JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS
2021, VOL. 00, NO. 0, 1–10
https://doi.org/10.1080/10618600.2021.1923521



# A Single-Index Model With a Surface-Link for Optimizing Individualized Dose Rules

Hyung Park[a], Eva Petkova[a], Thaddeus Tarpey[a], and R. Todd Ogden[b]

[a]Division of Biostatistics, Department of Population Health, New York University, New York, NY; [b]Department of Biostatistics, Columbia University, New York, NY

**ABSTRACT**

This article focuses on the problem of modeling and estimating interaction effects between covariates and a continuous treatment variable on an outcome, using a single-index regression. The primary motivation is to estimate an optimal individualized dose rule and individualized treatment effects. To model possibly nonlinear interaction effects between the patients' covariates and a continuous treatment variable, we employ a two-dimensional penalized spline regression on an index-treatment domain, where the index is defined as a linear projection of the covariates. The method is illustrated using two applications as well as simulation experiments. A unique contribution of this work is in the parsimonious (single-index) parameterization specifically defined for the interaction effect term, that can be used to assess the treatment benefit. Supplemental materials for this article are available online.

**ARTICLE HISTORY**
Received March 2020
Revised March 2021

**KEYWORDS**
Heterogeneous dose effects; Individualized dose rules; Single-index models; Tensor product P-splines

## 1. Introduction

In precision medicine, a primary goal is to characterize the individuals' heterogeneity in treatment responses so that individual-specific treatment decisions can be made (Murphy 2003; Robins 2004). Most work on developing methods for individualized treatment decisions has focused on a finite number of treatment options. The focus on this article is to develop individualized treatment decision methodology in the realm of a continuous treatment. Specifically, we consider a semiparametric regression approach for developing optimal individualized dosing rules based on baseline patient characteristics. Often in clinical practice, the maximum dose that a patient can tolerate is the most effective one, however, there are situations where this is not the case. In the example section, we present a study of warfarin (an anticoagulant), where too high doses lead to severe bleeding and thus the highest dose is not the optimal dose. In finding the optimal dose, there is an essential nonmonotone and nonlinear relationship that needs to be accounted for. A similar case is with insulin for controlling blood glucose levels.

To establish notation, let $X = (X_1, \ldots, X_p)^\top \in \mathcal{X}$ be the set of baseline covariates, $Y \in \mathbb{R}$ be the outcome variable, and $A \in \mathcal{A}$ denote the dose. Let $Y^*(a)$ be the potential outcome when a dose level $a \in \mathcal{A}$ is given. Throughout, we assume: (i) consistency, that is, $Y = \int_{\mathcal{A}} \delta(A = a) Y^*(a) \mathrm{d}a$, where $\delta(\cdot)$ is the Dirac delta function; (ii) no unmeasured confoundedness, that is, $\{Y^*(a), a \in \mathcal{A}\}$ is conditionally independent of $A$ given $X$; (iii) positivity, that is, $p(A = a|X = x) \geq c$, for all $a \in \mathcal{A}, x \in \mathcal{X}$, for some $c > 0$ (where $p(a|x)$ is the conditional density of $A = a$ given $X = x$), which are standard assumptions adopted in the causal inference literature (Gill and Robins 2001). Without loss of generality, we assume that a larger value of the outcome $Y$ is better. The goal is then to find an optimal individualized dose rule $f : \mathcal{X} \mapsto \mathcal{A}$ such that for a patient with covariate $X$, the dose assignment $A = f(X)$ maximizes the expected response, the so-called value function, $\mathcal{V}(f) = \mathbb{E}[Y^*(f(X))]$, that is,

$$\mathcal{V}(f) = \mathbb{E}[\mathbb{E}[Y|A = f(X), X]] \tag{1}$$

which holds and can be empirically approximated under the above three assumptions. In settings in which the treatment can be administered at continuous doses (i.e., when $\mathcal{A}$ is an interval), Chen, Zeng, and Kosorok (2016) proposed to optimize the individualized dosing rule $f$ by maximizing a local approximation of the value function (1), optimized under the framework of outcome weighted learning (Zhao et al. 2012). Laber and Zhao (2015) proposed a tree-based decision rule for treatment assignment with minimal impurity dividing patients into subgroups with different discretized doses. Kallus and Zhou (2018) developed an inverse propensity weighted estimator of Equation (1) for continuous treatments with the doubly robust property (Dudík et al. 2014), and recently, kernel-assisted learning with linear dimension reduction (Zhou, Zhu, and Zeng 2020; Zhu et al. 2020) for direct optimization of Equation (1) have been developed. However, implementation of these approaches for general exponential family distributions is not straightforward and has not been accomplished. In this article, we consider a regression-based approach to optimizing $f$ that uses a semiparametric regression model for $\mathbb{E}[Y|A, X]$. There is also extensive literature on multi-armed bandit (e.g., Lattimore and Szepesvari

**CONTACT** Hyung Park syhyunpark@gmail.com Division of Biostatistics, Department of Population Health, New York University, New York, NY 10012-1126.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JCGS.

2019) problems in the context of reinforcement learning (e.g., Kaelbling, Littman, and Moore 1996), incorporating context (i.e., feature $X \in \mathcal{X}$) (see, e.g., Lu, Pal, and Pal 2010; Perchet and Rigollet 2013; Slivkins 2014; Jun et al. 2017; Li, Lu, and Zhou 2017; Kveton et al. 2020; Chen, Lu, and Song 2021) for making a sequential decision that minimizes the notion of cumulative regret, with relatively fewer works on contextual bandits with continuous actions (see, e.g., Kleinberg, Slivkins, and Upfal 2019; Krishnamurthy et al. 2020; Majzoubi et al. 2020). However, these works are focused on optimizing online performance (addressing the exploration issue) and considerably different from personalized dose-finding focused on a single stage with feature $X$. Kennedy et al. (2017) considered a method for estimating the average dose effect allowing for flexible doubly robust covariate adjustment, but the method is not intended for optimal dose finding for individual patients. For multi-stage personalized dose-finding, Rich, Moodie, and Stephens (2014) proposed adaptive strategies, and more recently Schulz and Moodie (2021) developed a doubly robust estimation approach based on a linear model. All these approaches are limited by the stringent linear model assumptions for the heterogeneous dose effects.

While the methods of directly optimizing the value function (1), including the outcome weighted learning of Chen, Zeng, and Kosorok (2016) and the tree-based method of Laber and Zhao (2015), are highly appealing, the proposed semiparametric regression modeling has the advantage of being easy to implement and readily generalizable to an exponential family response.

It is straightforward to see that, given $X$, the optimal dose $f_{\text{opt}}(X)$ (i.e., that which maximizes the value function (1)) is

$$f_{\text{opt}}(X) \;=\; \underset{a\in\mathcal{A}}{\text{argmax}}\, m(a,X), \tag{2}$$

where $m(a,X) = \mathbb{E}[Y|A = a, X]$. If we estimate $m(a,X)$ with $\hat{m}(a,X)$, then the optimal rule $f_{\text{opt}}$ in Equation (2) can be approximated as

$$\hat{f}(X) \;=\; \underset{a\in\mathcal{A}}{\text{argmax}}\, \hat{m}(a,X). \tag{3}$$

Methodologies for optimizing individualized treatment rules $f$ in the precision medicine literature are developed predominantly for the cases in which the treatment variable $A$ is binary or discrete-valued. Regression-based methodologies first estimate the treatment $a$-specific mean response functions $m(a,X)$ and then obtain a treatment decision rule, that is, the left-hand side of (3) (e.g., see Qian and Murphy 2011; Zhang et al. 2012; Gunter, Zhu, and Murphy 2011; Lu, Zhang, and Zeng 2013; Park et al. 2020) given $X$. In particular, Qian and Murphy (2011) showed that the optimal individualized treatment rule (2) depends only on the interaction between treatment $A$ and covariates $X$, and not on the main effects of $X$ in the mean models $m(a,X)$. For regression-based methodologies, a successful estimation of the function $f_{\text{opt}}$ in Equation (2) boils down to efficiently estimating the $A$-by-$X$ interaction effects on the treatment response. In this article, we consider a semiparametric model that is useful for estimating such interactions in the case where $A$ is a continuous dose variable.

## 2. Models

Our goal is to provide an interpretable and flexible approach to modeling the $A$-by-$X$ interaction effects on $Y$. To achieve this goal, we consider the following additive single-index model:

$$\mathbb{E}[Y|X,A] \;=\; \mu(X) \,+\, g(\beta^\top X, A) \tag{4}$$

where $\mu(X)$ represents an unspecified main effect of $X$, and $g(\beta^\top X, A)$ models the $A$-by-$X$ interaction effects. Here, $g(\cdot,\cdot)$ is an unspecified smooth two-dimensional *surface* link function of the variable $A$ and a single index $\beta^\top X$. We shall call model (4) a *single-index model with a surface-link* (SIMSL). We restrict $\beta \in \Theta := \{\beta = (\beta_1,\ldots,\beta_p)^\top \in \mathbb{R}^p : \|\beta\|_2 = 1, \beta_1 > 0\}$, as $\beta$ in Equation (4) is only identifiable up to a scale constant without further constraint, due to the unspecified nature of $g$.

Without loss of generality, we assume $\mathbb{E}[\mu(X)] = 0$ and $\mathbb{E}[g(\beta^\top X, A)] = 0$ (where the expectation is with respect to $X$ and $A$), that is, each of the additive components in model (4) has mean 0, and that these components have finite variances, as typically assumed in generalized additive models (GAM; Hastie and Tibshirani 1999). That is, let $\mathcal{H}_1$ and $\mathcal{H}_2^{(\beta)}$ (for a fixed $\beta \in \Theta$) denote the $L^2$ spaces of measurable functions $\mu(X)$ on $X$ and measurable functions $g(\beta^\top X, A)$ on $(\beta^\top X, A)$, respectively, and we assume $\mu \in \mathcal{H}_1$ and $g \in \mathcal{H}_2^{(\beta)}$.

Due to the unspecified nature of $\mu$ and $g$ (and that $X$ is involved in both $\mu$ and $g$), model (4) is not identifiable without further constraints. We will constrain the smooth function $g \in \mathcal{H}_2^{(\beta)}$ to satisfy

$$\mathbb{E}[g(\beta^\top X, A)|X] = 0, \quad \text{a.s. } (X) \quad X \in \mathcal{X},\ \beta \in \mathbb{R}^p, \tag{5}$$

which acts as an identifiability condition of model (4).

Applying the constraint (5) to the function $g$ in Equation (4) essentially reparameterizes the model (4), by replacing $g(\beta^\top X, A)$ with $g_0(\beta^\top X, A) = g(\beta^\top X, A) - \mathbb{E}[g(\beta^\top X, A)|X]$, and $\mu(X)$ with $\mu_0(X) = \mu(X) + \mathbb{E}[g(\beta^\top X, A)|X]$. This yields an equivalent model of Equation (4), $\mathbb{E}[Y|X,A] = \mu_0(X) + g_0(\beta^\top X, A)$, where the term $g_0(\beta^\top X, A)$ satisfies the identifiability condition (5). Since any arbitrary $(\mu, g)$ in Equation (4) can be rearranged to give such reparameterized components $(\mu_0, g_0)$, without loss of generality, we will represent $(\mu_0, g_0)$ as $(\mu, g)$ subject to Equation (5).

Under the SIMSL (4) (subject to constraint (5)), the optimal individualized dose rule, $f_{\text{opt}}$, is specified as: $f_{\text{opt}}(X) = \underset{a\in\mathcal{A}}{\text{argmax}}\, g(\beta^\top X, a)$, which does not involve the component $\mu$. Therefore, in terms of estimating $f_{\text{opt}}$ in Equation (2), our modeling focus is on estimating $(g, \beta)$ in Equation (4).

Using the constrained least-square framework, the right-hand side of Equation (4), subject to constraint (5), can be optimized by solving:

$$\begin{aligned}(\mu^*, g^*, \beta^*) = \quad & \underset{\mu\in\mathcal{H}_1, g\in\mathcal{H}_2^{(\beta)}, \beta\in\Theta}{\text{argmin}} && \mathbb{E}\big[(Y - \mu(X) - g(\beta^\top X, A))^2\big] \\ & \text{subject to} && \mathbb{E}\Big[g(\beta^\top X, A)|X\Big] = 0.\end{aligned} \tag{6}$$

Constraint (5) ensures that $\mathbb{E}\big[\mu(X)g(\beta^\top X, A)\big] = \mathbb{E}\big[\mu(X)\,\mathbb{E}[g(\beta^\top X, A)|X]\big] = 0$ (in which we apply the iterated expectation rule to condition on $X$), which implies the orthogonality,

$$\mu(X) \quad \perp \quad g(\beta^\top X, A), \tag{7}$$

in $L^2$. The orthogonality (7) implies that the optimization for $\mu^*$ and that for $(g^*, \beta^*)$ on the left-hand side of Equation (6) can be performed separately, without iterating between the two optimizations. Specifically, we can solve for the $X$ main effect component

$$\mu^* = \underset{\mu \in \mathcal{H}_1}{\operatorname{argmin}} \ \mathbb{E}\big[(Y - \mu(X))^2\big], \tag{8}$$

and separately solve for the $A$-by-$X$ interaction effect component

$$\begin{aligned}(g^*, \beta^*) = \underset{g \in \mathcal{H}_2^{(\beta)}, \beta \in \Theta}{\operatorname{argmin}} \quad & \mathbb{E}\big[(Y - g(\beta^\top X, A))^2\big] \\ \text{subject to} \quad & \mathbb{E}\big[g(\beta^\top X, A)|X\big] = 0,\end{aligned} \tag{9}$$

without alternating between the two optimization procedures. We can fit $f_{\text{opt}}$ by: $\hat{f}(X) = \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, \hat{g}^*(\hat{\beta}^{*\top} X, a)$, where $(\hat{g}^*, \hat{\beta}^*)$ denotes an estimate of $(g^*, \beta^*)$ in Equation (9). This approach using optimization (9) is appealing, because, due to orthogonality (7), misspecification of the functional form for $\mu$ in Equation (6) does not affect specification of $(g^*, \beta^*)$ on the left-hand side of Equation (9). If primary interest is in estimating $f_{\text{opt}}$ in Equation (2), then using $\hat{f}$ based on optimization (9) circumvents the need to estimate the term $\mu$ in (4), obviating the need to specify its form and thus avoiding the issue of model misspecification on the $X$ main effect. The equivalence between $(g^*, \beta^*)$ on the left-hand side of Equation (9) and $(g, \beta)$ in Equation (4) is given in Proposition 1 in Section 4 (in the context where $Y$ follows an exponential family response).

Since our primary focus is on estimating $f_{\text{opt}}$, we focus on solving Equation (9). However, we also note that modeling the term $\mu$ in Equation (4) can generally improve the performance of the estimator of $(g, \beta)$, as discussed in supplementary materials Section C.1. For each fixed $\beta$, the term $g(\beta^\top X, A)$ depends the covariates $X \in \mathcal{X}$ only through the one-dimensional projection $\beta^\top X$. Therefore, for each fixed $\beta$, the distribution of $g(\beta^\top X, A)|X$ is equal to that for $g(\beta^\top X, A)|\beta^\top X$, which implies $\mathbb{E}\big[g(\beta^\top X, A)|X\big] = \mathbb{E}\big[g(\beta^\top X, A)|\beta^\top X\big]$, for each fixed $\beta$. Then, for each fixed $\beta \in \Theta$, the following constraint on $g \in \mathcal{H}_2^{(\beta)}$,

$$\mathbb{E}\Big[g(\beta^\top X, A)|\beta^\top X\Big] = 0, \quad X \in \mathcal{X}, \tag{10}$$

is a sufficient condition for the orthogonality constraint (5). Thus, for each fixed $\beta$, the constraint (5) can be simplified to (10). The following iterative procedure will be used to solve Equation (9):

1. For fixed $\beta$, optimize the smooth $g(\cdot, \cdot)$ by solving
$$\underset{g \in \mathcal{H}_2^{(\beta)}}{\operatorname{argmin}}\, \mathbb{E}\big[(Y - g(\beta^\top X, A))^2\big], \tag{11}$$
subject to the constraint (10).
2. For fixed $g$, optimize the coefficient $\beta \in \Theta$ by minimizing the squared error criterion of Equation (11).
3. Iterate steps 1 and 2 until convergence with $\beta \in \Theta$.

The data version of optimizing $(g, \beta)$ can be derived as an empirical counterpart of the iterative procedure given above. Details on implementing this algorithm are given below.

## 3. Estimation

### 3.1. Representation of Link Surface

Suppose we have observed data $(Y_i, A_i, X_i)$ $(i = 1, \ldots, n)$. For each candidate vector $\beta \in \Theta$, let

$$u_i = u_i^{(\beta)} = \beta^\top X_i \quad (i = 1, \ldots, n),$$

where (on the left-hand side), for the notational simplicity, we suppress the dependence of the linear predictor $u^{(\beta)} \in \mathbb{R}$ on the candidate vector $\beta$.

Eilers and Marx (2003) used tensor products of $B$-splines (de Boor 2001) to represent two-dimensional surfaces, which they termed tensor product $P$-splines, with separate difference penalties applied to the coefficients of the $B$-splines along the covariate axes. Although alternative nonparametric methods could also be used to estimate the smooth function $g \in \mathcal{H}_2^{(\beta)}$ given each coefficient vector $\beta$ in model (4), in this article we focus on one smoother, the tensor-product $P$-splines, for the ease of presentation.

Specifically, for each $u = \beta^\top X$, to represent the two-dimensional function $g(u, A)$ in (11), we consider the tensor product of the two sets of univariate cubic $B$-spline basis functions, say $B$ and $\check{B}$, with $N$ (and $\check{N}$) $B$-spline knots for the basis functions that are placed along the $u$ (and $A$) axis. The number of knots $N$ (and $\check{N}$) is chosen to be large, that is, to allow the surface much flexibility. Associated with the basis representation defined by the marginal basis function $B$ (resp., $\check{B}$) is an $N \times N$ (resp., $\check{N} \times \check{N}$) roughness penalty matrix, which we denote by $\mathbb{P}$ (and $\check{\mathbb{P}}$). The penalty matrix $\mathbb{P}$ (and $\check{\mathbb{P}}$) can be easily constructed, for example, based on a second-order difference matrix (see, e.g., Eilers and Marx 2003).

For each fixed $u_i = \beta^\top X_i$ $(i = 1, \ldots, n)$, let us write the $n \times N$ (and $n \times \check{N}$) $B$-spline evaluation matrix $\boldsymbol{B}$ (and $\check{\boldsymbol{B}}$), in which its $i$th row is $\boldsymbol{B}_i = B(u_i)^\top$ (and $\check{\boldsymbol{B}}_i = \check{B}(A_i)^\top$). Given a knot grid, a flexible surface can be approximated (Marx 2015) at $n$ points $(u_i, A_i)$ $(i = 1, \ldots, n)$:

$$g(u_i, A_i) = \sum_{r=1}^{N} \sum_{s=1}^{\check{N}} B_r(u_i) \check{B}_s(A_i) \gamma_{rs} = (\boldsymbol{B}_i \otimes \check{\boldsymbol{B}}_i)\boldsymbol{\theta} \quad (i = 1, \ldots, n), \tag{12}$$

where the vector $\boldsymbol{\theta} = \big(\gamma_{11}, \ldots, \gamma_{1\check{N}}; \ldots; \gamma_{N1}, \ldots, \gamma_{N\check{N}}\big)^\top \in \mathbb{R}^{N\check{N}}$ corresponds to an unknown (vectorized) coefficient vector of the tensor product representation of $g$, and $\otimes$ represents the usual Kronecker product. Equation (12) can be compactly written as:

$$\operatorname{vec}\big\{g(u_i, A_i)\big\} = g(u_{n\times 1}, A_{n\times 1}) = \boldsymbol{D}\boldsymbol{\theta}, \tag{13}$$

where

$$\boldsymbol{D} = \boldsymbol{B} \square \check{\boldsymbol{B}} = \left(\boldsymbol{B} \otimes \mathbf{1}_{\check{N}}^\top\right) \odot \left(\mathbf{1}_N^\top \otimes \check{\boldsymbol{B}}\right), \tag{14}$$

in which the symbol $\odot$ denotes element-wise multiplication of matrices. In Wood (2017), the symbol $\square$ in Equation (14) is called the row-wise Kronecker product, which results in an $n \times N\check{N}$ tensor product design matrix $\boldsymbol{D}$ from the two marginal design matrices $\boldsymbol{B}$ and $\check{\boldsymbol{B}}$.

Similarly, the roughness penalty matrices associated with the tensor product representation (12) can be constructed from

the roughness penalty matrices $\mathbb{P}$ and $\check{\mathbb{P}}$ associated with the univariate (marginal) basis matrices $\boldsymbol{B}$ and $\check{\boldsymbol{B}}$, and are given by $\boldsymbol{P} = \mathbb{P} \otimes \boldsymbol{I}_{\check{N}}$ and $\check{\boldsymbol{P}} = \boldsymbol{I}_N \otimes \check{\mathbb{P}}$, for the axis directions $u$ and $A$, respectively. Here, $\boldsymbol{I}$ denotes the identity matrix, and both $\boldsymbol{P}$ and $\check{\boldsymbol{P}}$ are square matrices with dimension $N\check{N}$.

We now need to impose the constraint (10) on the two-dimensional smooth function $g$ under the tensor product representation (13). For each fixed $\beta$, the constraint (10) on $g$ amounts to excluding the main effect of $u = \beta^\top X$ from the function $g$. We deal with this by a reparameterization of the representation (13) for $g$. Consider the following sum-to-zero (over the $n$ observed values) constraint for the marginal function of $A$:

$$\mathbf{1}^\top \check{\boldsymbol{B}} \check{\boldsymbol{\gamma}} = 0, \tag{15}$$

for any arbitrary $\check{\boldsymbol{\gamma}} \in \mathbb{R}^{\check{N}}$, where $\mathbf{1}$ is a length $n$ vector of 1's. With constraint (15), the linear smoother associated with the basis matrix $\check{\boldsymbol{B}}$ cannot reproduce constant functions (Hastie and Tibshirani 1999). That is, the linear constraint (15) removes the span of constant functions from the span of the marginal basis matrix $\check{\boldsymbol{B}}$ associated with $A$. Constraint (15) results in a tensor product basis matrix, $\boldsymbol{D} = \boldsymbol{B} \square \check{\boldsymbol{B}}$ in (13), that will not include the main effect of $u$ that results from the product of the marginal basis matrix $\boldsymbol{B}$ with the constant function in the span of the other marginal basis matrix $\check{\boldsymbol{B}}$. Therefore, the resultant fit, under representation (13) (subject to Equation (15)) of the smooth function $g$, excludes the main effect of $u$. See Section 5.6 of Wood (2017) for some more details.

We impose the linear constraint (15) on the matrix $\check{\boldsymbol{B}}$, and consequently, the resulting basis matrix $\boldsymbol{D}$ of representation of $g$ in Equation (13) becomes independent of the basis associated with the main effect of $u$. Imposition of such a linear constraint (15) on a basis matrix is routine. The key is to find an (orthogonal) basis for the null space of the constraint (15), and then absorb the constraint into the basis construction (14). To be specific, we can create a $\check{N} \times (\check{N}-1)$ matrix, which we denote as $\boldsymbol{Z}$, such that, given any arbitrary coefficient vector $\check{\boldsymbol{\gamma}}_0 \in \mathbb{R}^{\check{N}-1}$, if we set $\check{\boldsymbol{\gamma}} = \boldsymbol{Z}\check{\boldsymbol{\gamma}}_0$, then we have $\mathbf{1}^\top \check{\boldsymbol{B}} \check{\boldsymbol{\gamma}} = 0$, automatically satisfying the constraint (15). Such a matrix $\boldsymbol{Z}$ is constructed using a QR decomposition of $\check{\boldsymbol{B}}^\top \mathbf{1}$. Then we can reparameterize the marginal function of $A$ by setting its model matrix to $\check{\boldsymbol{B}} \leftarrow \check{\boldsymbol{B}}\boldsymbol{Z}$ (and its penalty matrix to $\check{\mathbb{P}} \leftarrow \boldsymbol{Z}^\top \check{\mathbb{P}} \boldsymbol{Z}$). From this point forward, for notational simplicity, we redefine the matrix $\check{\boldsymbol{B}}$ (and $\check{\mathbb{P}}$) to be this reparameterized, constrained marginal basis matrix (and the reparameterized constrained penalty matrix).

This sum-to-zero reparameterization of the marginal basis matrix $\check{\boldsymbol{B}}$ of $A$ to satisfy Equation (15) creates a term $\text{vec}\{g(u_i, A_i)\} \in \mathbb{R}^n$ in Equation (13) that specifies such a pure $A$-by-$X$ interaction (plus the $A$ main effect) component, that is also orthogonal to the $X$ main effect. In Wood (2006), this reparameterization approach is used to create an analysis of variance (ANOVA) decomposition of a smooth function of several variables. In this article, we use this same reparameterization to orthogonalize the interaction effect component $g(\beta^\top X, A)$ from the main effect, and to allow an unspecified/misspecified main effect for $X$ in the estimation of the SIMSL (4). Provided that the orthogonality constraint (i.e., Equation (15)) issue is addressed, the interaction effect term $g(\beta^\top X, A)$ of model (4), for each fixed $\beta$, can be represented using penalized regression splines and estimated based on penalized least squares, which we describe next.

### 3.2. Estimation Algorithm

We define the criterion function for estimating $(g, \beta)$ in the SIMSL (4)

$$\begin{aligned} Q(\boldsymbol{\theta}, \beta) &= \|Y_{n\times 1} - g(\boldsymbol{X}\beta, A_{n\times 1})\|^2 + \lambda \|\boldsymbol{P}\boldsymbol{\theta}\|^2 + \check{\lambda}\|\check{\boldsymbol{P}}\boldsymbol{\theta}\|^2 \\ &= \|Y_{n\times 1} - \boldsymbol{D}\boldsymbol{\theta}\|^2 + \lambda \|\boldsymbol{P}\boldsymbol{\theta}\|^2 + \check{\lambda}\|\check{\boldsymbol{P}}\boldsymbol{\theta}\|^2 \end{aligned} \tag{16}$$

subject to the constraint that the function $g(\cdot, \cdot)$ empirically satisfies Equations (5). In Equation (16), $\boldsymbol{X}$ is an $n \times p$ matrix whose $i$th row is $X_i^\top$. Since both $\boldsymbol{\theta}$ and $\beta$ are unknown in Equation (16), estimation of $\boldsymbol{\theta}$ and $\beta$ is conducted iteratively. We describe below the estimation procedure.

1. For a fixed estimate of $\beta$ (that defines the linear predictor $u$), minimize the following criterion function over $\boldsymbol{\theta} \in \mathbb{R}^{N\check{N}}$,

$$\|Y_{n\times 1} - \boldsymbol{D}\boldsymbol{\theta}\|^2 + \lambda \|\boldsymbol{P}\boldsymbol{\theta}\|^2 + \check{\lambda}\|\check{\boldsymbol{P}}\boldsymbol{\theta}\|^2, \tag{17}$$

   where $\boldsymbol{D}$ is given by Equation (14). Given tuning parameters $(\lambda, \check{\lambda})$, the minimizer $\hat{\boldsymbol{\theta}}$ of (17) is:

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{D}^\top \boldsymbol{D} + \lambda \boldsymbol{P}^\top \boldsymbol{P} + \check{\lambda}\check{\boldsymbol{P}}^\top \check{\boldsymbol{P}}\right)^{-1} \boldsymbol{D}^\top Y_{n\times 1}.$$

2. For a fixed estimate of the surface $g$ (i.e., given $\boldsymbol{\theta}$), perform a first-order Taylor approximation of $g(\boldsymbol{X}\beta, A_{n\times 1})$ in Equation (16) with respect to $\beta$, around the current estimate, say, $\tilde{\beta} \in \Theta$,

$$g(\boldsymbol{X}\beta, A_{n\times 1}) \approx g(\boldsymbol{X}\tilde{\beta}, A_{n\times 1}) + \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}(\beta - \tilde{\beta}), \tag{18}$$

   where $\dot{g}_{\partial_1}(u, a)$ denotes the partial first derivative of $g(u, a)$ with respect to $u$, that is, $\frac{\partial g(u,a)}{\partial u}$. Using (18), we approximate the quadratic loss in Equation (16) by (as a function of $\beta$ given $\boldsymbol{\theta}$)

$$\begin{aligned} &\left\| Y_{n\times 1} - g(\boldsymbol{X}\tilde{\beta}, A_{n\times 1}) - \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}(\beta - \tilde{\beta}) \right\|^2 \\ = &\left\| \begin{array}{l} Y_{n\times 1} - g(\boldsymbol{X}\tilde{\beta}, A_{n\times 1}) + \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}\tilde{\beta} \\ \quad - \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}\beta \end{array} \right\|^2 \\ = &\left\| Y^*_{n\times 1} - \boldsymbol{X}^*\beta \right\|^2, \end{aligned} \tag{19}$$

   where $Y^*_{n\times 1} = Y_{n\times 1} - g(\boldsymbol{X}\tilde{\beta}, A_{n\times 1}) + \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}\tilde{\beta}$, and $\boldsymbol{X}^* = \text{diag}\left\{\dot{g}_{\partial_1}(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})\right\}\boldsymbol{X}$. The minimizer $\hat{\beta}$ of Equation (19) is

$$\hat{\beta} = \left(\boldsymbol{X}^{*\top}\boldsymbol{X}^*\right)^{-1} \boldsymbol{X}^{*\top} Y^*_{n\times 1}. \tag{20}$$

   Then we scale $\hat{\beta}$ to have unit $L^2$ norm, that is, $\hat{\beta}/\|\hat{\beta}\|$, and enforce a positive first element to restrict the estimate of $\beta$ to be in $\Theta$.

These two steps can be iterated until convergence to obtain an estimate of $(g^*, \beta^*)$ in Equation (9), which we denote as $(\hat{g}^*, \hat{\beta}^*)$. For Step 1, the tuning parameters $(\lambda, \check{\lambda})$ can be automatically selected, for example, by the generalized cross-validation (GCV) or the restricted maximum likelihood (REML) methods. In this article, we use REML for the simulation examples and the applications.

Finally, for model hierarchy, it is common practice to include all lower order effects of variables if there are higher-order interaction terms including that set of variables. Once convergence of the estimate $\hat{\beta}$ is reached in the above algorithm and the single-index $\beta^\top X$ in the term $g(\beta^\top X, A)$ of model (4) is estimated, we recommend fitting one final (unconstrained) smooth function $g$ of $A$ and $\hat{\beta}^\top X$, without enforcing the constraint (15) on $g$. Given the final estimate of $\beta$, the unconstrained final surface-link $g(\cdot, \cdot)$ retains the main effect of $\beta^\top X$ and preserves model hierarchy.

## 4. Generalized Single-Index Models for Optimizing Dose Rules

The proposed approach to optimizing the heterogeneous dose effect (i.e., the $X$-by-$A$ interaction effect) term of model (4) can be extended to a more general setting in which the response $Y$ follows an exponential family distribution. Given $(X, A)$, we assume an additive single-index model (4)

$$m(X, A) = \mathbb{E}[Y|X, A] = \mu_0(X) + g_0(\beta_0^\top X, A), \tag{21}$$

(in which subscript (0) is used to indicate the "true" value), and the variance of $Y$ is determined based on the exponential family density of the form:

$$\exp\left\{\left[Yh(m(X, A)) - b(h(m(X, A)))\right]/a(\phi) + c(Y, \phi)\right\}, \tag{22}$$

where $h$ is the *canonical* link function associated with the assumed distribution of $Y$, and the functions $a$, $b$ and $c$ are distribution-specific known functions. For model identifiability, we assume $\beta_0 \in \Theta$, and $g_0 \in \mathcal{H}_2^{(\beta_0)}$ to satisfy $\mathbb{E}[g_0(\beta_0^\top X, A)|X] = 0$. The dispersion parameter $\phi > 0$ in Equation (22) takes on a fixed, known value in some families (e.g., $\phi = 1$, in Bernoulli and Poisson), while in other families, it is an unknown parameter (e.g., in Gaussian).

We estimate $(g_0, \beta_0)$ in model (21) by using the following working mean model:

$$m(X, A) = h^{-1}(g(\beta^\top X, A)) \qquad (g \in \mathcal{H}_2^{(\beta)};\ \beta \in \Theta), \tag{23}$$

subject to the constraint $\mathbb{E}[g(\beta^\top X, A)|X] = 0$ on $g$, and the distribution of $Y$ is described by the exponential family chosen in (22). We optimize the population logarithmic version of Equation (22) over the unknowns $(g, \beta)$ of Equation (23)

$$\begin{aligned}(g^*, \beta^*) = \underset{g \in \mathcal{H}_2^{(\beta)}, \beta \in \Theta}{\text{argmax}} \quad & \mathbb{E}\left[Yg(\beta^\top X, A) - b(g(\beta^\top X, A))\right] \\ \text{subject to} \quad & \mathbb{E}\left[g(\beta^\top X, A)|X\right] = 0,\end{aligned} \tag{24}$$

where convention that the terms that do not contain the parameter of interest (i.e., $a(\phi)$ and $c(Y, \phi)$) can be dropped from the expression of a log-likelihood) the expectation in the criterion function is with respect to the joint distribution of $(Y, A, X)$. Thus, the solution $(g^*, \beta^*)$ on the left-hand side of Equation (24) corresponds to the minimizer of the Kullback–Leibler (KL) divergence between the distributions with the working mean model (23) and the true mean model (21). In Equation (24), $b(s) = s^2/2$ for a Gaussian $Y$ (for which the optimization (9) is a special case of Equation (24)), $b(s) = \log\{1 + \exp(s)\}$ for a Bernoulli $Y$, and $b(s) = \exp(s)$ for a Poisson $Y$.

The constraint $\mathbb{E}\left[g(\beta^\top X, A)|X\right] = 0$ in optimization (24) implies

$$\begin{aligned}&\mathbb{E}\left[Yg(\beta^\top X, A) - b(g(\beta^\top X, A))\right] \\ &= \mathbb{E}\left[\left\{\mu_0(X) + g_0(\beta_0^\top X, A)\right\}g(\beta^\top X, A) - b(g(\beta^\top X, A))\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mu_0(X)g(\beta^\top X, A)|X\right]\right] + \mathbb{E}\left[g_0(\beta_0^\top X, A)(g(\beta^\top X, A) \right. \\ &\quad \left. - b(g(\beta^\top X, A))\right] \\ &= \mathbb{E}\left[g_0(\beta_0^\top X, A)(g(\beta^\top X, A) - b(g(\beta^\top X, A))\right],\end{aligned} \tag{25}$$

which is free of $\mu_0$ in model (21). Therefore, the left-hand side, $(g^*, \beta^*)$, of Equation (24) does not depend on the unspecified $X$ "main" effect function $\mu_0$.

*Proposition 1.* The solution $(g^*, \beta^*)$ of the constrained optimization problem (24) satisfies:

$$g_0 = h^{-1} \circ g^* \quad \text{and} \quad \beta_0 = \beta^* \tag{26}$$

(a.s.), where $g_0 \in \mathcal{H}_2^{(\beta_0)}$ and $\beta_0 \in \Theta$ are given from the true mean model (21), and the function $h^{-1}$ is the inverse of the canonical link function associated with the assumed exponential family distribution and the operator $\circ$ represents the composition of two functions.

The proof of Proposition 1 is in Section A.1 (supplemental materials). In Equation (26), $h^{-1}(s) = s$ (the identity function) for a Gaussian $Y$, $h^{-1}(s) = \exp(s)/\{1 + \exp(s)\}$ for a Bernoulli $Y$, and $h^{-1}(s) = \exp(s)$ for a Poisson $Y$.

To solve (24) based on observed data $(Y_i, A_i, X_i)$ $(i = 1, \ldots, n)$, we replace the quadratic loss term in Equation (16) by the negative of the log-likelihood of the data. For a fixed $\beta \in \Theta$, given the representation of $g$ with $\boldsymbol{D\theta}$ in Equation (13), the basis coefficient $\boldsymbol{\theta}$ is estimated by the *inner* iteratively reweighted least squares (IRLS) and the associated smoothing parameters ($\lambda$ and $\check{\lambda}$ in Equation (16)) are estimated by the *outer* optimization of, for example, REML or GCV, as part of Step 1 (of Section 3.2) of the model fitting. The only adjustment to be made to the conventional GAM estimation is to enforce the constraint $\mathbb{E}[g(\beta^\top X, A)|\beta^\top X] = 0$ on $g$. As in Section 3.1, this constraint can be absorbed into the tensor product basis representation (13). For Step 2 (of Section 3.2), once $g$ is profiled out (by Step 1), we replace the *normal residual* vector in Equation (19), that is, $Y_{n\times 1} - g(\boldsymbol{X}\tilde{\beta}, A_{n\times 1})$, by the *working residual* from the final IRLS fit of Step 1. Then we perform a weighted least squares (instead of the least squares (20)) for $\beta$, where the weights are given from the final IRLS fit of Step 1. We alternate between the Steps 1 and 2 until convergence of $\hat{\beta}$, as in Section 3.2. The resulting estimate for $(g^*, \beta^*)$ in Equation (24) is then used to estimate $(g_0, \beta_0)$ in (21), based on the relationship (26). We also note that the framework (24) can be extended to incorporate a multinomial response, an instance of which we discuss in the following remark.

*Remark 1.* The approach (24) to optimizing the component $(g, \beta)$ of SIMSL can be generalized to the context of a proportional odds single-index model, where we have an ordinal response $Y$, in which its value exists on an arbitrary scale (in $K$ categories), with only the relative ordering between different values being important. To deal with such a case, we introduce a length-$K$ response vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_K)^\top$, where its component $Y_j$ denotes the indicator for category $j$, with associated vector of probabilities $(p_1, p_2, \ldots, p_K)^\top$, in which $p_K = 1 - \sum_{j=1}^{K-1} p_j$ (and $p_j > 0$), and their cumulative probabilities: $P(Y \leq j) = q_j = \sum_{s=1}^{j} p_s$ $(j = 1, 2, \ldots, K-1)$. We can model these cumulative probabilities $(q_1, q_2, \ldots, q_{K-1}, 1)^\top$ by a cumulative logit SIMSL:

$$h(P(Y \leq j|X, A)) = h(q_j(X, A)) = \alpha_j + \mu(X) + g(\beta^\top X, A) \quad (j = 1, 2, \ldots, K-1) \tag{27}$$

where $\alpha_j \in \mathbb{R}$ $(\alpha_1 < \alpha_2 < \ldots < \alpha_{K-1})$ are unknown cut-point parameters associated with the ordered response categories $j = 1, 2, \ldots, K-1$, and $h(s) = \log(s/(1-s))$ is the logit link. For the $K$-category multinomial exponential family response $\boldsymbol{Y}$, let us consider its canonical parameter $\boldsymbol{\eta} := (\eta_1, \eta_2, \ldots, \eta_{K-1}, 0)^\top \in \mathbb{R}^K$, where its nonzero components are specified by: $\eta_j = \log\left(p_j/p_K\right) = \log\left((q_j - q_{j-1})/(1 - q_{K-1})\right)$ $(j = 1, \ldots, K-1)$ (with $q_0 = 0$), in which the cumulative probabilities $(q_1, q_2, \ldots, q_{K-1}, 1)^\top$ are specified by model (27). This multinomial exponential family representation for the distribution of $Y$ allows us to use the optimization framework (24), with its criterion function extended to incorporate the multivariate response: $\mathbb{E}\left[\boldsymbol{Y}^\top \boldsymbol{\eta} - b(\boldsymbol{\eta})\right]$, where $b(\boldsymbol{\eta}) = \log\{1 + \sum_{j=1}^{K-1} \exp(\eta_j)\}$, for optimization of the cumulative logit SIMSL (27). The threshold-point parameters $\alpha_j \in \mathbb{R}$ $(\alpha_1 < \alpha_2 < \cdots < \alpha_{K-1})$ in model (27) are estimated as part of Step 1 of model fitting in Section 3.2 (alongside the model smoothing parameters), in which the model (27), for each fixed $\beta$ that is estimated as part of Step 2, is optimized via the penalized IRLS with an empirical version of $-\mathbb{E}\left[\boldsymbol{Y}^\top \boldsymbol{\eta} - b(\boldsymbol{\eta})\right]$; the Steps 1 and 2 are iterated until convergence. The heterogeneous treatment effect $g(\beta^\top X, A)$ in model (27) does not depend on $j$ nor $\mu(X)$; this allows us to develop an individualized dose rule independently of the arbitrary categorization of $Y$, and of the unspecified $X$ main effect $\mu(X)$ that does not influence the treatment effect. Given $X$, the cumulative logit SIMSL-based individualized dose rule is $f(X) = \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, g(\beta^\top X, a)$.

In Section C (supplemental materials), we provide a real data example illustrating the approach (24) to modeling interaction effects between $X$ and $A$ on a count response variable $Y$, and a simulation illustration for the utility of the proportional odds SIMSL (27) when the treatment response $Y$ is ordinal, which is common in biomedical and epidemiological studies.

## 5. Simulation Example

In this section, we consider a set of simulation studies with data generated from the four scenarios described in Chen, Zeng, and Kosorok (2016). We generate $p$-dimensional covariates $X = (X_1, \ldots, X_p)^\top$, where each entry is generated independently from Uniform$[-1, 1]$. In Scenarios 1 and 2, the treatment $A$ is generated from Uniform$[0, 2]$ independently of $X$, mimicking a randomized trial. In Scenarios 3 and 4, the distribution of $A$ (described below) depends on $X$, mimicking an observational study setting. In each scenario, the outcome $Y$ given $(X, A)$ is generated from the standard normal distribution, with the following four different mean function scenarios:

1. *Scenario 1:* $\mathbb{E}[Y|X, A] = 8 + 4X_1 - 2X_2 - 2X_3 - 25(f_{\text{opt}}(X) - A)^2$, where $f_{\text{opt}}(X) = 1 + 0.5X_1 + 0.5X_2$. Here, the optimal individualized dose rule is a linear function of $X$.
2. *Scenario 2:* $\mathbb{E}[Y|X, A] = 8 + 4\cos(2\pi X_2) - 2X_4 - 8X_5^3 - 15|f_{\text{opt}}(X) - A|$, where

$$f_{\text{opt}}(X) = 0.6(-0.5 < X_1 < 0.5) + 1.2(X_1 > 0.5) + 1.2(X_1 < -0.5) + X_4^2 + 0.5\log(|X_7| + 1) - 0.6.$$

   Here, the optimal individualized dose rule is a nonlinear function of $X$.
3. *Scenario 3* is the same as in Scenario 2, except that $A$ depends on $X$ as follows:

$$A \sim \begin{cases} \text{TruncN}\left(-0.5 + 0.5X_1 + 0.5X_2, 0, 2, 0.5\right), & \text{when } X_3 < 0 \\ \text{TruncN}\left(|0.5 + 1.5X_2|, 0, 2, 1\right), & \text{when } X_3 > 0 \end{cases}$$

   where TruncN$(\mu, a, b, \sigma)$ denotes the truncated normal distribution with mean $\mu$, lower bound $a$ and upper bound $b$, and standard deviation $\sigma$.
4. *Scenario 4* is the same as in Scenario 2, except that $A$ depends on $X$ as follows:

$$A \sim \text{TruncN}\left(f_{\text{opt}}(X), 0, 2, 0.5\right).$$

Following Chen, Zeng, and Kosorok (2016), we set $p = 30$ in Scenario 1, and $p = 10$ for Scenarios 2, 3, and 4. (In Section A.2 of supplementary materials, we describe how these scenarios mean models can be reparameterized to fit in the framework (4).) For each simulated dataset, we apply the proposed method of estimating the $A$-by-$X$ interaction term in the SIMSL (4) based on Equation (9), and the optimal dose rule $f_{\text{opt}}$ by $\hat{f}(X) = \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, \hat{g}^*(\hat{\beta}^{*\top} X, a)$. We simulated 200 datasets for each scenario. For comparison, we report results of the estimation approaches considered in Chen, Zeng, and Kosorok (2016), including their Gaussian kernel-based outcome-weighted learning (K-O-learning) and linear kernel-based outcome-weighted learning (L-O-learning). We also report a support vector regression (SVR; Vapnik 1995; Smola and Scholopf 2004) with a Gaussian kernel to estimate the nonlinear relationship between $Y$ and $(X, A)$ (Zhao, Kosorok, and Zeng 2009) that was used for comparison. In Scenario 1, we used $(X, A)$ as the predictors for the outcome in the SIMSL. In Scenarios 2, 3 and 4, we used $(X, X^2, A)$ (i.e., including a quadratic term in $X$) as the predictors for the SIMSL.

Since we are simulating data from known models in which the true relationship $\mathbb{E}[Y|X, A]$ is known, we can compare the estimated dose rules $\hat{f}$ derived from each method in terms of the value (1). Specifically, an independent test set of size $\tilde{n} = 5000$ was generated and the value (1) of $\hat{f}$ was approximated using $\hat{\mathcal{V}}(\hat{f}) = \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \mathbb{E}[Y_i|X_i, A_i = \hat{f}(X_i)]$, for each simulation

**Table 1.** Average (sd) value $\hat{\mathcal{V}}(f)$ from 200 replicates from randomized trials.

| | $n$ | SIMSL | K-O-learning | L-O-learning | SVR |
|---|---|---|---|---|---|
| Scenario 1 | 50 | 1.04 (4.06) | 4.78 (0.48) | **4.83 (1.40)** | −12.21(7.53) |
| | 100 | **6.63 (0.63)** | 5.69 (0.40) | 5.39 (0.93) | −2.57(6.34) |
| | 200 | **7.45 (0.20)** | 6.68 (0.26) | 6.85 (0.34) | 3.46(1, 97) |
| | 400 | **7.77 (0.08)** | 7.28 (0.15) | 7.41 (0.14) | 6.13(0.47) |
| | 800 | **7.88 (0.04)** | 7.54 (0.08) | 7.67 (0.08) | 7.36(0.12) |
| Scenario 2 | 50 | 0.90 (2.04) | **2.00 (0.29)** | 1.16 (0.71) | −1.96(1.70) |
| | 100 | **3.65 (0.76)** | 2.19 (0.43) | 1.57 (0.52) | 0.24(1.42) |
| | 200 | **4.71 (0.41)** | 2.84 (0.37) | 2.02 (0.30) | 2.01(0.84) |
| | 400 | **5.25 (0.20)** | 3.69 (0.27) | 2.30 (0.18) | 3.47(0.37) |
| | 800 | **5.59 (0.12)** | 4.41 (0.19) | 2.49 (0.10) | 4.35(0.19) |

NOTE: In both settings, the oracle $f_{\text{opt}}$ attains a value $\mathcal{V}(f_{\text{opt}}) = 8$ (boldface denotes the largest in each row).

**Table 2.** Average (sd) value $\hat{\mathcal{V}}(f)$ from 200 replicates from observational studies.

| | $n$ | SIMSL | K-O-learning | K-O-learning(Prp) | SVR |
|---|---|---|---|---|---|
| Scenario 3 | 200 | **4.03 (0.97)** | 2.68 (0.30) | 2.74 (0.29) | 1.99(0.83) |
| | 800 | **5.46 (0.20)** | 4.06 (0.30) | 4.19 (0.20) | 4.09(0.28) |
| Scenario 4 | 200 | **4.07 (0.72)** | 3.29 (0.28) | 3.23 (0.28) | −0.95(1.57) |
| | 800 | **5.51 (0.19)** | 4.91 (0.14) | 4.73 (0.17) | 3.04(0.52) |

NOTE: In both settings, the oracle $f_{\text{opt}}$ attains a value $\mathcal{V}(f_{\text{opt}}) = 8$ (boldface denotes the largest in each row).

run. Given each scenario and a training sample size $n$, we replicate the simulation experiment 200 times, each time estimating the value. Again, following Chen, Zeng, and Kosorok (2016), we report the averaged estimated values (and standard deviations) for the cases where $\hat{f}$ is estimated from a training set of size $n = 50, 100, 200, 400$, and 800 for Scenarios 1 and 2, and the cases with $n = 200$ and 800 for Scenarios 3 and 4. The simulation results are given in Tables 1 and 2.

The results in Tables 1 and 2 indicate that the proposed regression method for optimizing dose rules outperforms the alternative approaches presented in Chen, Zeng, and Kosorok (2016) in all cases except when the training sample size is very small ($n = 50$). In Table 2, K-O-learning(Prp) refers to the propensity score-adjusted K-O-learning of Chen, Zeng, and Kosorok (2016). When $n = 50$, the outcome-weighted learning methods outperform the regression-based approaches (i.e., SIMSL and SVR), especially for Scenario 1 (with $p = 30$) where the regression approaches exhibit large variances. However, when $n = 100$, the performance of SIMSL improves dramatically in terms of both value and small variance. We also note that using $(A, X, X^2)$ instead of $(A, X)$ as predictors of SIMSL in Scenario 2 lead to a substantial improvement in performance. If $(X, A)$ is used for SIMSL in Scenario 2, the estimated values (and sd) are: −0.98(2.15), 0.56(1.54), 1.91(0.89), 2.70(0.64), and 3.23(0.41), for $n = 50, 100, 200, 400$ and 800, respectively.

## 6. Application to Optimization of the Warfarin Dose With Clinical and Pharmacogenetic Data

In this section, the utility of the SIMSL approach to personalized dose finding is illustrated from an anticoagulant study. Warfarin is a widely used anticoagulant to treat and prevent blood clots. The therapeutic dosage of warfarin varies widely across patients. Our analysis of the data will broadly follow that of Chen, Zeng, and Kosorok (2016). After removing patients with missing data, the dataset provided by International Warfarin Pharmacogenetics Consortium et al. (2009) (publicly available to download from *https://www.pharmgkb.org/downloads/*) consists of 1780 subjects, including information on patient covariates ($X$), final therapeutic dosages ($A$), and patient outcomes (INR, International Normalized Ratio). INR is a measure of how rapidly the blood can clot. For patients prescribed warfarin, the target INR is around 2.5. To convert the INR to a measurement responding to the warfarin dose level, we construct an outcome $Y = -|2.5 - \text{INR}|$, and a larger value of $Y$ is considered desirable.

There were 13 covariates $X = (X_1, \ldots, X_{13})^\top$ in the dataset (both clinical and pharmacogenetic variables): weight ($X_1$), height ($X_2$), age ($X_3$), use of the cytochrome P450 enzyme inducers ($X_4$; the enzyme inducers considered in this analysis includes phenytoin, carbamazepine, and rifampin), use of amiodarone ($X_5$), gender ($X_6$; 1 for male, 0 for female), African or black race ($X_7$), Asian race ($X_8$), the VKORC1 A/G genotype ($X_9$), the VKORC1 A/A genotype ($X_{10}$), the CYP2C9 *1/*2 genotype ($X_{11}$), the CYP2C9 *1/*3 genotype ($X_{12}$), and the other CYP2C9 genotypes (excluding the CYP2C9 *1/*1 genotype which is taken as the baseline genotype) ($X_{13}$). Further details on these covariates are given in International Warfarin Pharmacogenetics Consortium et al. (2009). The first 3 covariates (weight, height, and age) were treated as continuous variables, standardized to have mean zero and unit variance; the other 10 covariates are indicator variables.

In estimating the optimal individualized dose rule $f_{\text{opt}}$, modeling the drug (dose level $A$) interactions with the patient covariates $X$ is essential. Under the proposed SIMSL approach (4), $f_{\text{opt}}(X) = \underset{a \in \mathcal{A}}{\text{argmax}}\, g(\beta^\top X, a)$ and thus the $A$-by-$X$ interaction effect term $g(\beta^\top X, A)$ is the target component of interest. In SIMSL, due to orthogonality (7), we can solve Equation (9) for $(g, \beta)$, without having to model the $\mu$ term in Equation (4). However, modeling the $\mu$ term, even with a misspecified working model, can generally improve the efficiency of the estimator of $(g, \beta)$ (i.e., yielding smaller variances for estimators; see Park et al. (2020) for a theoretical justification in the case where treatment $A$ is a discrete or binary variable), which leads to improve the efficiency of the estimator of $f_{\text{opt}}$ (see Section C.1 (supplementary materials) for a simulation illustration in which an enhanced estimation performance is illustrated when the $X$ main effect is incorporated to the estimation of $f_{\text{opt}}$).

Thus, in this application, we model the $\mu$ term of Equation (4) with a possibly misspecified working model, which consists of a set of linear terms for the indicators $X_4, \ldots, X_{13}$ and a set of cubic $P$-spline smooth terms for the continuous covariates $X_1$, $X_2$ and $X_3$. These terms are estimated alongside the heterogeneous treatment effect term $g(\beta^\top X, A)$ by the procedure described in Section C.1 (supplementary materials), which is a slight modification to that described in Section 3.2. (The estimated coefficient $\beta$, with or without incorporating the $\mu$ term in the estimation, along with their bootstrap confidence intervals, are provided in Section D of supplementary materials.) The third panel in Figure 1 displays a surface plot of the estimated 2-dimensional link function $g(\beta^\top X, A)$, showing an interactive relationship on the index-treatment domain. The first two panels in Figure 1 display the estimated marginal effect function for the dose $A$ and that for the estimated index $\beta^\top X$.

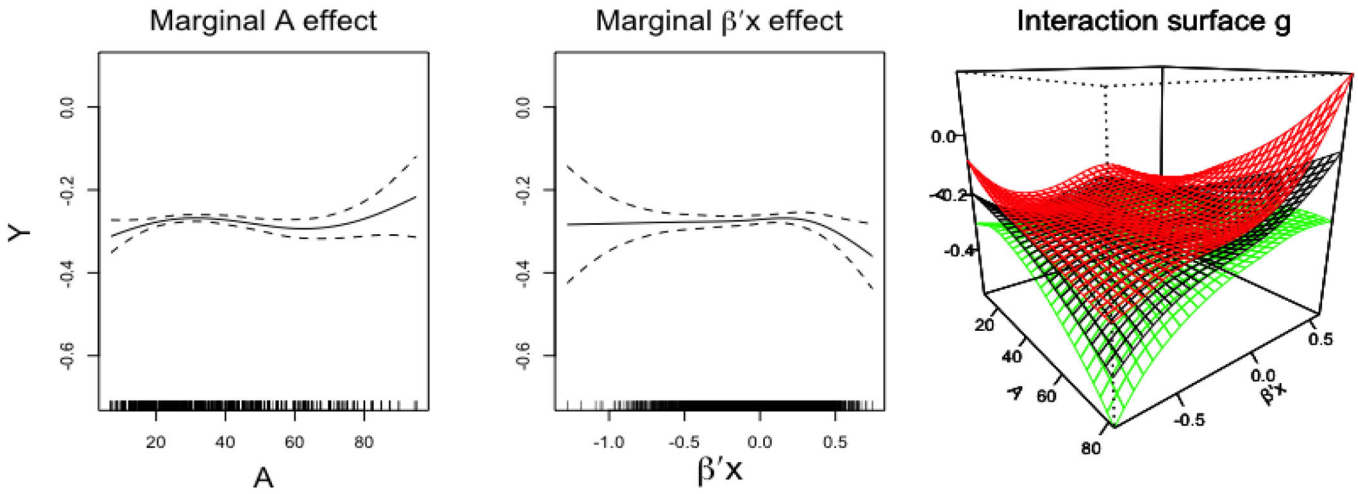


**Figure 1.** The first two panels: the marginal effect of dose $A$ (left panel) and that of the estimated single-index (middle panel) with 95% confidence bands (dashed curves) given the estimated $\beta^\top X$. The third panel: the estimated link surface ($g$) for the dose ($A$) and index ($\beta^\top X$) interaction; the red and green surfaces are at $\pm 2$ standard error from the estimated surface (the black) in the middle, conditioning on the estimated single index.

We construct a 95% normal approximation bootstrap confidence interval for $\beta$, based on 500 bootstrap replications (see Section C.3 (supplementary materials) for the confidence interval construction and for a coverage probability simulation). The confidence intervals for the $\beta_j$'s associated with the covariates weight ($X_1$), height ($X_2$), the use of the cytochrome P450 enzyme inducers ($X_4$), Asian race ($X_8$), the VKORC1 A/G genotype ($X_9$), and the CYP2C9 *1/*2 genotypes ($X_{11}$) do not include 0. We infer that these covariates are potentially clinically important drug effect modifiers, interacting with warfarin in their effects on INR.

Chen, Zeng, and Kosorok (2016) noted that the analysis results from International Warfarin Pharmacogenetics Consortium et al. (2009), as well as their linear kernel-based outcome-weighted learning results, suggest increasing the dose if patients are taking Cytochrome P450 enzyme ($X_4$). Roughly speaking, the interaction surface $g$ (the right-most panel) in Figure 1 indicates that for a smaller value of $\beta^\top X$ (e.g., $\beta^\top X < -0.3$), a moderate or a relatively low dose $A$ (e.g., $A < 50$) may be preferred, whereas for a larger value of $\beta^\top X$ (e.g., $\beta^\top X > 0$), a relatively high dose $A$ (e.g., $A > 50$) may be preferred. Considering the sign of the estimated coefficient ($\hat{\beta}_4 = -0.60$) associated with $X_4$, this is roughly consistent with International Warfarin Pharmacogenetics Consortium et al. (2009) and Chen, Zeng, and Kosorok (2016).

To evaluate the performance of the individualized dose rules estimated from the six methods, including the propensity score-adjusted outcome-weighted learning with a linear/Gaussian kernel, denoted as L-O-learning (Prp) and K-O-learning (Prp), respectively) considered in Section 5, we randomly split the dataset at a ratio of one-to-one into a training set and a testing set, replicated 100 times, each time estimating $f_{\text{opt}}$ using the 6 methods based on the training set, and estimating the value (1) of each estimated $f_{\text{opt}}$ based on the testing set. Unlike the simulated data in Section 5, the true relationship between the covariate-specific dose and the response is unknown. Therefore, for each dose rule $f$, we need to estimate the value (1) from the testing data. Given a dose rule $f$, only a very small proportion (or none) of the observations will satisfy $A_i = f(X_i)$, and thus only a very small proportion (or none) of the observations in the testing data will contribute information to estimate the value (1). However, Cai and Tian (2016) noted that the value (1) for each $f$ can also be written as $\mathcal{V}(f) = \mathbb{E}[\mathbb{E}[Y|A = f(X), f(X)]]$. Therefore, using a two-dimensional smoother of $A$ and $f(X)$ for $Y$, one may first obtain a nonparametric estimate of $\mathbb{E}[Y|A, f(X)]$, denoted as $\hat{m}(A, f(X))$, and then $\mathcal{V}(f)$ may be estimated as $\hat{\mathcal{V}}(f) = n^{-1} \sum_{i=1}^{n} \hat{m}(f(X_i), f(X_i))$. Specifically, given a dose rule $f$ estimated from a training set, we can estimate $\mathbb{E}[Y|A, f(X)]$ based on $(Y_i, A_i, f(X_i))$ from a test set, using a set of thin plate regression spline bases obtained from a rank-100 eigen-approximation to a thin plate spline, with the smoothness parameter selected by REML, implemented via the R (R Core Team 2019) function `mgcv::gam` (Wood 2019). A thin plate spline is an isotropic smooth; isotropy is often appropriate for two variables observed on the same scale, which is the case here.

Figure 2 displays boxplots describing the distributions for the estimated values (1) of the approaches "SIMSL(w. $X$ main)" (SIMSL *with* the $\mu$ term in the estimation) and "SIMSL(w.o. $X$ main)" (SIMSL *without* the $\mu$ term in the estimation), and the other five estimation methods described in Section 5, obtained from the aforementioned 100 random training/testing splits. The boxplots indicate that the proposed SIMSL methods (note that "SIMSL(w. $X$ main)" slightly outperforms "SIMSL(w.o. $X$ main)") and the propensity-score adjusted K-O-learning of Chen, Zeng, and Kosorok (2016) perform at a similar level, while outperforming all other approaches, illustrating the potential utility of the proposed method. In comparison to the outcome-weighted learning approach of Chen, Zeng, and Kosorok (2016), one advantage of the proposed approach is that it allows visualization of the estimated interactive structure on the dose-index domain, as illustrated in the right panel of Figure 1. Additionally, if each of the covariates is standardized to have, say, unit variance, then the relative importance of each covariate in characterizing the heterogeneous dose response can be determined by the magnitude of the estimated coefficients in

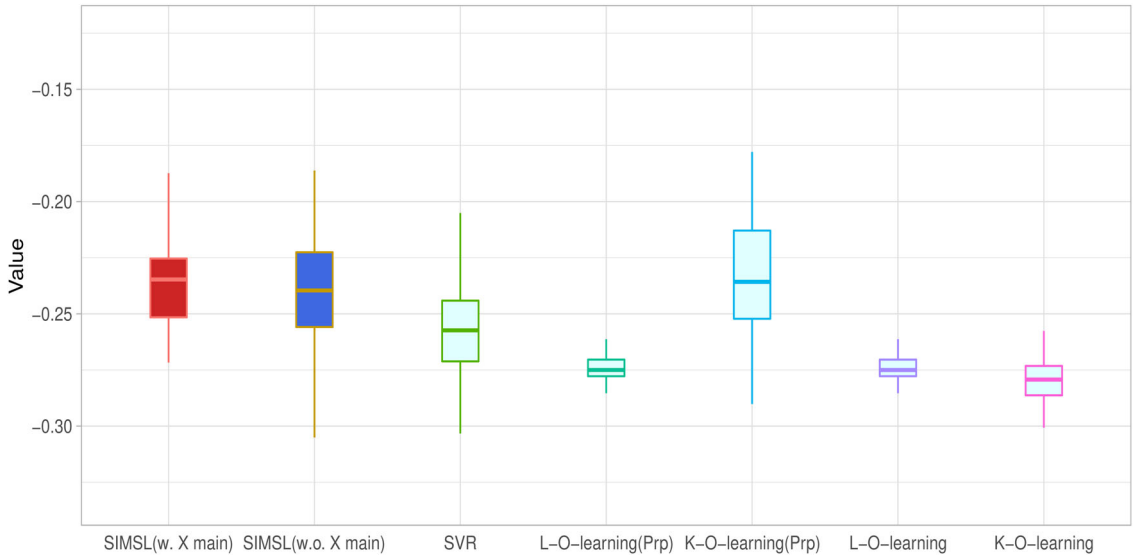


**Figure 2.** Boxplots of the estimated values of the individualized dose rules using 7 approaches, obtained from 100 randomly split testing sets. Mean (and standard deviation) of the value estimates: SIMSL(w. *X* main): -0.231 (0.03); SIMSL(w.o. *X* main): −0.237 (0.03); SVR: −0.254 (0.02); L-O-learning(Prp): −0.274 (0.01); K-O-learning(Prp): −0.234 (0.03); L-O-learning: −0.274 (0.01); and K-O-learning: −0.279 (0.01).

$\beta$, rendering a potentially useful interpretation when examining the drug–covariates interactions.

## 7. Discussion

In this article, we proposed a single-index model that utilizes a surface link-function as a function of a linear projection of covariates and a continuous "treatment" variable, which parsimoniously represents the interaction effect between covariates and a treatment defined on a continuum. The model provides an intuitive tool for investigating personalized dose finding in precision medicine, without the need for a significant change in the established generalized additive regression modeling framework.

One important limitation is that the confidence band associated with the estimated surface $g$ is computed conditional on the estimated $\beta^\top X$, and the uncertainty in $\beta$ is not accounted for. The fact that the domain of $g$ varies depending on the estimate of $\beta$ complicates the confidence band construction for $g$. One potential approach is to consider a Bayesian framework and a posterior distribution of $g(\beta^\top X, A)$, and make probabilistic statements about the prediction of the component $g(\beta^\top X, A)$ given $(X, A)$. Furthermore, depending on context, for scientific interpretability of the model, it may be sometimes desirable to consider shape constraints such as monotonicity or convexity/concavity (see Section C.4 (supplementary materials) for discussion on shape constraints), as well as optimization under safety constraints (Laber et al. 2018). The development of a Bayesian model estimation and inference, with potential monotonicity or convexity/concavity constraints on the link surface is currently under investigation. In many applications, only a subset of variables may be useful in determining an optimal individualized dose rule. Also, high-dimensional settings can lead to instabilities and issues of overfitting. Forthcoming work will introduce a regularization method that can both avoid overfitting and choose among multiple potential covariates by obtaining a sparse estimate of the single-index coefficient $\beta$. Future extensions of this work could also include an extension to incorporate a functional covariate.

## Supplemental Materials

*Supplementary Materials:* a pdf file containing supporting information for the main manuscript, including the proof of Proposition 1, a real data analysis and additional simulations illustrating an application of the generalized single-index regression approach in Section 4, and construction of bootstrap confidence intervals and supplementary information for Section 6.

*R-package for SIMSL routine:* R-package `simsl` (Park et al. 2021) available on CRAN containing code to perform the proposed single-index regression method, and the datasets and the simulation examples illustrated in this article.

## Funding

This work was supported by National Institute of Mental Health (NIH grant no. 5 R01 MH099003) and National Center for Advancing Translational Sciences (grant no. 3 UL1TR001445-06A1S2).

## References

Cai, T., and Tian, L. (2016), "Comment: Personalized Dose Finding Using Outcome Weighted Learning," *Journal of the American Statistical Association*, 111, 1521–1524. [8]

Chen, G., Zeng, D., and Kosorok, M. R. (2016), "Personalized Dose Finding Using Outcome Wieghted Learning," *Journal of the American Statistical Association*, 111, 1509–1547. [1,2,6,7,8]

Chen, H., Lu, W., and Song, R. (2021), "Statistical Inference for Online Decision Making: In a Contextual Bandit Setting," *Journal of the American Statistical Association*, 116, 240–255, doi:10.1080/01621459.2020.177009. [2]

de Boor, C. (2001), *A Practical Guide to Splines*, New York: Springer-Verlag. [3]

Dudík, M., Erhan, D., Langford, J., and Li, L. (2014), "Doubly Robust Policy Evaluation and Optimization," *Statistical Science*, 29, 485–511. [1]

Eilers, P., and Marx, B. (2003), "Multivariate Calibration With Temperature Interaction Using 2-Dimensional Penalized Signal Regression," *Chemometrics and Intelligence Laboratory Systems*, 66, 159–174. [3]

Gill, R. D., and Robins, J. M. (2001), "Causal Inference for Complex Longitudinal Data: The Continuous Case," *Annals of Statistics*, 29, 1785–1811. [1]

Gunter, L., Zhu, J., and Murphy, S. (2011), "Variable Selection for Qualitative Interactions in Presonalized Medicine While Controlling the Family-Wise Error Rate," *Journal of Biopharmaceutical Statistics*, 21, 1063–1078. [2]

Hastie, T., and Tibshirani, R. (1999), *Generalized Additive Models*, London: Chapman & Hall Ltd. [2,4]

Klein, T., Altman, R., Eriksson, N., Gage, B., Kimmel, S., Lee, M., Limdi, N., Page, D., Roden, D., Wagner, M., Caldwell, M., and Johnson, J., International Warfarin Pharmacogenetics Consortium. (2009), "Estimation of the Warfarin Dose With Clinical and Pharmacogenetic Data," *The New England Journal of Medicine*, 360, 753–674. [7,8]

Jun, K. S., Bhargava, A., Nowak, R., and Willett, R. (2017), "Scalable Generalized Linear Bandits: Online Computation and Hashing," *Advances in Neural Information Processing Systems*, 30, 99–109. [2]

Kaelbling, L., Littman, M., and Moore, A. (1996), "Reinforcement Learning: A Survey," *Journal of Artificial Intelligence Research*, 4, 237–285. [2]

Kallus, N., and Zhou, A. (2018), "Policy Evaluation and Optimization With Continuous Treatments," *International Conference on Artificial Intelligence and Statistics*, 84, 1243–1251. [1]

Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. (2017), "Nonparametric Methods for Doubly Robust Estimation of Continuous Treatment Effects," *Journal of Royal Statistical Society*, Series B, 79, 1229–1245. [2]

Kleinberg, R., Slivkins, A., and Upfal, E. (2019), "Bandits and Experts in Metric Spaces," *Journal of the ACM*, 66, 77. [2]

Krishnamurthy, A., Langford, J., Slivkins, A., and Zhang, C. (2020), "Contextual Bandits With Continuous Actions: Smoothing, Zooming, and Adapting," *Journal of Machine Learning Research*, 21, 1–45. [2]

Kveton, B., Zaheer, M., Szepesvari, C., Li, L., Ghavamzadeh, M., and Boutilier, C. (2020), "Randomized Exploration in Generalized Linear Bandits," *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 108, 2066–2076. [2]

Laber, E., Wu, F., Munera, C., Lipkovich, I., Colucci, S., and Ripa, S. (2018), "Identifying Optimal Dosage Regimes Under Safety Constraints: An Application to Long Term Opioid Treatment of Chronic Pain," *Statistics in Medicine*, 37, 1407–1418. [9]

Laber, E. B., and Zhao, Y. (2015), "Tree-Based Methods for Individualized Treatment Regimes," *Biometrika*, 102, 501–514. [1,2]

Lattimore, T. and Szepesvari, C. (2019), *Bandit Algorithms*, Cambridge: Cambridge University Press. [2]

Li, L., Lu, Y., and Zhou, D. (2017), "Provably Optimal Algorithms for Generalized Linear Contextual Bandits," *Proceedings of the 34th International Conference on Machine Learning*, 70, 2071–2080. [2]

Lu, T., Pal, D., and Pal, M. (2010), "Contextual multi-Armed Bandits," *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings*, 9, 485–492. [2]

Lu, W., Zhang, H., and Zeng, D. (2013), "Variable Selection for Optimal Treatment Decision," *Statistical Methods in Medical Research*, 22, 493–504. [2]

Majzoubi, M., Zhang, C., Chari, R. K., A. Langford, J., and Slivkins, A. (2020), "Efficient Contextual Bandits With Continuous Actions," arXiv no: abs/2006.06040 . [2]

Marx, B. (2015), "Varying-Coefficient Single-Index Signal Regression," *Chemometrics and Intellegence Laboratory Systems*, 143, 111–121. [3]

Murphy, S. A. (2003), "Optimal Dynamic Treatment Regimes," *Journal of the Royal Statistical Society*, Series B, 65, 331–355. [1]

Park, H., Petkova, E., Tarpey, T., and Ogden, R. T. (2020), "A Constrained Single-Index Regression for Estimating Interactions Between a Treatment and Covariates," *Biometrics*. doi:10.1111/biom.13320 . [2,7]

——— (2021), *simsl: Single-Index Models With a Surface-Link* (R package version 0.2.1). [9]

Perchet, V., and Rigollet, P. (2013), "The Multi-Armed Bandit Problem With Covariates," *The Annals of Statistics*, 41, 693–721. [2]

Qian, M., and Murphy, S. A. (2011), "Performance Guarantees for Individualized Treatment Rules," *The Annals of Statistics*, 39, 1180–1210. [2]

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. [8]

Rich, B., Moodie, E. E., and Stephens, D. A. (2014), "Simulating Sequential Multiple Assignment Randomized Trials to Generate Optimal Personalized Warfarin Dosing Strategies," *Clinical Trials*, 11, 435–444. [2]

Robins, J. (2004), *Optimal Structural Nested Models for Optimal Sequential Decisions*, New York: Springer. [1]

Schulz, J., and Moodie, E. E. (2021), "Doubly Robust Estimation of Optimal Dosing Strategies," *Journal of the American Statistical Association*, 116, 256–268, doi:10.1080/01621459.2020.1753521. [2]

Slivkins, A. (2014), "Contextual Bandits With Similarity Information," *Journal of Machine Learning Research*, 15, 2533–2568. [2]

Smola, A., and Scholopf, B. (2004), "A Tutorial on Support Vector Regression," *Statistics and Computing*, 14, 199–222. [6]

Vapnik, V. N. (1995), *The Nature of Statistical Learning Theory* (Vol. 8), New York: Springer. [6]

Wood, S. N. (2006), "Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models," *Biometrics*, 62, 1025–1036. [4]

——— (2017), *Generalized Additive Models: An Introduction With R* (2nd ed.), London: Chapman & Hall/CRC. [3,4]

——— (2019), *mgcv: Mixed GAM Computation Vehicle With Automatic Smoothness Estimation* (R package version 1.8.28). [8]

Zhang, B., Tsiatis, A. A., Davidian, M., Zhang, M., and Laber, E. (2012), "Estimating Optimal Treatment Regimes From Classification Perspective," *Statistics*, 1, 103–114. [2]

Zhao, Y., Kosorok, M., and Zeng, D. (2009), "Reinforcement Learning Design for Cancer Clinical Trials," *Statistics in Medicine*, 28, 3294–3315. [6]

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012), "Estimating Individualized Treatment Rules Using Outcome Weighted Learning," *Journal of the American Statistical Association*, 107, 1106–1118. [1]

Zhou, W., Zhu, R., and Zeng, D. (2020), "A Parsimonious Personalized Dose-Finding Model Via Dimension Reduction," *Biometrika*. doi:10.1093/biomet/asaa087. [1]

Zhu, L., Lu, W., Kosorok, M. R., and Song, R. (2020), "Kernel Assisted Learning for Personalized Dose Finding," arXiv:2007.09811 . [1]