BIOMETRIC METHODOLOGY

Relative contrast estimation and inference for treatment recommendation

Muxuan Liang1 | Menggang Yu2

1Department of Biostatistics, University of Florida, Gainesville, Florida, USA
2Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin, USA

Correspondence
Menggang Yu, Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA. Email: meyu@biostat.wisc.edu

Abstract
When there are resource constraints, it may be necessary to rank individualized treatment benefits to facilitate the prioritization of assigning different treatments. Most existing literature on individualized treatment rules targets absolute conditional treatment effect differences as a metric for the benefit. However, there can be settings where relative differences may better represent such benefit. In this paper, we consider modeling such relative differences formed as scale-invariant contrasts between the conditional treatment effects. By showing that all scale-invariant contrasts are monotonic transformations of each other, we posit a single index model for a particular relative contrast. We then characterize semiparametric estimating equations, including the efficient score, to estimate index parameters. To achieve semiparametric efficiency, we propose a two-step approach that minimizes a doubly robust loss function for initial estimation and then performs a one-step efficiency augmentation procedure. Careful theoretical and numerical studies are provided to show the superiority of our proposed approach.

KEYWORDS
individualized treatment rule, observational study, precision medicine, semiparametric efficiency, single index model

1 | INTRODUCTION

In this era of precision medicine, methods for constructing individualized treatment rules (ITRs) are flourishing. Most of these methods estimate conditional average treatment effect (CATE), the difference of conditional means of the outcomes under different treatments. Q-learning estimates the CATE based on separate outcome models for the treated and control groups (Chakraborty et al., 2010; Laber et al., 2014; Qian & Murphy, 2011; Watkins & Dayan, 1992). A-learning, on the other hand, directly models the CATE and therefore typically makes fewer modeling assumptions (Lu et al., 2013; Murphy, 2003; Shi et al., 2016, 2018; Song et al., 2017). In particular, Song et al. (2017) proposed a single index model for the CATE. Liang and Yu (2022) considered more general semiparametric models for the CATE. Zhao et al. (2012) pioneered an outcome weighted learning (O-learning) approach for ITR from a classification perspective. The work spurred many subsequent papers (Chen et al., 2017; Tian et al., 2014; Xu et al., 2015; Zhao et al., 2012, 2019; Zhang et al., 2012).

When resource constraints prohibit assigning all patients to their optimal treatments, it becomes necessary to prioritize assignments according to magnitudes of individualized treatment effects (ITEs). In particular, the CATEs can be compared, and patients with larger CATEs may have a higher priority for assignment to their optimal treatments. For example, Fan et al. (2017) proposed to use a linear index that is concordant with the CATE to rank the ITEs. Under resource constraints, many existing studies...
considered an optimal ITR maximizing the averaged outcome, or equivalently the CATE over patients who will be treated (Luedtke & Laan, 2018; Qiu et al., 2022, e.g.).

However, the studied CATE in the existing literature above is an absolute difference. There can be settings where relative differences are better individualized metrics to reflect treatment benefits for outcomes such as cost-saving, quality of life improvement, and disease response rate. In particular, a fixed or absolute amount of cost-saving may carry different levels of importance for subjects with different wealth. Similarly, an absolute amount of decrease in health outcome measurements such as blood pressure, blood sugar level, and cholesterol level can have different values for patients depending on their baseline levels of these measurements. In an analysis of a multiple sclerosis dataset, Yadlowksy et al. (2021) argued that the ratio of the expected relapse rate under different treatments is a better individualized metric for treatment recommendation.

Our motivating example is on recommending a tailored phone counseling session over the standard non-tailored care to improve mammography screening rates (Champion et al., 2016). The tailored counseling intervention for mammography includes messages based on subjects’ behavioral constructs of perceived susceptibility, benefits, barriers to screening, and self-efficacy. In general, tailored counseling is resource-demanding but can be more effective than non-tailored interventions (Champion et al., 2000, 2003, 2007). One interesting question raised by the study is whether the intervention effect depends on key patient characteristics (Champion et al., 2016). Because women from different social and economic backgrounds can have very different mammography screening rates, relative improvement is deemed as a more appropriate metric for intervention recommendation.

This paper focuses on estimating a class of relative contrasts of treatment effects on non-negative outcomes. In a related work, Kosorok and Moodie (2015) extended the O-learning approach to incorporate a reject option for treatment assignment recommendation. The recommended subsets are shown to be determined by a specific relative contrast. Yadlowksy et al. (2021) considered an exponential regression model for a ratio-based contrast, assuming homogeneous variances of the outcomes in both treatment and control groups.

While our method is linked with these existing works, we tackle this problem in a very general framework by considering all possible scale-invariant contrasts of treatment effects. By showing that all such contrasts are monotonic transformations of each other, we posit a single index model with a monotone link function for a particular relative contrast. Under this model, the estimated index can be used to rank patients as long as some scale-invariant contrast is an appropriate metric for the individualized treatment benefits. To estimate the index parameters, we derive a semiparametrically efficient score and adopt a two-step approach for computation. In the first step, we work with a doubly robust loss motivated by a simplification of the efficient score to obtain an initial estimator. In the second step, we improve the estimator by a one-step efficiency augmentation. The improved estimator is shown to be semiparametrically efficient and robust to possible heterogeneous variances.

### 2 | MODELING THE RELATIVE CONTRASTS

An ITR is a decision rule that maps from the space of the covariates \( \mathcal{X} \) to the treatments \( \mathcal{T} = \{-1, 1\} \). Given a decision rule \( d \) and patient covariates \( X \), treatment recommendation is given by \( d(X) \). Let \( (Y_1, Y_{-1}) \) be the potential outcomes corresponding to the treatments \( T = 1 \) and \( T = -1 \), respectively. The ITE is commonly reflected through the CATE defined as \( \Delta(X) = E[Y_1 | X] - E[Y_{-1} | X] \), where \( E[Y_1 | X] = E[Y | X] \). The optimal ITR aims to select the better treatment for each individual, thus \( d_{opt}(X) = \text{sign}(\Delta(X)) \).

In this work, assuming that \( \mu_1(X) > 0 \) and \( \mu_{-1}(X) > 0 \), we focus on the causal estimand:

\[
\theta(X) = \log \left\{ \frac{\mu_1(X)}{\mu_{-1}(X)} \right\}.
\] (1)

The causal estimand \( \theta(X) \) plays a key role in a class of causal estimands that is scale-invariant. To start with, we define a scale-invariant relative (SIR) contrast function \( h : R^+ \times R^+ \rightarrow R \) as a continuous function on positive numbers satisfying

- **Property 1.** \( h(a, a) = 0; \)
- **Property 2.** \( h(a, b) \) is strictly increasing in \( a \) for any fixed \( b; \)
- **Property 3.** \( h(ka, kb) = h(a, b), \) for all \( k > 0. \)

Property 1 imposes that if \( \mu_1(X) = \mu_{-1}(X) \), we have \( h(\mu_1(X), \mu_{-1}(X)) = 0 \). Property 2 guarantees that \( h(\mu_1(X), \mu_{-1}(X)) > 0 \iff \Delta(X) > 0 \). Property 3 implies that \( h \) is invariant to scale changes of the outcomes. Corresponding to a given SIR contrast function \( h \), \( h(\mu_1(X), \mu_{-1}(X)) \) is defined as an SIR contrast. Especially, \( \theta(X) \) is an SIR contrast with \( h(a, b) = \log(a/b) \).

In addition to \( \theta(X) \), there are many choices of SIR contrast functions. For example, \( (a - b)/b \) and \( (a - b)/(a + b) \) are both SIR contrast functions. Their corresponding
SIR contrasts are $\Delta(X)/\mu_-(X)$ and $\Delta(X)/S(X)$, where $S(X) \equiv \mu_1(X) + \mu_-(X)$. In practice, the choice of $h$ can vary upon scientific needs. The following important result indicates that all SIR contrasts are intrinsically related to $\theta(X)$.

**Theorem 1.** Any SIR contrast is a strictly increasing transformation of $\theta(X)$; for any two SIR contrast functions $h$ and $\tilde{h}$, there exists a strictly increasing mapping $g$ such that their corresponding SIR contrasts satisfy $\tilde{h}(\mu_1(X), \mu_-(X)) = g(h(\mu_1(X), \mu_-(X)))$.

Theorem 1 is proved in the online Supporting information of this paper. It motivates us to model SIR contrasts using single index models. Suppose that there exists an SIR contrast based on $h$ that follows a single index model $h(\mu_1(X), \mu_-(X)) = \psi(X^\top \beta)$, where $\psi$ is an increasing function and $\beta$ is a $p$-dimensional coefficient vector. From Theorem 1, for any other SIR contrast function $\tilde{h}$, there exists a strictly increasing $g$ such that $\tilde{h}(\mu_1(X), \mu_-(X)) = g(\psi(X^\top \beta))$. As such, a single index model assumption for one SIR contrast is equivalent to single index models for all SIR contrasts.

Therefore, we focus on modeling $\theta(X)$ using a single-index model, that is,

$$\theta(X) = \psi_0(X^\top \beta_0),$$  

(2)

where $\psi_0$ is an increasing unknown wrapper function and $\beta_0$ is a $p$-dimensional vector. Note that because $\psi_0$ is unknown, $\beta_0$ is not identifiable up to a positive multiplicative scale. Therefore, we force $\|\beta_0\| = 1$ to make it identifiable similar to Carroll et al. (1997), Liang et al. (2010), Ma and He (2016), and Song et al. (2017). In Yadlowsky et al. (2021), authors assumed that $\mu_1(X)/\mu_-(X) = \exp[\theta(X)]$, where $\theta(X)$ following a linear model; thus, our work can also be considered as an extension to Yadlowsky et al. (2021).

### 3.1 Orthogonal tangent space and efficient score

The following theorem characterizes the orthogonal tangent space of the nuisance parameters of observed data under Model (2).

**Theorem 2.** Under Model (2), the likelihood of the observed data is

$$p_X(X)p_T(X)p_{\epsilon}(Y - TS(X)e^{\psi_0(X^\top \beta)} - 1)/((2e^{\psi_0(X^\top \beta)} + 2 - S(X))/2, X, T),$$

(3)

where $p_X(\cdot)$ and $p_T(\cdot, X, T)$ are the density function of $X$ and $\epsilon \equiv Y - E(Y | X, T)$. The orthogonal tangent space corresponding to the sum of the tangent spaces of $\pi_T(X) = P(T | X)$, $S(X) = \mu_1(X) + \mu_-(X)$, and $\psi_0$ is

$$\left\{ \pi_T^{-1}(X)^T \left[ Y \left\{ 1 + e^{-T\psi_0(X^\top \beta)} \right\} - S(X) \right] \right\} \alpha(X) - \frac{E[S(X)\alpha(X) | X^\top \beta]}{E[S(X) | X^\top \beta]} : \alpha(X) \in L_2(X) \right\}.$$

(4)

By projecting the score function of $\beta_0$ onto the orthogonal tangent space, we can obtain the efficient score for $\beta_0$.

**Theorem 3.** Denote $\sigma_T^2(X) = E[\epsilon^2 | T, X]$. The efficient score for $\beta_0$ under Model (2) is

$$\pi_T^{-1}(X)^T \left[ Y \left\{ 1 + e^{-T\psi_0(X^\top \beta)} \right\} - S(X) \right] \psi_0(X^\top \beta) \alpha_\psi(X),$$

where $\psi_0(\cdot)$ is the derivative function of $\psi_0(\cdot)$,

$$\alpha_\psi(X) = V(X)S(X) \left[ X - \frac{E[V(X)S^2(X)|X^\top \beta]}{E[V(X)S^2(X)|X^\top \beta]} \right],$$

where $V(X)$ and $S(X)$ are the density function of $X$ and $\epsilon$ respectively, $\psi_0$ is an increasing unknown wrapper function, and $\alpha_\psi(X) \in L_2(X)$.
and

\[ V^{-1}(X) = \left\{ \pi_1^{-1}(X)e^{-\psi_0(X^T\beta)}\sigma_1^2(X) + \pi_2^{-1}(X)e^{\psi_0(X^T\beta)}\sigma_2^{-1}(X) \right\} \]

\[ \left\{ e^{\psi_0(X^T\beta)/2} + e^{-\psi_0(X^T\beta)/2} \right\}^2 \].

Proofs of Theorems 2 and 3 are in the online Supporting information. To use the efficient score directly for estimating \( \beta_0 \), one faces a couple of challenges. First, the efficient score involves \( \pi_0(X) \), whose estimation can be challenging as both \( E[V(X)S^2(X) \mid X^T\beta] \) and \( E[V(X)S^2(X)X \mid X^T\beta] \) involve \( \beta \). Second, the gradient of the efficient score is not of full rank due to the constraint \( \|\beta\|_2 = 1 \). Thus, we consider an alternative two-step solution to avoid direct solving the efficient score. In the first step, we estimate an initial value of \( \beta_0 \) through a doubly robust loss function. In the second step, we perform a one-step efficiency augmentation procedure. Both steps are motivated by Theorem 3.

### 3.2 The first step: minimizing a doubly robust loss function

In this section, we propose a doubly robust loss function to provide an initial estimate of \( \beta_0 \). We start with a simplified estimating equation that takes a similar form as the efficient score,

\[ \pi_T^{-1}(X)T \left[ Y \left\{ 1 + e^{-\psi_0(X^T\beta)} \right\} - S(X) \right] \psi_0(X^T\beta) \alpha(X). \] (5)

In general, we can replace \( X \) by \( \alpha(X) \) in this function and use

\[ \pi_T^{-1}(X)T \left[ Y \left\{ 1 + e^{-\psi_0(X^T\beta)} \right\} - S(X) \right] \psi_0(X^T\beta) \alpha(X) \]

as an estimating equation, where \( \alpha(X) \in L_2(X) \) is any function. For example, the estimating equation in Yadlowsky et al. (2021) corresponds to the choice of

\[ \alpha(X) = X\pi_3(X)\pi_{-3}(X) \left\{ 1 + e^{\psi_0(X^T\beta)}\pi_1(X) + e^{-\psi_0(X^T\beta)}\pi_{-1}(X) \right\}^{-1}. \]

Let \( \ell(\psi;\pi_T, S) = \pi_T^{-1}(X) Y \exp[-T\psi] + \pi_T^{-1}(X)T(S(X) - Y)\psi \) and \( \hat{\ell} \) and \( \hat{\ell} \) are first derivative and second derivative of \( \ell \) with respect to \( \psi \), respectively. Equation (5) can be viewed as the gradient of \( \ell(\psi_0(X^T\beta);\pi_T, S) \) with respect to \( \beta \). Therefore under Model (2), we propose to minimize \( E_n[\ell(\psi(X^T\beta);\pi_T, S)] \) with respect to \( \psi \) and \( \beta \). Theorem 4 implies that the true \( (\psi_0, \beta_0) \) is the unique minimizer of the proposed loss function when either \( \pi_T(X) \) or \( S(X) \) is correctly specified (therefore doubly robust). The proof of Theorem 4 can be found in the online Supporting information.

**Theorem 4.** If either \( \pi_T(X) \) or \( S(X) \) is correctly specified, under the assumption that \( \mu_1(X) > 0 \) and \( \mu_{-1}(X) > 0 \), \( \hat{\ell}(X) \) is the unique minimizer (almost surely) of the loss function

\[ E[\ell(f(X); \pi_T, S)] \]

over all bounded measurable functions, where

\[ \ell(f(X); \pi_T, S) = \pi_T^{-1}(X) Y \exp[-T f(X)] + \pi_T^{-1}(X) T(S(X) - Y) f(X). \]

To estimate the wrapper function \( \psi_0 \) in Model (2), similar to Song et al. (2017), we use monotone B-splines. That is, we let \( \psi(t) \sim \sum_{j=1}^{K_n+M} \xi_j N_j(t) \), where \( N_j \)'s are B-spline bases, \( K_n \) is the number of interior knots in an bounded interval containing \( X^T\beta \), and \( M \) is the order of the B-spline bases. Let \( \xi \equiv (\xi_1, \ldots, \xi_d)^T \) with \( d = K_n + M \). In general observational studies, in addition to \( \xi \) and \( \beta \), the loss function \( \ell(\psi(X^T\beta); \pi_T, S) \) involves two more unknown nuisance parts: \( \pi_T \) and \( S \). For the ease of presentation, we first assume that these nuisance parts are known and then consider the setting when they need to be estimated.

To ensure the monotonicity of \( \psi \), we impose two constraints on \( \xi \) following Leitenstorfer and Tutz (2007). Specifically, let \( B \) be a \((d-1) \times d \) matrix with \( B(i, i) = 1, B(i, i+1) = -1 \), and 0 for other entries. Then, the first constraint is \( B \xi \leq 0 \) which requires \( \xi \) to be non-decreasing. The second constraint is on the L-1 norm of \( \xi \), \( \sum_{j=1}^{d} |\xi_j| \leq C \), where \( C \) is a sufficient large constant. Combining this approximation with the proposed loss function, we propose the following optimization problem:

\[ \min_{\xi, \beta} E_n[\ell(\xi, \beta; \pi_T, S)] \]

subject to \( B \xi \leq 0, \sum_{j=1}^{d} |\xi_j| \leq C, \) and \( \|\beta\|_2 = 1, \)

where

\[ \ell(\xi, \beta; \pi_T, S) = \pi_T^{-1}(X)Y \exp \left\{ -T \sum_{j=1}^{d} \xi_j N_j(X^T\beta) \right\} + \pi_T^{-1}(X) T(S(X) - Y) \sum_{j=1}^{d} \xi_j N_j(X^T\beta). \] (6)

To solve this optimization problem, we iterate between the two following steps until convergence. In the first step, we fix \( \beta \) and optimize over \( \xi \). This step solves a convex optimization problem with linear constraints. In the second step, we fix \( \xi \) and optimize \( \beta \) under the constraint \( \|\beta\|_2 = 1 \). This step solves a nonlinear optimization
Algorithm 1 Estimation of $\psi(X^\top \beta)$ with known $\pi_T$ and $S$.
1. Obtain an initial estimator $\hat{\beta}^{(0)}$; optimize a linear form $\psi(X^\top \beta) = X^\top \beta$ and then normalize it with $\|\hat{\beta}^{(0)}\|_2 = 1$. Set $l = 0$;
2. Given $\hat{\beta}^{(l)}$, solve the following to obtain $\hat{\xi}^{(l)}$:
   \[
   \min E_n \left[ \ell(\xi; \hat{\beta}^{(l)}; \pi_T, S) \right] \text{ subject to } B \xi \leq 0 \text{ and } \sum_{j=1}^d \xi_j \leq C;
   \]
3. Given $\hat{\xi}^{(l)}$, solve the following to obtain $\hat{\beta}^{(l+1)}$:
   \[
   \min_{B \beta} E_n \left[ \ell(\xi; \beta; \pi_T, S) \right] ;
   \]
4. Iterate Steps 2 and 3 until convergence. Let $\hat{\xi}$ and $\hat{\beta}$ be the estimates at the last iteration. We set $\hat{\psi}(\cdot) = \sum_{j=1}^d \hat{\xi}_j N_j(\cdot)$;

Algorithm 2 Estimation of $\psi(X^\top \beta)$ with unknown $\pi_T$ and $S$.
1. Randomly split data into two parts $Z^{(1)}$ and $Z^{(2)}$;
2. Estimate $\pi_T(X)$ and $S(X)$ on $Z^{(2)}$, denote the estimators as $\hat{\pi}_T(X)$ and $\hat{S}(X)$;
3. Obtain the estimates $\hat{\xi}^{(2)}$ and $\hat{\beta}^{(2)}$ on $Z^{(2)}$ by Algorithm 1 with the estimated $\hat{\pi}_T$ and $\hat{S}$ from Step 2. In particular, $\hat{\xi}^{(2)} = \sum_j \hat{\xi}_j^{(2)} N_j(\cdot)$;
4. Switch the order of $Z^{(1)}$ and $Z^{(2)}$, and repeat Steps 2 and 3 to obtain $\hat{\xi}^{(3)}$ and $\hat{\beta}^{(3)}$;
5. Aggregate to obtain $\hat{\beta} = \frac{1}{2} \sum_{i=1}^3 \hat{\beta}^{(i)}$ and $\hat{\psi} = \frac{1}{2} \sum_{i=1}^3 \hat{\psi}^{(i)}$.

### 3.3 The second step: one-step efficiency augmentation procedure

In this section, we investigate possible efficiency improvement for estimating $\beta_0$ using the efficient score. In Section 3.2, we propose a doubly robust loss function to obtain the estimators $\hat{\beta}$ and $\hat{\xi}$. Due to the misspecification of $\alpha(X)$, the $\hat{\beta}$ may not be semiparametrically efficient. Our efficiency augmentation procedure is a one-step procedure to obtain a semiparametrically efficient estimator based on $\hat{\beta}$.

From Theorem 3, a key element in using the efficient score requires evaluation of $\alpha_j(X)$, which depends on $V(X)$, $E[V(X)S(X)|X^\top \beta]$, and $E[V(X)S^2(X)|X^\top \beta]$. Suppose that we have initial estimators $\hat{\xi}$, $\hat{\beta}$, $\hat{\pi}(X)$, $\hat{S}(X)$, and estimators for the conditional variance terms $\sigma_j^2(X)$, then we can obtain a plug-in estimator for $V(X)$. We propose to estimate $\sigma_j^2(X)$ for $T = 1$ and $T = -1$ separately. For $T = 1$, after obtaining the estimated outcome model $\hat{\mu}_1(X)$, we can regress the squared-residual $[Y - \hat{\mu}_1(X)]^2$ on $X$ using kernel regression to estimate $\sigma_1^2(X)$. Similarly, $\sigma_{-1}^2(X)$ can be estimated from the $T = -1$ group. Further, we can estimate $E[V(X)S(X)|X^\top \beta]$ and $E[V(X)S^2(X)|X^\top \beta]$ by fitting $\hat{V}(X)S(X)|X^\top \beta$ and $\hat{V}(X)S^2(X)|X^\top \beta$ using kernel regressions. An estimator for $\alpha_j(X)$ can be formed by plugging in the resulting components.

Note that the derivative of Equation (6), with respect to $\sum_{j=1}^d \xi_j N_j(X^\top \beta)$ is

\[
\hat{\epsilon}(\xi; \beta; \pi_T, S) = \pi_T^{-1}(X)T \left\{ S(X) - Y(1 + e^{-T \sum_{j=1}^d \xi_j N_j(X^\top \beta)}) \right\},
\]

which is the negation of the first term of the efficient score. Therefore, we write the efficient score in Theorem 3 as

\[
-E_n \left[ \hat{\epsilon}(\xi, \beta; \pi_T, S) \hat{\psi}_0(X^\top \beta) \alpha_j(X) \right],
\]

where $\xi$ comes from the B-spline representation of $\psi$ as in Equation (6). Denote the estimator for $\alpha_j(X)$ as $\hat{\alpha}_j(X)$. We propose a modified one-step estimator based on the first-order expansion of Equation (7) as follows,

\[
E_n \left[ \hat{\epsilon}(\xi, \beta; \pi_T, S) \hat{\psi}(X^\top \beta) \hat{\alpha}(X) \right] +
\]
where $\lambda_n$ is a parameter, $I$ is an identity matrix with dimension $p \times p$.

$$\hat{\psi}(t) = \sum_{j=1}^{K_0+M} \hat{\xi}_j N_j(t), \quad \ell(\xi; \beta; \pi, S) = \pi^{-1}(X)Ye^{-T \sum_{j=1}^{d} \hat{\xi}_j N_j(X^T \beta)}.$$ 

Compared with the usual one-step estimator (Bolthausen et al., 2002), we have the extra term $\lambda_n I$ due to the matrix $\Sigma_1 \equiv E\left[\ell(\psi_0(X^T \beta_0); \pi_T, S)\left(\psi_0(X^T \beta_0)\right)^2 X\alpha_0^T(X)\right]$ not of full rank (the rank is $p - 1$), and its approximation

$$\tilde{\Sigma}_1 \equiv E_n\left[\ell(\xi; \beta; \pi_T, S)\left(\hat{\psi}(X^T \beta)\right)^2 X\hat{\alpha}(X)\right]$$

may not be invertible numerically. By adding the extra term $\lambda_n I$, the estimator $\tilde{\Sigma}_1$ can be invertible. Then, we define the one-step efficiency augmentation estimator $\tilde{\beta}$ as the solution of Equation (8),

$$\tilde{\beta} = \hat{\beta} + (\tilde{\Sigma}_1 + \lambda_n I)^{-1}E_n \left[\ell(\xi; \beta; \pi_T, S)\hat{\psi}(X^T \beta)\hat{\alpha}(X)\right].$$

In Section 4, we show that $\tilde{\beta}$ is semiparametrically efficient. Its asymptotic variance can be estimated by

$$\tilde{\Sigma} = (\tilde{\Sigma}_1 + \lambda_n I)^{-1}E_n \left[\ell(\xi; \beta; \pi_T, S)\hat{\psi}(X^T \beta)\hat{\alpha}(X)\right] \tilde{\alpha}(X)\hat{\alpha}_n^T(X)$$

$$= (\tilde{\Sigma}_1 + \lambda_n I)^{-1}.$$

To accommodate nonparametric estimators for $\pi_T(X)$, $S(X)$, and $\alpha(X)$, we extend the cross-fitting procedure by splitting the entire dataset into three parts. On the first part, we fit $\pi_T(X)$, $\mu_1(X)$, and $\mu_0(X)$; on the second part, we utilize the proposed doubly robust loss with estimated $\pi_T(X)$ and $S(X)$ to get initial estimators $\hat{\xi}$ and $\hat{\beta}$; on the third part, we improve the estimator using the efficiency augmentation procedure. To make up for the efficiency loss of sample-splittings, we rotate the role of each part and average the results. Details of this algorithm are listed in Algorithm 3.3.

**Algorithm 3** Semiparametrically efficient estimation of $\beta$ based on sample-splitting

1. Randomly split data into three parts $Z^{(i)}$, $Z^{(2)}$, and $Z^{(3)}$.
2. Denote $S = ((1, 2, 3), (2, 3, 1), (3, 1, 2))$ and $(k_1, k_2, k_3) = (1, 2, 3)$. Use $Z^{(i)}$ to obtain $\hat{\pi}^{(i)}(X)$ and $\hat{S}^{(i)}(X)$. Then use $Z^{(j)}$ to obtain $\hat{\xi}^{(i)}$ and $\hat{\beta}^{(i)}$ from

$$\min_{\beta \in \mathbb{R}^p} E_n \left[\ell(\xi, \pi, S)\right]$$

subject to $B \xi \leq 0$, $\sum_{i=1}^{d} |\xi_i| \leq C$, and $|B\xi|_1 = 1$, where $E_n(\cdot)$ represents the empirical mean on $Z^{(i)}$.
3. Obtain $\hat{\pi}^{(1)}(X)$, $\hat{S}^{(1)}(X)$, and $\hat{\alpha}^{(1)}(X)$ using the estimators in Step 2. To obtain $\hat{\alpha}^{(1; i)}(X)$, estimate $\hat{S}^{(1)}(X)\pi^{(1)}(X)$ using kernel regressions on $Z^{(i)} \cup Z^{(j)}$ based on $\hat{\pi}^{(1)}(X)$, $\hat{S}^{(1)}(X)$, $\hat{\beta}^{(1)}$ and $\hat{\beta}^{(j)}$. Likewise, estimate $\hat{\beta}^{(1)}(X)$ using kernel regressions on $Z^{(i)} \cup Z^{(j)}$.
4. Switch the order of $(k_1, k_2, k_3)$ and repeat Steps 2 - 3.
5. Let $\tilde{\beta} = \frac{1}{|S|} \sum_{(i, j) \in S} \beta_j, \tilde{\alpha} = \frac{1}{|S|} \sum_{(i, j) \in S} \alpha_i$.
6. Obtain $\tilde{\psi}(X)\pi^{(1)}(X)$ using kernel regressions on $Z^{(i)} \cup Z^{(j)}$.
7. Calculate $\hat{\beta} = \hat{\beta} + \tilde{\beta} + (\tilde{\Sigma}_1 + \lambda_n I)^{-1}\tilde{\alpha} \tilde{\psi}(X)$.
8. Aggregate to obtain $\hat{\beta} = \frac{1}{|S|} \sum_{(i, j) \in S} \hat{\beta}_i$ and $\hat{\psi}(X) = \hat{\psi}(X^T \beta)$.

4 | THEORETICAL PROPERTIES

In this section, we present the theoretical properties of our proposed estimation procedures in Algorithms 3.2 and 3.3. The proof of the main theorem can be found in the online Supporting information. Let $k$ be the smoothness parameter of the wrapper function $\psi_0$, which indicates that $\psi_0$ has at least $k$th bounded derivative over the support of $X^T \beta_0$.

The conditions that we need for our theoretical results are listed in the online Supporting information.

For the estimated propensity and outcome models, we also require the following conditions on the convergence rate of these estimated nuisance parameters. Note that these conditions are rather mild, and many machine learning methods satisfy these conditions. Due to the sample-splitting procedures in Algorithms 3.2 and 3.3, these conditions and corresponding conclusions are expected to hold for split datasets as their sample sizes are in the same order of the whole dataset.

**Assumption 1.** Suppose that there exist some $\alpha, \beta > 0$ such that $\|\hat{\pi}_a(x) - \pi_a(x)\|_\infty = O_p(n^{-\alpha})$ and $\|\hat{\mu}_a(x) - \mu_a(x)\|_\infty = O_p(n^{-\beta})$ for $a = 1$ and $-1$, we require that $\alpha + \beta > 1/2$.

**Assumption 2.** There exists two constants $\pi_{\text{min}}$ and $\pi_{\text{max}}$ satisfying $0 < \pi_{\text{min}} \leq \pi_a(x), \pi_{\text{max}} < 1$ for all $x$ and $a$.
Assumption 1 requires that the estimators for \( \pi_T(X) \) and \( S(X) \) are consistent, and the convergence rates can be as slow as \( n^{-1/4} \). In addition, the assumption mitigates the requirement of the convergence rate on \( S(X) \) when we have an estimator of \( \pi_T(X) \) with a relatively fast convergence rate and vice versa. For example, if \( \pi_T(X) \) can be estimated by logistic regression and the convergence rate of \( \hat{\pi}_T(X) \) is \( O_p(n^{-1/2}) \), we only require the convergence rate of \( \hat{S}(X) \) to be \( o_p(1) \). This indicates that \( S(X) \) can be obtained by many nonparametric methods. Further, if we use the kernel regression with the optimal bandwidth, the convergence rate of \( \hat{S}(X) \) is as fast as \((n^{-1} \log n)^{1/(2l+p)}\) for \( l \)-th order Lipschitz continuous \( E[Y \mid X] \) (Stone, 1982). If both nuisance parts are estimated by the kernel regression and have the same order of smoothness, we have \( n^{-s-p} = (n^{-1} \log n)^{1/(2l+p)} \) which is faster than the required \( n^{-1/2} \) when \( p < 2l \). Assumption 2 is commonly known as the positivity assumption for the propensity score (Rosenbaum & Rubin, 1983).

Under these assumptions and assumptions listed in the online Supporting information, we first establish consistency and convergence rates of the initial estimators \((\tilde{\beta}_n, \hat{\psi}_n)\). Details are relegated to the online Supporting information. Based on the consistency of the initial estimator, the following theorem provides the asymptotic normality of the one-step estimator \( \tilde{\beta} \) from Algorithm 3.3.

**Theorem 5.** Let \( K_n = C_1 n^\gamma \) for some positive constants \( C_1 \) with \( \gamma < 1/10 \), where \( \gamma \) controls the increase of \( K_n \) with respect to \( n \). Assume that \( n^{1/2} h(n^{-1/2+\eta/2+2\gamma} + n^{-(k-3/2)\gamma}) \to 0 \) for some \( \eta \in (0,1/2) \), where \( h = n^{-1/2} \min\{\alpha, \beta\} + n^{-1/2+\eta/4+\gamma} + n^{-(k-3/2)\gamma} + (n/\log n)^{-1/4(p+4)} \). If we set \( \lambda_n = h \), then \( \sqrt{n}(\beta - \beta_0) \) converges in distribution to a mean-zero normal distribution with covariance matrix \( \Sigma = \Sigma_1^+ \Sigma_2 \Sigma_1^+ \), where \( \Sigma_1^+ \) is the pseudo inverse of \( \Sigma_1 \), and the formulas of \( \Sigma_1 \) and \( \Sigma_2 \) are given in the online Supporting information. In addition, \( \tilde{\beta} \) is semiparametrically efficient.

Theorem 5 requires that \( n^{1/2} h(n^{-1/2+\eta/2+2\gamma} + n^{-(k-3/2)\gamma}) \to 0 \) for some \( \eta \in (0,1/2) \). The \( \eta \) is introduced to indicate that there can be many choices of \( h \) that satisfy this requirement. To see this, suppose that the requirement is satisfied for a constant \( \eta_0 \in (0,1/2) \), then the requirement is satisfied with any \( \eta \leq \eta_0 \) and

\[
h = n^{-1/2} \min\{\alpha, \beta\} + n^{-1/4+\eta/4+\gamma} + n^{-(k-3/2)\gamma} + (n/\log n)^{-1/4(p+4)}.
\]

This requirement is also related to the convergence rate of the asymptotic variance and the convergence rate of estimator for the derivative of the link function, that is, \( \hat{\psi} \). In Theorem 5, the asymptotic variance comes from a degenerated Gaussian distribution because \( \tilde{\beta}_0^T \Sigma_1 \tilde{\beta}_0 = 0 \) and \( \tilde{\beta}_0^T \Sigma_2 \tilde{\beta}_0 = 0 \). In the proof of Theorem 5, we show that the variance estimator in Algorithm 3.3 is consistent and has a convergence rate of \( h \). The requirement that \( n^{1/2} h(n^{-1/2+\eta/2+2\gamma} + n^{-(k-3/2)\gamma}) \to 0 \) for some \( \eta \in (0,1/2) \) guarantees that the product of the estimation error of the variance estimate and that of \( \hat{\psi} \) is negligible. This requirement can be easily satisfied when the smoothness parameter \( k \) of \( \psi_0 \) is large. To see this, suppose that the convergence rate of \( h \) is dominated by \( n^{-1/2 \min\{\alpha, \beta\}} \), then when \( k \) is large, we can choose a small \( \gamma \), or less number of basis functions, to approximate \( \psi \) and \( \hat{\psi} \) and enforce the requirement.

## 5 | SIMULATION

In this section, we conduct simulations to compare our proposed method with other methods. Specifically, we compare with the contrast regression approach (CRA) (Yadlowsky et al., 2021), the Q-learning method (Qian & Murphy, 2011), the single index model on the absolute difference function \( \Delta(X) \) (\( \Delta \)-learning) proposed in Song et al. (2017) and Liang and Yu (2022), the concordance-learning method (C-learning) (Fan et al., 2017), and the efficient augmentation and relaxation learning (EARL) method (Zhao et al., 2019). Q-learning is a regression-based method, which posits parametric assumptions on the outcome models. \( \Delta \)-learning is an extension of Song et al. (2017) and Liang and Yu (2022) by incorporating the cross-fitting procedure for improving efficiency and safeguard model misspecification. The C-learning method promotes a decision rule according to the rank of the contrast function rather than the relative contrast function (Fan et al., 2017). As an ad hoc procedure, we also implement the cross-fitting algorithm with C-learning to accommodate nuisance parameters fitted by nonparametric methods.

The EARL method is an efficiency augmented approach that extends the O-learning method of Zhao et al. (2012).

In these comparisons, we focus on two metrics evaluated on a testing dataset. The first metric is the rank correlation between the estimate \( \hat{\theta}(X) \) and the true \( \theta(X) \). The second metric is the value function defined as \( V(d) = E[Y_{d(X)}] \), where \( d(X) = \text{sgn}(\delta(X)) \). Given a testing dataset, we can empirically estimate this value function by \( \hat{V}(d) = E_n[Y_{d(X)}] \). In many practical settings, including our motivating example, the rank correlation is more interesting than the value function because the former facilitates the prioritization of patient recommendations. In addition to these comparisons, we also provide the coverage of the interval estimations.

For our proposed approach, we obtain \( \hat{\theta}(X) \) following Algorithm 3.3. Kernel regressions are used to estimate the propensity and outcome models. For the Q-learning
approach, we estimate \( \mu_1(\mathbf{X}) \) and \( \mu_{-1}(\mathbf{X}) \) first from linear models with all the covariates and the interaction terms between the covariates and treatment. Then, we construct a plug-in estimator for \( \theta \). For the \( \Delta \)-learning approach, we posit a single index model on the absolute difference \( \Delta(\mathbf{X}) = \mu_1(\mathbf{X}) - \mu_{-1}(\mathbf{X}) \) following Song et al. (2017). To obtain the relative contrast \( \theta(\mathbf{X}) \), we estimate \( S(\mathbf{X}) \) following a single index model using the relationship \( E[Y/\pi_T(\mathbf{X}) | \mathbf{X}] = S(\mathbf{X}) \). For the proposed method, the Q-learning, and the \( \Delta \)-learning, we use the estimated \( \theta(\mathbf{X}) \) to calculate the rank correlation and the value function. For the C-learning approach, instead of estimating \( \theta(\mathbf{X}) \), we use the estimated index as a substitute of \( \theta(\mathbf{X}) \) to calculate the rank correlation. For the EARL method, we consider the class of linear decision rules because the optimal ITR is linear in these settings. To calculate the rank correlation, we use the resulting estimate \( \mathbf{X}^\top \hat{\beta}_{\text{EARL}} \) as a substitute for \( \theta(\mathbf{X}) \) to calculate the rank correlation. For the C-learning and the EARL approach, the value function is calculated based on the learned decision rules.

In the simulations, we consider the following two settings:

**O1** All covariates \( \mathbf{X} \) are generated from independent uniform distribution on \([-0.5, 0.5]\). We set \( \theta(\mathbf{X}) = \log((1.5 + \Phi(\mathbf{X}^\top \hat{\beta}_0))/\{2.5 - \Phi(\mathbf{X}^\top \hat{\beta}_0)\}) \), where \( \hat{\beta}_0 = (1, -1, 1, -1)^\top \) and \( \Phi(\cdot) \) is the cumulative distribution function of standard normal \( N(0, 1) \). We further set \( S(\mathbf{X}) = \|\mathbf{X}\|_2 \). Then from the relationship \( \mu_1(\mathbf{X}) = e^{\theta(\mathbf{X})} S(\mathbf{X})/(e^{\theta(\mathbf{X})} + 1) \) and \( \mu_{-1}(\mathbf{X}) = \mu_1(\mathbf{X})/e^{\theta(\mathbf{X})} \), we generate \( Y_1 = \mu_1(\mathbf{X}) + N(0, 0.01) \) and \( Y_{-1} = \mu_{-1}(\mathbf{X}) + N(0, 0.01) \).

**O2** All covariates \( \mathbf{X} \) are generated from independent uniform distribution on \([-1, 1]\). Here, \( \theta(\mathbf{X}) = \log((0.5 + \Phi(0.8 \mathbf{X}^\top \hat{\beta}_0))/\{1.5 - \Phi(0.8 \mathbf{X}^\top \hat{\beta}_0)\}) \), with \( \hat{\beta}_0 = (-1, -1, 1, -1)^\top \). We set \( S(\mathbf{X}) = \exp(0.8 \mathbf{X}^\top \hat{\beta}_0) + 1 \). We generate \( Y_1 \) and \( Y_{-1} \) according to Poisson distributions with mean \( \mu_1(\mathbf{X}) \) and \( \mu_{-1}(\mathbf{X}) \).

For the propensity model, we consider the following two settings: **PS1** \( P(T = 1|\mathbf{X}) = e^{\mathbf{X}^\top \hat{\beta}_\pi}/(1 + e^{\mathbf{X}^\top \hat{\beta}_\pi}) \), where \( \hat{\beta}_\pi = (0.2, -0.2, -0.2, 0.2)^\top \); **PS2** \( P(T = 1|\mathbf{X}) = \exp\{0.2(X_1^2 + X_2^2 + X_3X_4)/[1 + \exp\{0.2(X_1^2 + X_2^2 + X_3X_4)]\} \). Therefore, we consider four simulation scenarios (O1)+PS1), (O2)+PS1), (O1)+PS2), and (O2)+PS2). In the online Supporting information, more simulation with correlated variables are available. In each simulation, the sample size in the training dataset varies from 800, 1,000, to 1,500. To evaluate different methods, we generate a testing dataset with a sample size of \( 10^3 \). Each simulation is repeated 500 times.

| Sample size | O1+PS1 | O2+PS1 | O1+PS2 | O2+PS2 |
|-------------|--------|--------|--------|--------|
| 800         | 0.974  | 0.930  | 0.956  | 0.952  |
| 1000        | 0.972  | 0.948  | 0.974  | 0.958  |
| 1500        | 0.964  | 0.930  | 0.988  | 0.954  |

For the propensity model, we consider the following two settings: **PS1** \( P(T = 1|\mathbf{X}) = e^{\mathbf{X}^\top \hat{\beta}_\pi}/(1 + e^{\mathbf{X}^\top \hat{\beta}_\pi}) \), where \( \hat{\beta}_\pi = (0.2, -0.2, -0.2, 0.2)^\top \); **PS2** \( P(T = 1|\mathbf{X}) = \exp\{0.2(X_1^2 + X_2^2 + X_3X_4)/[1 + \exp\{0.2(X_1^2 + X_2^2 + X_3X_4)]\} \). Therefore, we consider four simulation scenarios (O1)+PS1), (O2)+PS1), (O1)+PS2), and (O2)+PS2). In the online Supporting information, more simulation with correlated variables are available. In each simulation, the sample size in the training dataset varies from 800, 1,000, to 1,500. To evaluate different methods, we generate a testing dataset with a sample size of \( 10^3 \). Each simulation is repeated 500 times.

**Figures 1 and 2** show the results for the rank correlation and the value function, respectively. For comparison, we also plot the oracle value of the value function (labeled as ‘Optim’) in Figure 2. We see that our proposed method outperforms Q-learning, and C-learning in all scenarios in terms of the rank correlation. Compared with the EARL and the \( \Delta \)-learning, our proposed method performs comparably in Setting O1), but much better in Setting O2).

These results show that our proposed method performs consistently well under our simulation scenarios. In terms of the inference results, Table 1 shows the coverage of the interval estimations. They appear to work very well.

### 6 REAL-DATA EXAMPLE

In this section, we compare different approaches to analyzing data from our motivating example, which is a study of the intervention effect of phone counseling versus usual care on the adherence to mammography screening guidelines (Champion et al., 2007). Participants include female subjects who do not follow the mammogram guidelines in the year before entering the study. One primary outcome of the study is whether the participant took mammography screening during a 21-month follow-up post-baseline period. Among 530 subjects, 259 participants receive the phone intervention and 271 the usual care. Each participant has 18 covariates available, including demographics, number of years having had at least one mammogram, recommendation by a doctor or nurse, and mammography screening and breast cancer belief and knowledge scales.

We randomly split the whole dataset into training and testing datasets with a sample size ratio of 4:1. On the
training dataset, our proposed method, the Q-learning, the \( \Delta \)-learning, the C-learning, and the EARL are implemented. Similar to our simulation study, we utilized the kernel regressions to estimate the propensity score \( \pi \) and the nuisance part \( S \). On the testing datasets, we calculate two metrics to compare. The entire procedure is repeated 1,000 times.

The first metric is the value function \( E[Y_{d(X)}] \) for a given decision rule \( d(X) \). The second metric is related to the ranks of \( \hat{\theta}(X) \). Because the true \( \theta(X) \) is unknown in real data, we propose the following averaged relative change (ARC). Given a \( \tau \in (0, 1) \), we define

\[
ARC(\hat{\theta}; \tau) = \frac{\sum_{i=1}^{n} I\{X : \hat{\theta}(X) \geq \hat{\theta}_{\tau}\} Y_{i}}{\sum_{i=1}^{n} I\{X : \hat{\theta}(X) \geq \hat{\theta}_{\tau}\}},
\]

where \( \hat{\theta}_{\tau} \) is the \( \tau \)-th quantile of \( \hat{\theta}(X) \) evaluated on a testing dataset. It is easy to see that the ARC can be written as the ratio of the average treatment effect, \( Y_{1} - Y_{-1} \), on \( \{X : \hat{\theta}(X) \geq \hat{\theta}_{\tau}\} \) and the average main effect, \( Y_{1} + Y_{-1} \), on \( \{X : \hat{\theta}(X) \geq \hat{\theta}_{\tau}\} \). Therefore, if \( \hat{\theta} \) possesses good ranking ability, we would expect \( ARC(\hat{\theta}; \tau) \) to be non-decreasing as a function of \( \tau \). To calculate \( ARC(\hat{\theta}; \tau) \), for our proposed approach, the \( \Delta \)-learning, and the Q-learning, \( \hat{\theta}(X) \) is directly obtained to calculate the ARC. For the EARL and the C-learning, we use \( X^{T}\hat{\beta}_{EARL} \) or the estimated index as a substitute for \( \hat{\theta}(X) \). The parameter \( \tau \) is chosen as 0.3, 0.4, ..., 0.7.

Table 2 and Figure 3 show that our proposed method achieves the highest value function and shows a monotone pattern for \( ARC(\hat{\theta}; \tau) \) compared with other methods. We also implemented our proposed method on the entire dataset to identify the important covariates. The identified important covariates and their \( p \)-values are listed in Table 3. In particular, we find that higher relative benefits from the tailored phone intervention are likely for patients who are currently working for pay, have more number of years having had at least one mammogram, have low household income, are non-Causian, have more than a high school degree, are contemplating screening, have a family history of breast cancer, have high perceived benefits of screening, and have low fear, high fatalism, high perceived susceptibility of breast cancer.
FIGURE 2  Value function comparison in simulated data. This figure appears in color in the electronic version of this paper, and any mention of color refers to that version.

TABLE 2  Comparison of the value in the real-data example summarized over 1000 cross-validated repeats

|               | Proposed | EARL   | QLearn  | ΔLearn | CRA   | CLearn |
|---------------|----------|--------|---------|--------|-------|--------|
| Value         | 0.4902   | 0.4874 | 0.4823  | 0.4526 | 0.4837| 0.4784 |
| Standard deviation | 0.0660   | 0.0683 | 0.0683  | 0.0541 | 0.0666| 0.0645 |

7  DISCUSSION

In this work, we have considered estimation of SIR contrasts for settings, where a relative difference may better represent treatment benefits. Because all SIR contrasts are monotonic transformations of each other, we posit a semiparametric single index model for a particular relative contrast. We then derive the efficient score and propose a two-step approach to obtain a semiparametrically efficient estimator for the index parameter.
Our work assumes positive conditional mean outcomes. For negative outcomes, we can shift the outcomes by a meaningful and interpretable positive constant which should be chosen based on domain knowledge. Unlike the outcome weighted learning where any shifts are theoretically allowed (Zhao et al., 2012), our proposed SIRs are not translation-invariant in general.

One important topic which is not covered in the work is the inference for the optimal value (Chakraborty et al., 2010, 2014; Goldberg et al., 2014; Laber et al., 2014; Luedtke & Van Der Laan, 2016; Zhang et al., 2012). Given the estimated \( \hat{\psi}(X^T \beta) \), the estimated decision rule for treatment recommendation can be defined as \( \text{sign}(\hat{\psi}(X^T \beta)) \). To obtain the value function inference under such estimated decision rule, existing methods can be directly applied.

In particular, the procedure proposed by Shi et al. (2020) should be applicable even though their results are based on \( \Delta(X) \). The conditions in Shi et al. (2020) are very general, and we can replace the estimator of \( \Delta(X) \) with \( \hat{\psi}(X^T \beta) \) to yield a valid inference procedure.

One possible future work is to extend our approach to accommodate high-dimensional covariates, for example, by adopting the de-correlated score proposed in Ning and Liu (2017). To circumvent the difficulty of enforcing the \( L_2 \)-norm constraint, one can replace it by assuming \( \theta(X) = \psi_0(1 + X^T \beta_0) \), where \( \psi_0 \) is a monotone function and \( \psi_0(0) = 0 \).

Another extension is to allow a non-monotone \( \psi_0 \). This extension is computationally straightforward by removing the constraint on \( \xi \). However, we may require the
convergence rates of the nuisance parameters to be faster than what is needed in this paper. It may also be hard to interpret $\beta_0$ when $\psi_0$ is non-monotonic, even if the inference is possible.

**DATA AVAILABILITY STATEMENT**

The data that support the findings in this paper are available from the corresponding author upon reasonable request.

**ORCID**

Muxuan Liang https://orcid.org/0000-0002-1351-7928
Menggang Yu https://orcid.org/0000-0002-7904-3155

**REFERENCES**

Bolthausen, E., Perkins, E. & Aad, v.d.V. (2002) *Lectures on Probability Theory and Statistics*. Springer Science & Business Media.

Carroll, R.J., Fan, J., Gijbels, I. & Wand, M.P. (1997) Generalized partially linear single-index models. *Journal of the American Statistical Association*, 92(438), 477–489.

Chakraborty, B., Laber, E.B. & Zhao, Y.Q. (2014) Inference about the expected performance of a data-driven dynamic treatment regime. *Clinical Trials*, 11(4), 408–417.

Chakraborty, B., Murphy, S. & Strecher, V. (2010) Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research*, 19(3), 317–343.

Champion, V., Maraj, M., Hui, S., Perkins, A.J., Tierney, W., Menon, U., et al. (2003) Comparison of tailored interventions to increase mammography screening in nonadherent older women. *Preventive Medicine*, 36(2), 150–158.

Champion, V., Skinner, C.S., Hui, S., Monahan, P., Julian, B., Daggy, J., et al. (2007) The effect of telephone versus print tailoring for mammography adherence. *Patient Education and Counseling*, 65(3), 416–423.

Champion, V.L., Rawl, S.M., Bourff, S.A., Champion, K.M., Smith, L.G., Buchanan, A.H., et al. (2016) Randomized trial of DVD, telephone, and usual care for increasing mammography adherence. *Journal of Health Psychology*, 21(6), 916–926.

Champion, V.L., Ray, D.W., Heilman, D.K. & Springer, J.K. (2000) A tailored intervention for mammography among low-income African–American women. *Journal of Psychosocial Oncology*, 18(4), 1–13.

Chen, S., Tian, L., Cai, T. & Yu, M. (2017) A general statistical framework for subgroup identification and comparative treatment scoring. *Biometrics*, 73(4), 1199–1209.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., et al. (2018) Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68.

Fan, C., Lu, W., Song, R. & Zhou, Y. (2017) Concordance-assisted learning for estimating optimal individualized treatment regimes. *Journal of the Royal Statistical Society, Series B*, 79(5), 1565–1582.

Goldberg, Y., Song, R., Zeng, D. & Kosorok, M.R. (2014) Comment on “dynamic treatment regimes: technical challenges and applications”. *Electronic Journal of Statistics*, 8(1), 1290–1300.

Imbens, G.W. & Rubin, D.B. (2015) *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Kosorok, M.R. & Moodie, E.E.M. (2015) *Adaptive treatment strategies in practice*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Laber, E.B., Linn, K.A. & Stefanski, L.A. (2014) Interactive model building for Q-learning. *Biometrika*, 101(4), 831–847.

Laber, E.B., Lizotte, D.J., Qian, M., Pelham, W.E. & Murphy, S.A. (2014) Dynamic treatment regimes: technical challenges and applications. *Electronic Journal of Statistics*, 8(1), 1225–1272.

Leitenstorfer, F. & Tutz, G. (2007) Generalized monotonic regression based on B-splines with an application to air pollution data. *Biostatistics*, 8(3), 654–673.

Liang, H., Liu, X., Li, R. & Tsai, C.L. (2010) Estimation and testing for partially linear single-index models. *Annals of Statistics*, 38(6), 3811–3836.

Liang, M. & Yu, M. (2022) A semiparametric approach to model effect modification. *Journal of the American Statistical Association*, 117(538), 752–764.

Lu, W., Zhang, H.H. & Zeng, D. (2013) Variable selection for optimal treatment decision. *Statistical Methods in Medical Research*, 22(5), 493–504.

Luedtke, A.R. & Van Der Laan, M.J. (2018) Optimal individualized treatments under limited resources. *Targeted learning in data science*. Springer, pp. 419–435.

Luedtke, A.R. & Van Der Laan, M.J. (2016) Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Annals of Statistics*, 44(2), 713–742.

Ma, S. & He, X. (2016) Inference for single-index quantile regression models with profile optimization. *Annals of Statistics*, 44(3), 1234–1268.

Murphy, S.A. (2003) Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B*, 65(2), 331–355.

Ning, Y. & Liu, H. (2017) A general theory of hypothesis tests and confidence regions for sparse high-dimensional models. *Annals of Statistics*, 45(1), 158–195.

Qian, M. & Murphy, S.A. (2011) Performance guarantees for individualized treatment rules. *Annals of Statistics*, 39(2), 1180–1210.

Qiu, H., Carone, M. & Luedtke, A. (2022) Individualized treatment rules under stochastic treatment cost constraints. arXiv:2201.06669.

Rosenbaum, P.R. & Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Rubin, D.B. (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.

Rubin, D.B. (2005) Causal inference using potential outcomes: design, modeling, decisions. *Journal of the American Statistical Association*, 100(469), 322–331.

Shi, C., Fan, A., Song, R. & Lu, W. (2018) High-dimensional a-learning for optimal dynamic treatment regimes. *Annals of Statistics*, 46(3), 925–957.

Shi, C., Lu, W. & Song, R. (2020) Breaking the curse of nonregularity with subagging: inference of the mean outcome under optimal treatment regimes. *Journal of Machine Learning Research*, 21(176), 1–67.

Shi, C., Song, R. & Lu, W. (2016) Robust learning for optimal treatment decision with np-dimensionality. *Electronic Journal of Statistics*, 10, 2894–2921.

**ORCID**

Muxuan Liang https://orcid.org/0000-0002-1351-7928
Menggang Yu https://orcid.org/0000-0002-7904-3155
Song, R., Luo, S., Zeng, D., Zhang, H.H., Lu, W. & Li, Z. (2017) Semiparametric single-index model for estimating optimal individualized treatment strategy. *Electronic Journal of Statistics*, 11(1), 364–384.

Stone, C.J. (1982) Optimal global rates of convergence for nonparametric regression. *Annals of Statistics*, 10, 1040–1053.

Tian, L., Alizadeh, A.A., Gentles, A.J. & Tibshirani, R. (2014) A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*, 109(508), 1517–1532.

Watkins, C.J. & Dayan, P. (1992) Q-learning. *Machine Learning*, 8(3–4), 279–292.

Xu, Y., Yu, M., Zhao, Y.Q., Li, Q., Wang, S. & Shao, J. (2015) Regularized outcome weighted subgroup identification for differential treatment effects. *Biometrics*, 71(3), 645–653.

Yadlowsky, S., Pellegrini, F., Lionetto, F., Braune, S. & Tian, L. (2021) Estimation and validation of ratio-based conditional average treatment effects using observational data. *Journal of the American Statistical Association*, 116(533), 335–352.

Zhang, B., Tsiatis, A.A., Davidian, M., Zhang, M. & Laber, E. (2012) Estimating optimal treatment regimes from a classification perspective. *Stat*, 1(1), 103–114.

Zhang, B., Tsiatis, A.A., Laber, E.B. & Davidian, M. (2012) A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4), 1010–1018.

Zhao, Y., Zeng, D., Rush, A.J. & Kosorok, M.R. (2012) Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499), 1106–1118.

Zhao, Y.Q., Laber, E.B., Ning, Y., Saha, S. & Sands, B.E. (2019) Efficient augmentation and relaxation learning for individualized treatment rules using observational data. *Journal of Machine Learning Research*, 20(1), 1821–1843.

**SUPPORTING INFORMATION**

Web Appendices, tables, and figures, referenced in Sections 3–5, are available with this paper at the Biometrics website on Wiley Online Library. An R package implementing the proposed method in this paper is posted online with this paper and can be downloaded at [https://github.com/muxuanliang/RSICF.git](https://github.com/muxuanliang/RSICF.git).

Web Figure 1. Comparison in simulated data.

Web Table 1. Coverages of the 95% confidence intervals from the proposed approach.

Data S1

**How to cite this article**: Liang, M. & Yu, M. (2023) Relative contrast estimation and inference for treatment recommendation. *Biometrics*, 79, 2920–2932. [https://doi.org/10.1111/biom.13826](https://doi.org/10.1111/biom.13826)