multi-stage optimal DTRs for complex diseases, leveraging massive EHR data when the endpoint is time-to-event, e.g. OS, subject to right-censoring. Compared to existing methods, our approach demonstrates substantial improvements in robustly identifying optimal DTRs that are nonlinear and complex. Additionally, it yields satisfactory results in optimizing clinical outcomes when the DTRs are linear. At each stage of the treatment decision, the proposed algorithm first estimates a counterfactual survival time for patients with censored outcomes through matching. Subsequently, it solves a weighted classification problem with weights constructed by the contrast between the counterfactual survival times. Under the specified assumptions, consistency of the proposed method can be achieved, when adequate confounders are considered in the matching steps, as well as an appropriate distance caliper is selected.[14,16] The proposed MLSurv approach does not require explicit modeling of the treatment or censoring mechanisms, which avoids the risk of instability caused by extreme weights. Additionally, it does not require explicit modeling of the outcome, which alleviates the model misspecification concerns. Backward induction was employed in the MLSurv method to address the complex confounding problem. We emphasize that the application in

Section 4 was aimed at demonstrating how MLSurv could generate evidence from EHR data to assist clinical decision-making. Nevertheless, in practice, the clinician's judgment should always take precedence.

Research using EHR data is highly likely to encounter challenges associated with high dimensionality, which may affect the performance of MLSurv. Notably, there have been attempts in matching methods literature to deal with high-dimensional data.[15,32,31] One approach involves projecting a large number of confounders to a lower dimensional space, such as computing prognostic risks[33] and propensity scores, and subsequently matching patients on those overall measures. Practically, we recommend incorporating subject-knowledge-based confounders, which can be directly included in matching steps or used to construct summary features for matching comprehensive measures. This may help us to improve the balance of covariates among arms. By following the recommendations, we may improve the efficiency and attain robustness of MLSurv in high-dimensional data. While in this article, we simplified the problem into two stages of treatment decisions between two options, it's noteworthy that MLSurv can be generalized to more than two stages, and importantly, accommodate more than two treatment choices per stage by adopting multicategory machine learning methods.[34] Dealing with missing data is another significant concern when working with EHR data. We believe that possible future research of extending our method involves leveraging matching techniques to handle missing data. We employed a weighted SVM with the hinge loss function to accommodate a soft margin in the MLSurv algorithm, which requires extensive efforts in tuning parameters, especially when utilizing complex kernels. To enhance efficiency, alternative machine learning methods, such as a weighted random forest, could be considered. It's worth noting that, in this article, we followed the common practice in SVM literature to decide the choice of kernels. Alternatively, one may explore rigorous kernel selection methods to determine the optimal choice in a data-adaptive manner.[35]

In summary, MLSurv exhibits robust performance by requiring neither prior knowledge regarding the shape of the decision boundaries, nor the modeling assumptions on treatment propensity and censoring. Notably, it effectively handles cases where extreme weights are likely to occur when applying IPW-based approaches.

Acknowledgments

Research reported in this publication utilized the Cancer Biostatistics Shared Resource at Huntsman Cancer Institute at the University of Utah and was supported by the National Cancer Institute of the National Institutes of Health under Award Number P30CA042014. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article.

ORCID iD


Xuechen Wang https://orcid.org/0000-0002-3584-7821

References

1. Chakraborty B and Murphy SA. Dynamic treatment regimes. *Annu Rev Stat Appl* 2014; **1**: 447.
2. Collins LM, Murphy SA and Strecher V. The multiphase optimization strategy (most) and the sequential multiple assignment randomized trial (SMART): New methods for more potent ehealth interventions. *Am J Prev Med* 2007; **32**: S112–S118.
3. Lei H, Nahum-Shani I, Lynch K, et al. A "SMART" design for building individualized treatment sequences. *Annu Rev Clin Psychol* 2012; **8**: 21–48.
4. Murphy SA. Optimal dynamic treatment regimes. *J R Stat Soc: Ser B (Statistical Methodology)* 2003; **65**: 331–355.
5. Robins JM. Optimal structural nested models for optimal sequential decisions. *In Proceedings of the second seattle Symposium in Biostatistics* New York: Springer, 2004. pp.189–326.
6. Watkins CJ and Dayan P. Q-learning. *Mach Learn* 1992; **8**: 279–292.
7. Zhang B, Tsiatis AA, Laber EB, et al. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika* 2013; **100**: 681–694.
8. Zhao YQ, Zeng D, Laber EB, et al. New statistical learning methods for estimating optimal dynamic treatment regimes. *J Am Stat Assoc* 2015; **110**: 583–598.
9. Austin PC and Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015; **34**: 3661–3679.
10. Bang H and Tsiatis AA. Estimating medical costs with censored data. *Biometrika* 2000; **87**: 329–343.
11. Robins JM and Finkelstein DM. Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; **56**: 779–788.

12. Robins JM and Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In: *AIDS epidemiology*. Springer, 1992, pp. 297–331.
13. Lee BK, Lessler J and Stuart EA. Improving propensity score weighting using machine learning. *Stat Med* 2010; **29**: 337–346.
14. Austin PC and Stuart EA. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Stat Methods Med Res* 2017; **26**: 1654–1670.
15. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci: Rev J Inst Math Stat* 2010; **25**: 1.
16. Wu P, Zeng D and Wang Y. Matched learning for optimizing individualized treatment strategies using electronic health records. *J Am Stat Assoc* 2019; **115**: 380–392.
17. Bellman R. Dynamic programming. *Science* 1966; **153**: 34–37.
18. Goldberg Y and Kosorok MR. Q-learning with censored data. *Ann Stat* 2012; **40**: 529.
19. Liu Y, Wang Y, Kosorok MR, et al. Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens. *Stat Med* 2018; **37**: 3776–3788.
20. Simoneau G, Moodie EE, Nijjar JS et al. Estimating optimal dynamic treatment regimes with survival outcomes. *J Am Stat Assoc* 2020; **115**: 1531–1539.
21. Rubin DB. Randomization analysis of experimental data: The fisher randomization test comment. *J Am Stat Assoc* 1980; **75**: 591–593.
22. Robins JM. Robust estimation in sequentially ignorable missing data and causal inference models. In: *Proceedings of the American Statistical Association*, volume 1999. Indianapolis, IN, 2000, pp. 6–10.
23. Karrison TG. Use of irwin's restricted mean as an index for comparing survival in different treatment groups–interpretation and power considerations. *Control Clin Trials* 1997; **18**: 151–167.
24. Zhao YQ, Zeng D, Laber EB, et al. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika* 2015; **102**: 151–168.
25. Sutton RS and Barto AG. *Reinforcement Learning: An Introduction*. Cambridge: MIT press, 2018.
26. Zhao Y, Zeng D, Rush AJ, et al. Estimating individualized treatment rules using outcome weighted learning. *J Am Stat Assoc* 2012; **107**: 1106–1118.
27. Ettinger DS, Wood DE, Aisner DL et al. Non–small cell lung cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2022; **20**: 497–530.
28. Birnbaum B, Nussbaum N, Seidl-Rathkopf K, et al. Model-assisted cohort selection with bias analysis for generating large-scale cohorts from the EHR for oncology research. *arXiv preprint arXiv:2001.09765*, 2020.
29. Ma X, Long L, Moon S, et al. Comparison of population characteristics in real-world clinical oncology databases in the US: Flatiron health, seer, and npcr. *Medrxiv*, 2020.
30. Zhang Q, Gossai A, Monroe S, et al. Validation analysis of a composite real-world mortality endpoint for patients with cancer in the united states. *Health Serv Res* 2021; **56**: 1281–1287.
31. Dehejia RH and Wahba S. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *J Am Stat Assoc* 1999; **94**: 1053–1062.
32. Smith HL. 6. matching with multiple controls to estimate treatment effects in observational studies. *Sociol Methodol* 1997; **27**: 325–353.
33. Wang X, Kerrigan K, Puri S, et al. Dynamic prediction of near-term overall survival in patients with advanced NSCLC based on real-world data. *Cancers* 2022; **14**: 690.
34. Huang GB, Zhou H, Ding X and Zhang R. Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst, Man, Cybernet, Part B (Cybernetics)* 2011; **42**(2): 513–529.
35. Burges CJ. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov* 1998; **2**(2): 121–167.