Journal of Data Science 24 (4), 703–720
October 2026
DOI: 10.6339/25-JDS1193
Statistical Data Science

Transfer Learning for Individualized Treatment Rules with Application to Sepsis Patients Data

Andong Wang[1,*], Kelly Wentzlof[2,†], Johnny Rajala[3,†], Miontranese Green[4], Dr.Yunshu Zhang[1], and Dr.Shu Yang[1]

[1]*Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA*
[2]*Department of Statistics, Indiana University, Bloomington, IN 47045, USA*
[3]*Department of Mathematics, University of Maryland, College Park, MD 20742, USA*
[4]*Department of Statistics, California State University, Long Beach, CA 90840, USA*

Abstract

Modern precision medicine aims to utilize real-world data to provide the best treatment for an individual patient. An individualized treatment rule (ITR) maps each patient's characteristics to a recommended treatment scheme that maximizes the expected outcome of the patient. A challenge precision medicine faces is population heterogeneity, as studies on treatment effects are often conducted on source populations that differ from the populations of interest in terms of the distribution of patient characteristics. Our research goal is to explore a transfer learning algorithm that aims to address the population heterogeneity problem and obtain targeted, optimal, and interpretable ITRs. The algorithm incorporates a calibrated augmented inverse probability weighting estimator for the average treatment effect and employs value function maximization for the target population using Genetic Algorithm to produce our desired ITR. To demonstrate its practical utility, we apply this transfer learning algorithm to two large medical databases, eICU Collaborative Research Database and Medical Information Mart for Intensive Care III. We first identify the important covariates, treatment options, and outcomes of interest based on the two databases, and then estimate the optimal linear ITRs for patients with sepsis. Our research introduces and applies new techniques for data fusion to obtain data-driven ITRs that cater to patients' individual medical needs in a population of interest. By emphasizing generalizability and personalized decision-making, this methodology extends its potential application beyond medicine to fields such as marketing, technology, social sciences, and education.

Keywords *augmented Inverse Probability Weighting; causal inference; generalizability; genetic algorithm; optimization; population heterogeneity; precision medicine*

1 Introduction

Under FDA's 21st-Century Cures Act, the field of precision medicine is developed to utilize real-world data to provide evidence-based and data-driven optimal treatment for each individual patient (Kosorok and Laber, 2019). A rich literature has been developed on estimating individualized treatment rules (ITRs) for precision medicine. For instance, Outcome-Weighted Learning (OWL) (Zhao et al., 2012) formulates the estimation of optimal treatment rules as a weighted

*Corresponding author. Email: andongwang7@gmail.com or andong@unc.edu.
†Equal Contribution.

© 2026 The Author(s). Published by the School of Statistics and the Center for Applied Statistics, Renmin University of China. Open access article under the CC BY license.
Received November 24, 2024; Accepted June 23, 2025

classification problem that directly targets the value function. It has since been extended to incorporate regularization, kernel methods (Zhou et al., 2017), and doubly robust estimation strategies (Zhang et al., 2012), offering both flexibility and robustness in observational settings.

An ongoing challenge facing modern precision medicine is that studies on outcomes of potential treatments are conducted for a source population that is different from some target population for which the treatment will be implemented in terms of the distribution of patients' demographics, characteristics, and health factors. While treatment may be proven effective after testing on a sample population, it may not have the same optimal effect for all patients within other populations. This problem is often referred to as population heterogeneity, or covariate shift and it is widely studied in social sciences and biological sciences (Nagin and Paternoster, 2000; Ryall et al., 2012). In medicine, a typical example of population heterogeneity consists of an experimental clinical trial population and a real-world population. Despite experimental data having strong internal validity, it has limited external validity and introduces bias when applied to other populations due to its limited sample pool from inclusion and exclusion criteria (Rothwell, 2005). As opposed to experimental data, using real-world data and observational data provides a sample that is representative of the larger population. Thus, the goal of this research is to use knowledge learned about one population to inform decisions for another, a technique referred to as transfer learning.

To achieve our research goals, we focus on the transfer learning framework. Based on the potential outcome framework in causal inference (Imbens and Rubin, 2015), the transfer learning framework combines two key components: the Augmented Inverse Probability Weighting (AIPW) and calibration weighting (CW). The AIPW estimator estimates the average treatment effect (ATE), which is the average expected outcome for a patient population under one specific treatment rule. The AIPW estimator has the desirable property of double robustness: it provides a consistent estimation even if either the outcome regression model or the propensity score model is misspecified (Bang and Robins, 2005; Glynn and Quinn, 2010). CW addresses covariate shift by calibrating the covariates of patients in the source population to assimilate the target population's patient covariates using entropy balancing (Hainmueller, 2012). Entropy balancing is a particularly effective calibration method, offering exact covariate balance through convex optimization. It has seen growing use in policy evaluation (Zubizarreta, 2015) and causal generalizability studies (Chu et al., 2023), and it enables valid inference under population heterogeneity without explicitly modeling the assignment mechanism.

To demonstrate the utility of our framework, we apply it to the problem of learning optimal linear treatment rules for sepsis patients using two large-scale medical databases: Medical Information Mart for Intensive Care (MIMIC-III) and electronic Intensive Care Unit Collaborative Research Database (eICU-CRD). We define the treatment, outcome, and covariates of interest, and construct optimal and interpretable linear ITRs tailored to the target population. In addition to the real-world application, we conduct a simulation study to validate the performance of our method under controlled conditions, where treatment effects and population distributions are known. This allows us to demonstrate the impact of calibration weighting and value function optimization in recovering target-optimal treatment rules. Through both the medical application and simulation study, we illustrate the flexibility, effectiveness, and generalizability of our proposed transfer learning framework for estimating individualized treatment rules under population heterogeneity. We hope to inspire other researchers to apply our framework to populations of patients with other diseases or adapt it to other fields of study.

The rest of the paper is organized as follows: In Section 2, we define notations and outline the theoretical assumptions as the premise of our research. We then compare estimators for the

ATE. In Section 3, we delve into the three main components of the methodology: CW, Value Function, and Genetic Algorithm. In Section 4, we apply the framework to real-world data of two populations of patients with sepsis and compile the application results. In Section 5, we validate the utility of our framework by conducting a simple simulation study. In Section 6, we discuss the advantages and limitations of our framework. All relevant R code for the medical application and the simulation study is provided in the Supplementary Material.

2 Statistical Framework

To study the average effect of a treatment rule and eventually optimize it, we adopt some basic concepts from the potential outcomes framework for causal inference in (Imbens and Rubin, 2015) as building blocks for our research. The ATE measures the difference between the average outcome that would be achieved if all individuals in the population were to receive treatment and if all were to receive control, given that the treatment is binary. We start with an overview of the notation that will be used throughout this paper. Then, we will move to the assumptions that must hold for us to proceed to develop our methodology.

2.1 Notation

Suppose we have a population of n patients and each patient is indexed by i. The baseline covariates of these patients are specified in n by p matrix X, where p is the number of covariates of interest. Thus, X_i is the vector of covariates for patient i. We refer to covariates as pre-treatment attributes or features of each patient such as age, gender, medical history, etc. Covariates help explain variation in outcomes, making estimates more precise and allowing researchers to identify subgroups in patients with different responses to treatment and control. The treatment assignment vector $A \subset \{0, 1\}^n$ denotes what intervention each patient gets. We will consider a binary treatment, with $A_i = 1$ if patient i receives treatment, and $A_i = 0$ if they receive control. Y denotes patient outcome, i.e., the greater the Y_i, the better patient i reacts to the treatment assigned to them, A_i.

Following the potential outcome framework proposed in Imbens and Rubin (2015), the individual treatment effect is defined as the difference in the outcome if the patient is given treatment, $Y^*(1)$, over the outcome if the same patient is given control, $Y^*(0)$. For each patient, one of these two outcomes can be observed, referred to as the "observed outcome"; the other outcome will be missing, referred to as the "missing outcome" or the "counterfactual outcome". For patient i, the potential outcome under treatment a is represented as $Y^*(A_i = a)$. Hence, we can represent the ATE by taking the expectation of differences between potential outcomes: $\mathbb{E}[Y^*(1) - Y^*(0)]$.

When estimating the ATE, τ, we consider the propensity score. The propensity score is the probability unit i receives active treatment given its covariates, $\pi(X_i) = \Pr(A_i = 1 \mid X_i)$. In randomized clinical trials, researchers control the probability of each patient receiving treatment, which is typically fixed across patients, regardless of patient characteristics; in an observational study, a patient's propensity score may vary depending on their characteristics, introducing assignment bias. In Section 2.3, we discuss a variety of estimators for the ATE, including the naive estimator, the inverse probability weighting (IPW) estimator, the outcome regression (OR) estimator, and the AIPW estimator. Then we explore how these estimators address the assignment bias.

Table 1: Structure of observed and counterfactual data across source and target populations. The source population includes treatment assignment and observed outcomes; counterfactuals are defined but partially unobserved. The target population includes only covariates.

Population	S_i	X_i	A_i	Y_i	$Y_i^*(1)$	$Y_i^*(0)$
Source (Treated)	1	✓	1	✓	✓	–
Source (Control)	1	✓	0	✓	–	✓
Target	0	✓	–	–	–	–

We denote the population to which each patient belongs using a binary indicator S, where $S_i = 0$ if patient i is from the source population and $S_i = 1$ if the patient is from the target population. For our transfer learning applications, the data structure for the source and target populations is illustrated in Table 1. The source population consists of patients for whom baseline covariates, treatment assignments, and outcomes have all been observed. In contrast, the target population consists of patients for whom only baseline covariates X are available, and the goal is to recommend treatment decisions that maximize expected outcomes. This asymmetry reflects real-world settings in which we must apply what is learned from fully labeled data (source) to guide treatment in a new population where outcomes are not yet realized. When population heterogeneity exists, the covariate distribution differs across the two groups, i.e., $\Pr(X = x \mid S = 1) \neq \Pr(X = x \mid S = 0)$. This covariate shift, also known as selection or sampling bias, is addressed in our framework using calibration weighting, described in Section 3.1.

We use d to denote an ITR. In this paper, we only consider the class of linear decision rules of the form $d_\eta(X) = \mathbb{I}\left\{\eta^\top X > 0\right\}$ where $\eta \in \mathbb{R}^p$ specifies the covariate coefficient vector that uniquely identifies the rule. To interpret the rule $d_\eta(X)$ for patient i, we substitute the patient's covariates vector X_i into $\mathbb{I}\left\{\eta^\top X_i > 0\right\}$ and evaluate the indicator function. If $d_\eta(X_i) = 1$, then the treatment yields good outcome and is recommended for patient i; if $d_\eta(X_i) = 0$, then the treatment does not yield a good outcome and is not recommended for patient i.

2.2 Assumptions

Assumption 1. *Stable Unit Treatment Value:* $Y = Y^*(1)(A) + Y^*(0)(1 - A)$.

Meeting the Stable Unit Treatment Value Assumption (SUTVA) allows researchers to use multiple units within one study. SUTVA incorporates two main components. The first component of SUTVA is that the treatment of one unit does not affect the outcome of another unit. This is typically ensured when researchers separate the participants of a study, therefore, reducing the likelihood of the participants' effects intermingling. If participants interact, resulting in outcomes that are different if the participants had not come in contact with each other, then SUTVA is violated. The second component of SUTVA is that the researchers must minimize any differences in the efficacy and the method of administering the treatment. For example, in a drug trial, researchers must ensure that patients in the treatment group all receive treatment of the same strength. If the efficacy of the treatment varies within this singular treatment group, then SUTVA is violated.

Assumption 2. *No Unmeasured Confounding:* $(Y^*(1), Y^*(0)) \perp\!\!\!\perp A \mid X$.

The No Unmeasured Confounding (Conditional Exchangeability) assumption entails that all the variables that influence both the treatment assignment and the outcome of interest

are measured and accessible in our data. If there are unmeasured confounding variables, our estimation of the causal effect will be biased. This assumption is one of the most commonly violated assumptions in causal inference.

Assumption 3. *Positivity of Treatment Assignment:* $0 < \Pr(A = a \mid X = x) < 1$ *for all* x *and* $a = 0, 1$.

The Positivity assumption requires every patient to have a positive probability of being assigned to treatment or control. If this assumption is violated, all patients could be assigned to one treatment group, rendering inference impossible. This is essential as we model the counterfactual and estimate the treatment effect.

Assumption 4. *Transportability:* $\mathbb{E}[Y^*(A = a) \mid S = 1, X = x] = \mathbb{E}[Y^*(A = a) \mid X = x]$ *for all* x *and* $a = 0, 1$.

Transportability describes the ability to "transport" causal effect estimated from a random clinical trial or observation study done on the source population to a target population. It requires the ATE to be consistent across populations. In recent literature, transportability is sometimes referred to as generalizability, external validity, or recoverability. These terms have slightly different definitions concerning the overlap between populations and there have been discussions of the differences as seen in Colnet et al. (2024).

Assumption 5. *Common Support:* $\Pr(S = 1|X) > 0$.

The common support assumption entails that for the inference to be transportable, the support for the source population covariate distribution is required to overlap the support for the target population covariate distribution.

2.3 Estimators

Now we will consider various estimators for the ATE, τ, and how each estimator handles assignment bias that is introduced when the propensity score is not predetermined. We include the naive, IPW, OR, and AIPW estimators not to suggest they are direct components of our proposed method, but to offer a conceptual progression that leads to the development of our calibrated AIPW estimator. In particular, the AIPW estimator is a cornerstone of our proposed method, and the others help illustrate the value of double robustness in addressing treatment assignment bias. This context helps situate our method as a natural extension of these classical estimators.

The naive estimator, $\hat{\tau}_0$, takes the difference between the average observed outcome of patients in the treatment group and the average observed outcome of patients in the control group to obtain the ATE. Let n_1 be the number of patients in the treatment group and n_0 be the number of patients in the control group, then the naive estimator is represented as:

$$\hat{\tau}_0 = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i - \frac{1}{n_0} \sum_{j=1}^{n_0} Y_j. \tag{1}$$

The naive estimator provides a relatively accurate estimation of the ATE for randomized experiments because there is no assignment bias that needs to be addressed. However, with observational data, the naive estimator is biased and performs poorly because there is no way to

consider the covariates that the treatment depends on (i.e., there is no way to minimize the assignment bias).

The IPW estimator, $\hat{\tau}_{\text{IPW}}$, estimates ATE with the objective to address the assignment bias and is formulated as:

$$\hat{\tau}_{\text{IPW}} = \frac{1}{n}\sum_{i=1}^{n}\frac{A_i Y_i}{\hat{\pi}(X_i)}. \tag{2}$$

The IPW estimator weights each patient's treatment effect based on their covariates. The weights, $1/\hat{\pi}(X_i)$, adjusts for the probability of patient i receiving treatment. The more likely a patient is to receive treatment, the higher the propensity score, the smaller the weight. The IPW estimator is unbiased if we correctly specify the propensity score model.

The OR estimator, $\hat{\tau}_{\text{OR}}$, estimates ATE by modeling the outcome based on observed values for covariates, as shown in Equation (3):

$$\hat{\tau}_{\text{OR}} = \frac{1}{n}\sum_{i=1}^{n}\hat{m}(X_i). \tag{3}$$

We denote m as the true mapping from covariates to outcome. However, in reality, it is more likely that m is unknown and needs to be estimated. Thus, we use $\hat{m}(X_i)$ to represent the estimated outcome for patient i based on the OR model and their covariates. The fitted OR model can be parametric (e.g., linear regression, logistic regression) or nonparametric (e.g., machine learning methods). If the model m is correctly specified, then we can better estimate the missing outcomes and the treatment effect. Although the OR estimator is efficient, it tends to be nonrobust due to the misspecification of the model.

The AIPW estimator, $\hat{\tau}_{\text{AIPW}}$, estimates the ATE by augmenting the IPW estimator with OR, as shown in Equation (4), achieving the doubly robust property:

$$\hat{\tau}_{\text{AIPW}} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{A_i Y_i}{\pi(X_i)} - \frac{A_i - \pi(X_i)}{\pi(X_i)}m(X_i)\right] = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{A_i[Y_i - m(X_i)]}{\pi(X_i)} + m(X_i)\right\}. \tag{4}$$

The double robustness of the AIPW estimator proposed in Glynn and Quinn (2010) suggests that if either the propensity score model or the outcome regression model is misspecified, the AIPW estimator will remain unbiased. In Equation (4) above, two mathematical representations of the AIPW estimator are shown to demonstrate the doubly robust property. In the expression on the left side, if the propensity score model is correctly specified, then $\pi(X_i)$ closely approximates A_i, which will cancel out the outcome regression component, leaving only the correctly specified IPW model. Similarly, in the expression on the right side, if the outcome regression model is correctly specified, then $m(X_i)$ closely approximates Y_i, which will cancel out the IPW estimator component, leaving only the correctly specified OR model. In our application, we employ the AIPW estimator to estimate the ATE for its double robustness.

3 Proposed Method

3.1 Calibration Weighting

When covariate heterogeneity exists between patients from the source population and the target population, the optimal ITR we find for the source population will likely not be optimal for the target population. In other words, the optimal ITR learned from clinical trials or observation

studies possesses internal validity and lacks external validity, which means this ITR will not lead to the best outcome for patients awaiting treatment outside of that clinical trial or observation study. This issue can be referred to as a selection bias or sampling bias since the sampling distribution of the source population differs from that of the target population.

To correct this sampling bias, we use entropy balancing methods introduced in Hainmueller (2012) to compute calibration weights based on the covariates information from patients of both the source and target populations. Entropy balancing weighting is a trusted method for balancing covariates and has been studied and applied recently in Chu et al. (2023) and Wu and Yang (2023). It reweighs units in the source population so that the weighted empirical distribution of their covariates matches that of the target population. Specifically, the weights are chosen to satisfy moment-matching constraints on covariates between the two populations. This is achieved through solving a constrained optimization problem that minimizes the relative entropy (i.e., Kullback-Leibler divergence) of the weights from uniformity, subject to the balancing constraints. Since computation of calibration weights only requires covariates and does not involve treatment or outcome data, causal identification assumptions are preserved. Once computed, these weights are used in downstream estimation tasks to ensure that inferences drawn from the source population generalize appropriately to the target population.

3.2 Value Function

The value function estimates the total treatment effect for the whole population given an ITR, d_η, thus it is used to evaluate the quality of an ITR. It is expressed as follows:

$$\hat{V}(d_\eta; \hat{w}, I_n) = \sum_{i \in I_n} \hat{w}_i \left\{ \left[\frac{A_i d_\eta(X_i)}{\hat{\pi}(X_i)} + \frac{(1 - A_i)\left[1 - d_\eta(X_i)\right]}{1 - \hat{\pi}(X_i)} \right] \left[Y_i - \hat{m}(X_i) \right] + \hat{m}(X_i) \right\}, \quad (5)$$

where I_n is the index set of all patients in the source population.

Equation (5) defines our proposed calibrated augmented inverse probability weighting (CAIPW) value function, which extends the classical AIPW estimator in Equation (4) by incorporating calibration weights $\hat{w}_i$ and by evaluating a specific proposed treatment rule d_η. While the AIPW estimator is doubly robust and addresses treatment assignment bias under the assumption of no unmeasured confounding, it does not account for distributional differences between a source population and a target population. In other words, it assumes equal weights for all individuals in the source population.

Additionally, a key component of the value function in Equation (5) evaluates as follows:

$$\frac{A_i d_\eta(X_i)}{\hat{\pi}(X_i)} + \frac{(1 - A_i)[1 - d_\eta(X_i)]}{1 - \hat{\pi}(X_i)} = \begin{cases} 0, & \text{if } A_i \neq d_\eta(X_i) \\ \frac{1}{\hat{\pi}(X_i)}, & \text{if } A_i = d_\eta(X_i) = 1 \\ \frac{1}{1 - \hat{\pi}(X_i)}, & \text{if } A_i = d_\eta(X_i) = 0 \end{cases}.$$

This highlights another key difference between the AIPW estimator and the CAIPW value function: the AIPW estimator estimates the ATE simply under the observed treatment assignment A, while the CAIPW value function estimates the expected outcome under a proposed treatment rule d_η, which may differ from A for many individuals.

Therefore, our proposed CAIPW value function simplifies to the AIPW estimator when the following two conditions are met: (i) all subjects have equal calibration weights, and (ii) the observed treatment assignment matches the proposed treatment rule. The AIPW estimator

can be considered as a special case of the CAIPW value function. This generalization of the AIPW estimator that makes up our proposed CAIPW value function is essential for learning individualized treatment rules (ITRs) that are tailored to the target population when only the source population has observed treatment and outcome data. Without such a generalization, an ITR optimized for the source may yield suboptimal or biased results when applied to the target population.

Note that, similar to the AIPW estimator, the CAIPW estimator consists of an estimate of the propensity score $\hat{\pi}(X_i)$ and an outcome estimate from outcome regression $\hat{m}(X_i)$, which can be obtained through logistic regression and linear regression respectively. Additionally, CAIPW weights are normalized, so the summation of the weighted treatment effect in the value function yields the ATE.

The variance of the estimated value function $\hat{V}(d)$ is formulated based on the influence function associated with the augmented inverse probability weighted (AIPW) estimator. For a given treatment rule $d(X)$, we estimate the variance of the value function as follows:

$$\widehat{\mathrm{Var}}(\hat{V}(d)) = \frac{1}{n^2}\sum_{i=1}^{n}\left\{\hat{w}_i\left[\left(\frac{A_i d(X_i)}{\hat{\pi}(X_i)} + \frac{(1-A_i)(1-d(X_i))}{1-\hat{\pi}(X_i)}\right)(Y_i - \hat{m}(X_i)) + \hat{m}(X_i)\right] - \hat{V}(d)\right\}^2, \tag{6}$$

where w_i is the calibration weight, $\hat{\pi}(X_i)$ is the estimated propensity score, and $\hat{m}(X_i)$ is the estimated outcome regression. This variance estimator accounts for the sampling variability of the value estimate and is used to construct Wald-style confidence intervals for policy evaluation.

3.3 Genetic Algorithm

The value function has a greater value for an ITR that performs better in optimizing patient outcomes. Our goal is to find the ITR that maximizes the value function. First, we find the optimal covariate coefficient vector:

$$\eta_{\mathrm{opt}} = \underset{\eta}{\mathrm{argmax}}\ \hat{V}(d_\eta; \hat{w}_i, I_n), \tag{7}$$

then we can substitute η^{opt} in our linear ITR formulation to get the optimal linear ITR:

$$d_{\mathrm{opt}} = \mathbb{I}\left\{\eta_{\mathrm{opt}}^{\top} X > 0\right\}. \tag{8}$$

In order to obtain the optimal covariate coefficient vector, η_{opt} in Equation (7), we use Genetic Algorithm (GA) to solve the value function maximization problem, specifically, using the `rgenoud` R package developed by Mebane Jr. and Sekhon (2011).

GA is a search-based optimization tool inspired by the mechanism of biological evolution and natural selection with a wide range of applications across disciplines, as discussed in Katoch et al. (2021). It is a population-based algorithm, which means it maintains a population of candidate solutions, allowing for both the search for a new solution space and the refinement of outstanding solutions in the current solution space. This makes the search process more robust and makes overcoming the local maxima easier. GA is flexible with respect to the type of objective function it can optimize. It is a suitable option for our non-differentiable, non-convex value function and can generate high-quality solutions.

4 Medical Application

Sepsis is a life-threatening complication of an infection, leading to 270,000 deaths each year in America. In every three hospital deaths, approximately one patient was suffering from sepsis at the time of death. Providing precise treatment for sepsis patients is crucial to preventing further medical complications such as organ failure and possible death, making sepsis a critical condition for studying individualized treatment strategies. In this section, we apply our proposed method to two large-scale public databases that contain sepsis patients' data to demonstrate its practical relevance in solving real-world problems.

4.1 Databases

MIMIC-III is a freely accessible single-center database consisting of desensitized health record data of over 40,000 patients who stayed in the intensive care units of the Beth Israel Deaconess Medical Center from 2001 to 2012. eICU-CRD is a freely accessible multi-center database for critical care research. It consists of desensitized health record data of over 200,000 patients who were treated in intensive care units all over the U.S. under the Phillips eICU program from 2014 to 2015.

A randomized clinical trial conducted in a single center tends to follow a simpler design compared to a multi-center trial, but it often has limited external validity. Since our methodology aims to expand external validity, in our application, we use the single-center data from MIMIC-III as our data for the source population and the multi-center data from eICU-CRD as our data for the target population.

4.2 Data Pre-Processing

After an Exploratory Data Analysis of the raw data, we removed duplicate observations of patients in both datasets to keep only the baseline entries of each patient before any treatment is applied. This leaves us with 20,955 unique patients' data in MIMIC-III and 21,995 unique patients' data in eICU-CRD.

We then reformatted and rescaled certain variables for consistency across the two datasets and removed all non-binary categorical variables due to their incompatibility with our methods. We also removed variables with over 60% missing entries and imputed the rest of the missing data with the MICE (Multivariate Imputation by Chained Equations) algorithm, introduced in Azur et al. (2011). The MICE algorithm fills in the incomplete variable based on observed variables multiple times in an iterative manner. Upon visually inspecting the covariate lists of both datasets, we kept the covariates that are shared by both datasets due to the requirement of calibration weighting. This leaves us with usable data from thirteen common variables in both datasets: re-admission (binary), age, weight at admission, temperature in Celsius, mean blood pressure, respiratory rate, sodium, glucose, blood urea nitrogen, creatinine, bilirubin, albumin, and white blood cell count. Two other variables, mechanical ventilation (binary) and death (binary) are only available in the MIMIC-III dataset but are kept as treatment and outcome variables in the source data. For more detailed reasoning for these decisions, see Section 4.3.

4.3 Treatment and Outcome

Out of the 15 variables available after data pre-processing, we chose mechanical ventilation as treatment and death as outcome.

The interplay between mechanical ventilation and sepsis has been discussed in recent medical research. Sepsis is responsible for approximately 70% of acute respiratory distress syndromes (ARDS). Zampieri and Mazza (2017) suggests that while a mechanical ventilator is necessary to support breathing for patients with sepsis-induced ARDS, suboptimal use of mechanical ventilators can cause lung injuries, which further contribute to a downward spiral of sepsis-related organ failures. The use of mechanical ventilation should be tailored to the patient's individual characteristics and conditions to improve. The emphasis on personalization makes mechanical ventilation a treatment of interest for our application.

The choice of patient death as the outcome variable in our application was an intuitive one. However, there are three caveats with this outcome variable:

1. Information on patient deaths is only available in MIMIC-III, our source population, which means we can only evaluate the effect of our optimal ITRs using estimated outcomes in our target population.

2. Death is coded in 3 scenarios in MIMIC-III: 1) Died in hospital; 2) Died within 48 Hours after discharge; 3) Mortality within 90 Days. We decided to integrate the 3 scenarios into a binary outcome variable that classifies a patient as dead (death = 1) if they fit at least one of the scenarios and not dead (death = 0) otherwise.

3. Since we are maximizing the value function to yield the best outcome for patients, we need to ensure that the outcome variable is greater in value when representing a "better" outcome. In our application, we use (1 - death) as the outcome variable, representing patient survival. Therefore, the outcome variable is 1 if the patient survives, 0 otherwise.

4.4 Covariates

The remaining 13 variables (hospital re-admission, mean blood pressure, body temperature, respiratory rate, sodium, glucose, blood urea nitrogen, creatinine, bilirubin, albumin, white blood cell count, weight, and age) are used as patient covariates, serving as the active components of our calibration weighting scheme.

In Figure 1, we compare the density curves for each of the 12 numeric variables of the 13 covariates between the target population based on the eICU-CRD database and the source population based on the MIMIC-III database. We notice that mean blood pressure, body temperature, respiratory rate, and albumin have clear distinctions between distributions for the two populations. The one categorical variable, hospital re-admission, has also shown a discrepancy between the two populations. In the source population, approximately 32.77% of patients had been re-admitted, whereas in the target population, only approximately 4.75% of patients had been re-admitted. The differences in distribution of these covariates indicate a population heterogeneity problem, and calibration weighting is needed to address such problem.

After computing and applying calibration weights to the source population, we examine their impact on the covariates that exhibited the greatest distributional discrepancies between the source and target populations—namely, mean blood pressure, body temperature, respiratory rate, and albumin (Figure 2). We observe that calibration weighting substantially improved alignment for respiratory rate and albumin, while the distributions of mean blood pressure and body temperature remained largely unchanged. This likely reflects the fact that entropy balancing is designed to match only the first few moments (e.g., mean, variance) of the covariate distributions. Moreover, each subject receives a single calibration weight to adjust the joint distribution of all covariates, rather than separate weights for each covariate. As a result, it is expected that not all covariates will achieve perfect alignment simultaneously. Nonetheless,

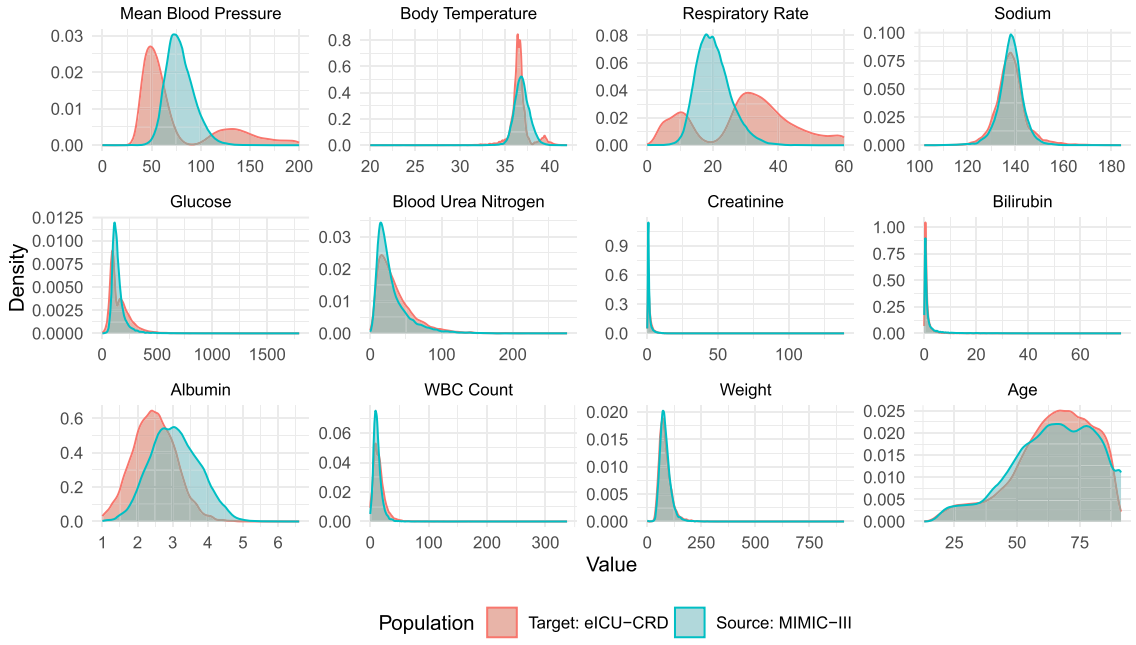


Figure 1: Density plots demonstrate covariate shifts between the two populations, i.e., population heterogeneity.

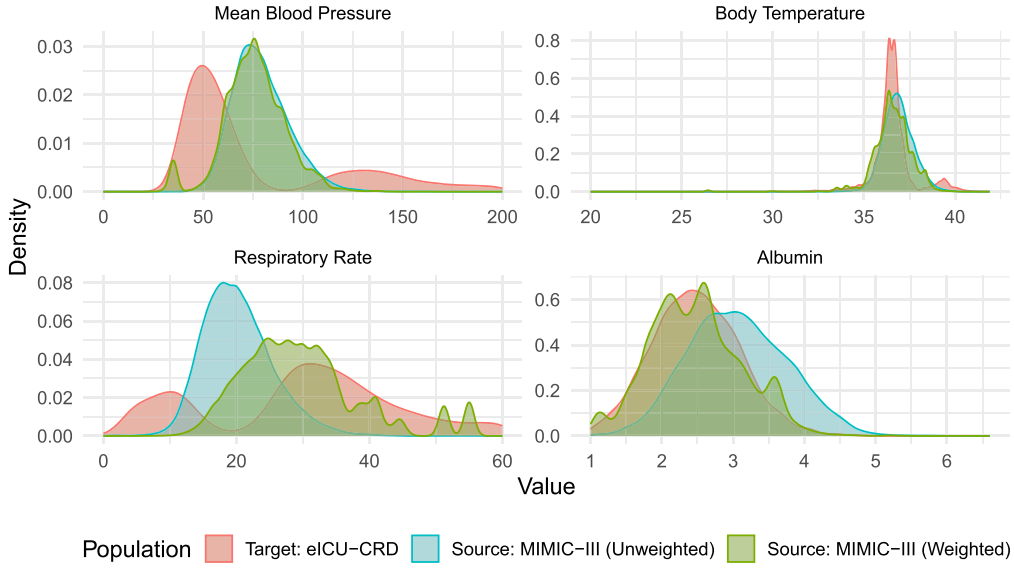


Figure 2: Density plots demonstrate the effect of calibration weighting on the 4 covariates with clear distribution shifts.

the overall improvement in distributional similarity demonstrates that calibration weighting is generally effective in adjusting the source population to more closely resemble the target population.

4.5 Results

Following our proposed methods, we obtained the entropy balancing calibration weights with covariates data from MIMIC-III and eICU-CRD. Then, we repeatedly ran GA that aims to maximize the value function: 1000 iterations with calibration weights and 1000 iterations with equal weights. Since we do not have data on treatment and outcome for the target population, we evaluate the optimal ITR produced by GA by computing its treatment effect with the value function.

With calibration weighting, the optimal ITR has a value of approximately 0.7616, while the unweighted optimal ITR has a value of approximately 0.7223. In the context of this application, our results imply that assigning mechanical ventilation treatments to patients using the optimal weighted ITR has led to an approximately 3.93% increase in survival rate compared to using the optimal ITR without calibration weighting. The resulting optimal weighted ITR is illustrated in Equation (9):

$$\begin{aligned} d_{\text{opt}} = \mathbb{I}\{&-0.3933 \cdot \text{Glucose} + 0.6507 \cdot \text{Blood Urea Nitrogen} + 0.6282 \cdot \text{Age} - 0.2484 \cdot \text{Weight} \\ &+ 0.4333 \cdot \text{Mean Blood Pressure} - 0.4738 \cdot \text{WBC Count} + 0.8800 \cdot \text{Respiratory Rate} \\ &+ 0.8830 \cdot \text{Bilirubin} - 0.6220 \cdot \text{Sodium} - 0.0565 \cdot \text{Creatinine} - 0.7644 \cdot \text{ReAdmission} \\ &+ 0.5545 \cdot \text{Body Temperature} + 0.3633 \cdot \text{Albumin} + 0.1634 > 0\}. \end{aligned} \tag{9}$$

After substituting a patient's baseline reading of the covariates, if $d_{\text{opt}} = 1$, then such a patient has a positive treatment effect and hence should be given the treatment of interest, and vice versa.

With the optimal weighted ITR, we can also make rudimentary analyses on which patient covariate is more important to the ITR by comparing the magnitudes of the coefficients. To account for the differences in the spreads of covariates' distributions, we adjusted the coefficients by multiplying the standard deviation of the corresponding covariate. Observing Figure 3, we found that Glucose is the primary covariate in our optimal ITR, and it has a negative coefficient, which indicates that the lower the patient's blood glucose reading, the more likely they will have a positive treatment effect and get assigned treatment. Blood Urea Nitrogen (BUN) is the secondary covariate in our optimal ITR, and it has a positive coefficient, which indicates that the higher the patient's BUN reading, the more likely they will have a positive treatment effect and get assigned treatment.

5 Simulation Study

To further demonstrate the performance of our methods under controlled conditions, we conducted a simulation study. While our medical application showcases the methods' utility in a real-world setting, the simulation study allows us to validate its functionality in a setting where covariate distributions and treatment effects are known and can be manipulated. Using generated data for covariates, treatment assignment, and outcome for both populations, we examine

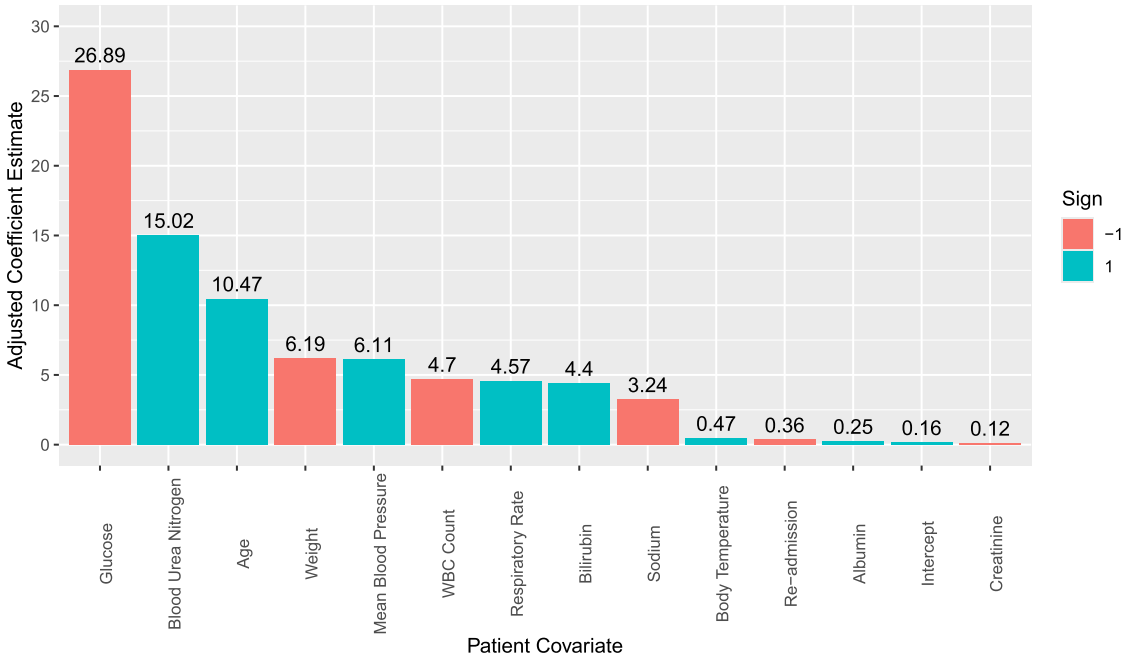


Figure 3: From left to right, the covariates are ranked by the magnitude of their corresponding coefficients in the optimal weighted ITR. Colored by the signs of the coefficients.

how well the CAIPW estimator estimates treatment effects as well as the maximization of the value function to obtain optimal ITRs.

5.1 Simulation Data

We generate a dataset with 50,000 rows with each row representing an individual patient. Each patient has height and age as covariates, and both covariates are uniformly distributed. This is our general population. We randomly sample 10,000 observations from the general population to use as Real World Data (RWD), i.e., the target population. While usually not available in real practice, we also simulated the target population's treatment and outcome data.

For treatment data, we first generate propensity scores for all patients in the target population based on a linear model of their covariates. With the propensity scores, we can then model individual Bernoulli trials to attach binary treatment assignments to the patients.

For outcome data, although in reality, we can only observe one outcome for each patient – under either treatment or control – we simulated both outcomes for each patient. In this simulation, we establish an individual treatment effect (ITE) condition that if a patient is taller than 55 inches and less than 41 years old, they will have a positive response to treatment; otherwise, the treatment will have a negative effect on them, and they should be assigned control. We first start with the control outcome, which is a linear combination of the patient covariates with the addition of a normal error term. Then, we formulate a contrast function that produces a positive value if a patient satisfies the positive treatment effect condition, and vice versa. Lastly, to generate the treatment outcome, we simply add the value of the contrast function to the control outcome. Since we have the treatment assignment and both outcomes, we can identify the observed outcome and the missing outcome for each patient.

We then generate sampling scores for all observations in the general population minus the target population based on a linear model of their covariates with a bias to observations with greater age value and greater height value. With sampling scores, we can then model individual Bernoulli trials, to decide whether each of these observations will be sampled in a Random Clinical Trial (RCT), i.e., the source population. Because of the bias in the sampling scores, patients sampled in the source population are more likely to be old and tall. Following an actual random clinical trial design, we give each patient in the source population an equal chance to be put into the treatment group (50%) or the control group (50%). The outcome data for the source population is generated the same way as for the target population.

5.2 Simulation Results

In our simulation, we generate treatment and outcome information for the target population. This would not be available in a real scenario since our end goal is to assign the best treatment to patients in the target population to maximize their outcome. However, this information is useful to establish the best ITR for the target population as a standard to compare to. The true optimal linear ITR for the target population (shown in red in Figure 4) yields a 94.5% correct classification rate (CCR), i.e., 94.5% of the patients in the target population received the treatments that gave them the better outcome under the true optimal ITR.

If we derive the ITR from only the source population information and there is a population heterogeneity problem between the source and target populations, the resulting ITR is only the optimal ITR for the source population, but not the optimal ITR for the target population. It can be considered as an unweighted ITR for the target population (shown in blue in Figure 4).

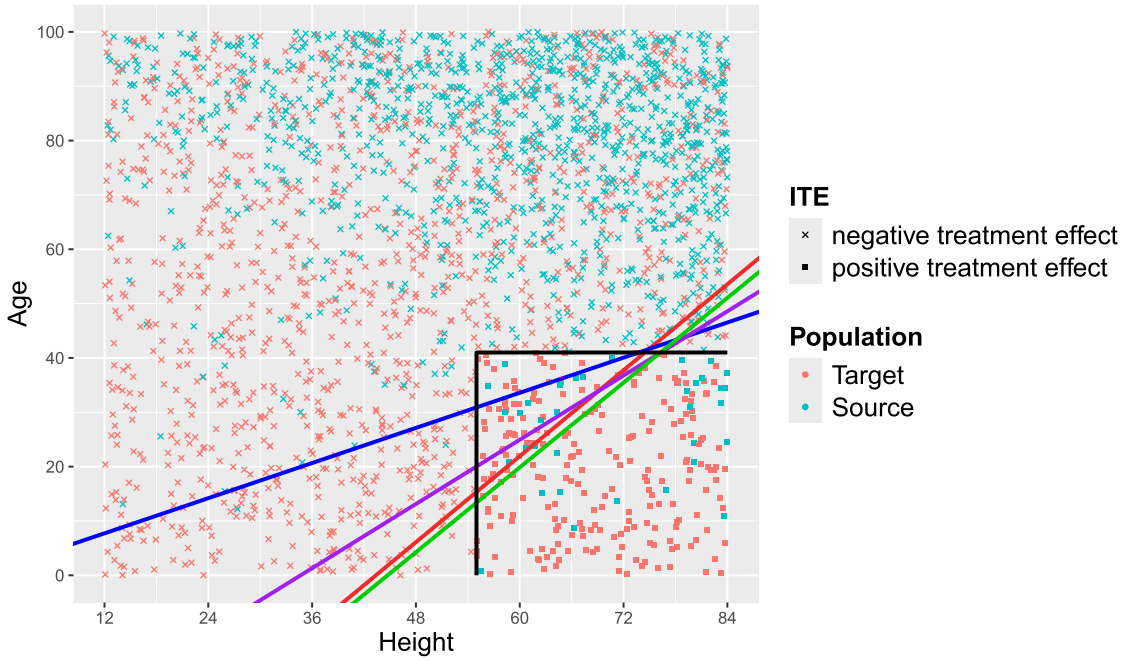


Figure 4: Calibration weighting and GA optimization improve estimation of ATE (blue line: unweighted ITR; purple line: one random weighted ITR; green line: weighted ITR optimized with GA; red line: true optimal ITR).

Table 2: Simulation study results comparing treatment rules in terms of classification accuracy, estimated value for the value function, uncertainty of value function using 95% confidence interval, false positive rate (FPR) and false negative rate (FNR).

Treatment Rule	**CCR**	**Value**	**Wald 95% CI**	**FPR**	**FNR**
Unweighted ITR	0.8598	146.3	[145.3, 147.3]	0.1512	0.0846
Calibration-weighted ITR	0.9360	146.8	[145.8, 147.8]	0.0394	0.1880
GA-optimized ITR	0.9431	146.9	[145.8, 147.9]	0.0177	0.2545
True Optimal ITR	0.9450	147.3	[146.2, 148.3]	–	–

In our simulation with heterogeneity of the covariate distribution, this ITR–although well fit for the source population–is not a good fit for the target population, having a CCR of 86.0%.

When we compute the calibration weights, we incorporate covariates from both the source and the target populations to address the population heterogeneity problem, taking full advantage of the information we can obtain in a real-world scenario. The resulting calibration-weighted ITR (shown in purple in Figure 4) has an improved CCR of 93.6%. Compared to the unweighted ITR, it is much closer to the true optimal ITR for the target population.

After demonstrating the effect of calibration weighting, we moved on to demonstrating the effect of using GA optimization in ITR development by producing the optimal weighted linear ITR that maximizes patient outcomes. After 100 iterations of GA with 100 different random seeds that generate different samplings of the population pool, the algorithm yields an ITR (shown in green in Figure 4) with the largest value for the value function, and a CCR of 94.31%, which closely approximates the true optimal ITR for the target population.

To quantify uncertainty in the estimated value function for each treatment rule, we compute a 95% confidence interval using an asymptotic Wald approach. These intervals reflect sampling variability in the estimated value of each treatment rule and allow for principled comparison of policy performance.

Observing Table 2, we notice that the 95% confidence intervals for different options of ITRs all overlap each other. While the GA-optimized ITR outperforms other ITRs in terms of having high CCR, high value for value function, low false positive rate, it also yielded higher false negative rate, which indicates that it is the most conservative rule that prevents overtreatment in a real-world healthcare application.

6 Discussion

6.1 Advantages

The transfer learning framework has many advantages and impacts for the statistical and clinical communities. When used appropriately, it can be a resourceful tool that bridges diverse sources and target populations, connecting data from various sources to account for patient heterogeneity, and informing more accurate decisions. It embraces the essence of precision medicine by providing precise treatment recommendations tailored to each patient's characteristics, assisting clinicians across health care systems globally to make data-driven treatment decisions for patients. In addition, transfer learning is a transferable framework and can be applied in other fields of study such as marketing, technology, social sciences, and education.

6.2 Limitations

There are some limitations to the work we have done. A limitation in our proposed methods resides in the Genetic Algorithm (GA) we used to optimize the value function. Like many other optimization algorithms, GA can become computationally expensive when applied to large datasets. It is also sensitive to the choice of initial values, i.e., the randomly selected initial population of candidate solutions. GA can only be applied under a moderate covariate size. When there are too many covariates, the search space can be too large to find the global maximum. Additionally, while our current approach uses GA for flexibility in optimizing complex, non-convex value functions, this method does not directly provide uncertainty quantification for the covariate coefficient point estimates due to its black-box nature.

In our real-world application, we removed covariates from either dataset due to either discreteness or a high level (greater than 60%) of missingness. The removed covariates include biomarkers such as arterial pH, PaO2 (the partial pressure of oxygen in the arterial blood), pCO2 (partial pressure of carbon dioxide), FiO2 (fraction of inspired oxygen), etc., which may provide useful information to assess whether mechanical ventilation should be used on individual patients to improve outcome. However, the large missingness of these biomarkers makes it difficult to perform data imputation. Therefore, we had to discard these biomarkers from our model. Additionally, due to the lack of survival outcome information in the real-world data, we are unable to confirm the real-world efficacy of the treatment rule generated by our method.

6.3 Comparison with a Similar Approach

An alternative approach to correcting population shift is the use of transfer weights, as introduced by (Wu and Yang, 2023). While our method adopts entropy balancing (Hainmueller, 2012) to obtain calibration weights that enforce exact balance on moments of covariates between the source and target populations, Wu and Yang consider a generalized formulation that allows approximate balance via inequality constraints. Specifically, they solve a constrained optimization problem that minimizes the relative entropy of weights subject to bounded deviations between weighted covariate moments in the experimental sample and unweighted moments in the real-world data. The bounds are defined by user-specified tolerance parameters, which offer a bias-variance tradeoff mechanism: tighter constraints reduce bias at the cost of higher variance, while looser constraints improve stability. In contrast, our calibration weights strictly enforce equality of covariate moments, which means all tolerance parameters are set to zero. This simplifies interpretation and guarantees exact balance, but can result in more extreme weights in finite samples.

6.4 Future Work

In our proposed framework, we used logistic regression and linear regression for estimation of the propensity score and outcome regression components of the CAIPW estimator. However, we plan to explore a variety of semi-parametric and nonparametric estimation techniques such as generalized additive models, spline regression, random forests, and Super Learners. These techniques may be more advantageous for capturing non-linear relationships, particularly within high-dimensional covariate-outcome spaces, further improving the robustness of the CAIPW estimator.

To address the lack of uncertainty quantification on the covariate coefficient estimates due to our use of GA, we plan to explore alternative optimization strategies such as gradient-based

methods and penalized M-estimation approaches, which enable valid statistical inference for the estimated ITR coefficients. Since gradient-based optimization requires the objective function to be differentiable, we may consider smooth approximations of the hard threshold function used in our linear ITR definition. Similarly, because penalized M-estimators rely on convexity, we may reformulate the ITR estimation problem as a risk minimization task with surrogate loss functions (e.g., logistic or hinge loss), facilitating both tractable optimization and inference.

In our current approach, we restrict the class of individualized treatment rules to binary linear rules due to their interpretability and clinical applicability. However, this restriction leads to a non-convex and non-differentiable value function, limiting our options for optimization algorithms for the value function. For example, gradient-based optimization methods such as Newton's method is not suitable for our approach. In the future, we plan to explore more flexible function classes for ITRs. For example, replacing the hard decision boundary with a smooth surrogate such as a sigmoid function yields a continuous and differentiable value function. Under such formulations, we can systematically compare the performance and efficiency of various optimization strategies, including Newton's method, quasi-Newton methods, coordinate descent, and stochastic gradient-based approaches.

Tree-based ITRs are a promising alternative to linear ITRs in terms of interpretability. ITRs in a decision-tree form more closely mimic human decision-making, which enhances communication between statisticians and medical professionals. Additionally, tree-based rules naturally perform variable selection and model simplification through pruning, making the resulting treatment rules both interpretable and parsimonious.

Supplementary Material

Supplementary materials include pre-processed eICU-CRD and MIMIC-III data files used in the medical application, an R script containing the R functions and R Markdown files for both the simulation study and the medical application.

Funding

The authors gratefully acknowledge the generous support from the National Science Foundation (NSF) grant DMS2051010 and National Security Agency (NSA) grant H98230-22-1-0006. This research is also supported in part by the National Institute of Environmental Health Sciences (NIEHS) training grant T32ES007018.

References

Azur M, Stuart E, Frangakis C, Leaf P (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1): 40–49. PMID: 21499542; PMCID, PMC3074241. https://doi.org/10.1002/mpr.329

Bang H, Robins JM (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4): 962–973. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Chu J, Lu W, Yang S (2023). Targeted optimal treatment regime learning using summary statistics. *Biometrika*, 110(4): 913–931. https://doi.org/10.1093/biomet/asad020

Colnet B, Mayer I, Chen G, Dieng A, Li R, Varoquaux G, et al. (2024). Causal inference methods for combining randomized trials and observational studies: A review. *Statistical Science*, 39(1): 165–191. https://doi.org/10.1214/23-STS889

Glynn A, Quinn KM (2010). An introduction to the augmented inverse propensity weighted estimator. *Political Analysis*, 18(1): 36–56. https://doi.org/10.1093/pan/mpp036

Hainmueller J (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1): 25–46. https://doi.org/10.1093/pan/mpr025

Imbens G, Rubin D (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. https://doi.org/10.1017/CBO9781139025751.

Katoch S, Chauhan S, Kumar V (2021). A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications*, 80: 8091–8126. https://doi.org/10.1007/s11042-020-10139-6

Kosorok MR, Laber EB (2019). Precision medicine. *Annual Review of Statistics and Its Application*, 6: 263–286. https://doi.org/10.1146/annurev-statistics-030718-105251

Mebane Jr WR, Sekhon JS (2011). Genetic optimization using derivatives: The rgenoud package for R. *Journal of Statistical Software*, 42(11): 1–26. https://doi.org/10.18637/jss.v042.i11

Nagin D, Paternoster R (2000). Population heterogeneity and state dependence: State of the evidence and directions for future research. *Journal of Quantitative Criminology*, 16(2): 117–144. https://doi.org/10.1023/A:1007502804941

Rothwell PM (2005). External validity of randomised controlled trials: "to whom do the results of this trial apply?". *The Lancet*, 365(9453): 82–93. https://doi.org/10.1016/S0140-6736(04)17670-8

Ryall B, Eydallin G, Ferenci T (2012). Culture history and population heterogeneity as determinants of bacterial adaptation: The adaptomics of a single environmental transition. *Microbiology and Molecular Biology Reviews*, 76(3): 597–625. https://doi.org/10.1128/mmbr.05028-11. https://doi.org/10.1128/MMBR.05028-11

Wu L, Yang S (2023). Transfer learning of individualized treatment rules from experimental to real-world data. *Journal of Computational and Graphical Statistics*, 32(3): 1036–1045. https://doi.org/10.1080/10618600.2022.2141752

Zampieri FG, Mazza B (2017). Mechanical ventilation in sepsis: A reappraisal. *Shock*, 47(1S): 41–46. PMID: 27454388. https://doi.org/10.1097/SHK.0000000000000702

Zhang B, Tsiatis AA, Laber EB, Davidian M (2012). A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4): 1010–1018. https://doi.org/10.1111/j.1541-0420.2012.01763.x

Zhao Y, Zeng D, Rush AJ, Kosorok MR (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499): 1106–1118. https://doi.org/10.1080/01621459.2012.695674

Zhou X, Mayer-Hamblett N, Kosorok MR (2017). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*, 112(517): 169–187. https://doi.org/10.1080/01621459.2015.1093947

Zubizarreta JR (2015). Stable weights that balance covariates for estimation with incomplete outcome data. *Journal of the American Statistical Association*, 110(511): 910–922. https://doi.org/10.1080/01621459.2015.1023805