Biostatistics, 2023, **00**, 1–15
https://doi.org/10.1093/biostatistics/kxad022
Article

Variable selection in high dimensions for discrete-outcome individualized treatment rules: Reducing severity of depression symptoms

Erica E. M. Moodie [1,*], Zeyu Bian[1], Janie Coulombe[2], Yi Lian[1], Archer Y. Yang[3], Susan M. Shortreed[4,5]

[1]McGill University, Department of Epidemiology & Biostatistics, 2001 McGill College Ave, Suite 1200, Montreal, QC Canada H3A 1G1
[2]Université de Montréal, Department of Mathematics & Statistics, Pavillon André-Aisenstadt, Montréal, QC Canada H3C 3J7
[3]McGill University, Department of Mathematics & Statistics, 805 Sherbrooke Street West Montreal, QC Canada H3A 0B9
[4]Kaiser Permanente Washington Health Research Institute, 1730 Minor Ave, Suite 1600, Seattle, WA 98101
[5]University of Washington, Department of Biostatistics, 1705 NE Pacific St, Seattle, WA 98195

*To whom correspondence should be addressed: Erica E. M. Moodie. Email: erica.moodie@mcgill.ca

ABSTRACT

Despite growing interest in estimating individualized treatment rules, little attention has been given the binary outcome setting. Estimation is challenging with nonlinear link functions, especially when variable selection is needed. We use a new computational approach to solve a recently proposed doubly robust regularized estimating equation to accomplish this difficult task in a case study of depression treatment. We demonstrate an application of this new approach in combination with a weighted and penalized estimating equation to this challenging binary outcome setting. We demonstrate the double robustness of the method and its effectiveness for variable selection. The work is motivated by and applied to an analysis of treatment for unipolar depression using a population of patients treated at Kaiser Permanente Washington.

KEYWORDS: Adaptive treatment strategies; Antidepressant treatment; Estimating equations; Precision medicine; Regularization.

1. INTRODUCTION

Many mental illnesses are only partially understood, and treatment can involve considerable trial and error to find the therapy (type, dose, or schedule) that works best for a given patient. Treatment adjustment can occur because mental illness diagnoses are often difficult and based on multiple symptoms that may represent a single illness with a heterogeneous presentation or an illness with several subtypes. Treatment of depression often requires adjustment because individual patients may respond to identical treatments in different ways. This variation motivates the quest for individualized treatment rules (ITRs) (Simon, 2001; Chakraborty, 2011; Coulombe *and others*, 2021), which uses their demographic, clinical, or other characteristics to better guide treatment choices that meet their needs (Murphy, 2003; Robins, 2004; Chakraborty and Moodie, 2013).

Received: December 21, 2022. **Revised:** July 14, 2023. **Accepted:** August 3, 2023

Herein, we focus on ITR (i.e., a single treatment decision) estimation rather than adaptive treatment strategies which consider sequences of decisions.

The statistical literature on ITRs is growing at a fast pace. However, a majority of the developments focus on the continuous-valued outcome setting, where the linear link function is the natural modeling choice. The binary outcome setting, in which non-linear links are common, has received far less attention, in spite of the ubiquity of binary outcomes in much of the medical and epidemiological literature in which ITRs might relevantly be applied. This may be explained in part by the fact that estimation is more challenging with discrete outcomes, where simple weighted regression can no longer be employed to achieve a doubly robust estimator (Tchetgen Tchetgen *and others*, 2010; Zetterqvist and Sjölander, 2015; Bian *and others*, 2022).

Penalization has been used in ITRs (and their sequential counterparts) for continuous outcomes by several authors, first in the the context of singly robust methods Qian and Murphy (2011); Lu *and others* (2013); Goldberg *and others* (2013); Song *and others* (2015); Jeng *and others* (2018) and more recently within doubly robust approaches (Shi *and others*, 2018; Zhang and Zhang, 2018; Bian *and others*, 2021). There has been little consideration of the estimation of ITRs for discrete outcomes in a high-dimensional setting. Recently, Bian *and others* (2022) proposed a penalization approach for a doubly robust ITR estimation procedure in a discrete-outcome setting. The approach is promising, but computationally challenging and limited by the requirement of a good initial estimate for the treatment rule parameters. As well-behaved initial estimates may be difficult to find, particularly in small samples or when there are few outcomes (e.g., the number of "successes" in the binary outcome is small), Bian *and others*'s approach may require unsatisfying, but pragmatic, solutions such as a "manual" pre-selection of candidate tailoring variables. In this work, we aim to address this limitation by deploying a new computational approach to solving Bian *and others*'s proposed doubly robust regularized estimating equation (REE).

This work was motivated by the question of how to choose between two antidepressant drug classes for patients initiating antidepressant medication. In a main analysis, we focus on the choice between selective serotonin reuptake inhibitors (SSRIs) and serotonin and norepinephrine reuptake inhibitors (SNRIs) to minimize the risk of a patient having severe depression symptoms in the 1 year following medication initiation, defined as a response of 15 or greater on the nine-item patient health questionnaire (PHQ) (Kroenke *and others*, 2010). We conducted the study using data from Kaiser Permanente Washington (KPWA), a large, non-profit health care organization in Washington state.

This article is organized as follows. Section 2 details notation and assumptions, then provides a brief review of existing approaches to variable selection for ITRs before describing our approach. We explore the performance of our proposed approach via simulation in Section 3, before addressing, in Section 4, our motivating research question. We conclude with limitations and directions for future investigations in Section 5.

2. METHODS

Our data are from n independent individuals indexed by i. Let $\boldsymbol{X}_i \in \mathbb{R}^p$ be a vector of pre-treatment covariates. We focus on a setting in which both treatment and outcome are binary, with $A_i \in \{0, 1\}$ and $Y_i \in \{0, 1\}$, where $Y_i = 1$ represents the observation of an event such as worsening of symptoms beyond a given threshold or hospitalization within 1 year of diagnosis and $Y_i = 0$ indicates censoring at or before the end of follow-up. We define the counterfactual event indicator Y_i^a as the potential outcome if, possibly contrary to fact, an individual received treatment $a \in \{0, 1\}$. Unless specified otherwise, uppercase, lowercase, and bold denote random variables, realizations of random variables, and vectors, respectively.

An ITR is a decision rule $d(\boldsymbol{X}) : \boldsymbol{X} \to \{0, 1\}$ that takes as arguments pre-treatment covariates and returns a recommended treatment. For an optimal ITR, the decision rule $d^{\text{opt}}(\boldsymbol{x}) := d^{\text{opt}}(\boldsymbol{X} = \boldsymbol{x})$ minimizes the proportion of the population expected to experience a negative outcome, as in our example of experiencing severe symptoms. (Alternatively, an optimal ITR may maximize the proportion expected to experience a positive outcome.) We suppose the axiom of consistency

linking counterfactual to actual outcomes, and assume (i) no interference; (ii) no unmeasured confounding (Robins, 2000); and (iii) positivity.

Linking this notation to our motivating example, we take Y to be an indicator of having a recorded PHQ response higher than 15 within 1 year of treatment initiation (excluding the PHQ score measured at treatment initiation), A to be a binary treatment indicator which, in our setting, takes values in $\{0,1\}$ where both levels are drug classes (e.g., SSRI or SNRI), and $\boldsymbol{X}$ to denote a vector of covariates that may predict patient depression severity and may also modify the effect of treatment. This treatment effect moderation is modeled through an interaction between the treatment and covariates, $\boldsymbol{X}$, which can be captured in a statistical regression model, e.g., via interaction terms included on the linear predictor scale in a generalized linear model. The covariates $\boldsymbol{X}$ include demographic, clinical (e.g., PHQ score at treatment initiation), and other health-related (e.g., body mass index) information.

We assume the following semiparametric regression model describes the relationship between Y and $(\boldsymbol{X}, A)$: $\text{logit}\big(\mathbb{E}(Y^a|\boldsymbol{X}=\boldsymbol{x})\big) = \text{logit}\big(\mathbb{E}(Y|\boldsymbol{X}=\boldsymbol{x}, A=a)\big) = f_0(\boldsymbol{x}) + \gamma\,(\boldsymbol{x}, a; \boldsymbol{\psi})$, where f_0 is an unknown baseline function and γ is the blip function (Robins, 2004) characterizing the impact of treatment on the outcome as a contrast on the scale of the linear predictor. We assume that functional form of γ is known, and that it is indexed by a finite-dimensional parameter of interest $\boldsymbol{\psi}$, which we call the blip parameter. The optimal ITR can be obtained by simply optimizing the blip function, e.g., $d^{\text{opt}}(\boldsymbol{x}) = \arg\min_a \gamma\,(\boldsymbol{x}, a; \boldsymbol{\psi})$ for cases that aim to minimize a negative outcome, as in our setting.

2.1. Doubly robust estimation methods for binary outcomes

Doubly robust estimation of ITRs for binary outcomes has been studied in Tchetgen Tchetgen *and others* (2010) and Bian *and others* (2022). The method of Bian *and others* (2022) incorporated a penalization framework for simultaneous variable selection and ITR estimation: the estimation problem is reformulated as an iteratively reweighted generalized linear model, which reduces the computational challenge to solving a minimization problem. Below, we describe the estimation methods proposed in Tchetgen Tchetgen *and others* (2010) and Bian *and others* (2022).

For binary outcomes modeled via a logit link, the blip parameter can be estimated by solving the following A-learning estimating equation:

$$\boldsymbol{U}_1(\boldsymbol{\psi}) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i(a_i - \widehat{\pi}_i^*)\left(y_i - \text{expit}(f(\boldsymbol{x}_i; \widehat{\boldsymbol{\beta}}) + \gamma(\boldsymbol{x}_i, a_i; \boldsymbol{\psi}))\right), \tag{2.1}$$

where $f(\boldsymbol{x}; \boldsymbol{\beta})$ is the posited baseline model (not necessarily identical to f_0), $\widehat{\boldsymbol{\beta}}$ is a plug-in estimator, and, for $\text{expit}(t) = \frac{\exp(t)}{1+\exp(t)}$ and $u(\boldsymbol{x}; \boldsymbol{\phi})$,

$$\widehat{\pi}^* = \left(1 + \frac{(1 - \text{expit}(u(\boldsymbol{x}; \widehat{\boldsymbol{\phi}}))\text{expit}(f(\boldsymbol{x}; \widehat{\boldsymbol{\beta}}))}{\text{expit}(u(\boldsymbol{x}; \widehat{\boldsymbol{\phi}}))\text{expit}(f(\boldsymbol{x}; \widehat{\boldsymbol{\beta}}) + \gamma(x, 0; \boldsymbol{\psi}))}\right)^{-1},$$

is the nuisance treatment model of $\mathbb{E}(A|X, Y=0)$. Note that $u(\boldsymbol{x}; \boldsymbol{\phi})$ is not a (typical) propensity score, but rather a model for the conditional probability of receiving treatment *among* patients who do not experience the outcome. The resulting estimator is consistent for $\boldsymbol{\psi}$ when at least one of $\mathbb{E}(Y|X, A=0)$ or $\mathbb{E}(A|X, Y=0)$ is correctly specified (Tchetgen Tchetgen *and others*, 2010).

To introduce sparsity in estimating the blip parameter $\boldsymbol{\psi}$, it is tempting to directly use a *regularized* A-learning approach that penalizes the estimating function $\boldsymbol{U}_1(\boldsymbol{\psi})$. Specifically, $\boldsymbol{\psi}$ can be estimated by solving the REE (Fu, 2003; Johnson *and others*, 2008; Wang *and others*, 2012): $\boldsymbol{0} \in \boldsymbol{U}_1(\boldsymbol{\psi}) + \partial\Omega_\lambda(\boldsymbol{\psi})$, where $\Omega_\lambda(\cdot)\colon \mathbb{R}^p \to \mathbb{R}$ is a convex sparsity-inducing penalty function, e.g., the LASSO penalty $\Omega_\lambda(\boldsymbol{\psi}) = \lambda\|\boldsymbol{\psi}\|_1$ (Tibshirani, 1996), and λ is the corresponding tuning parameter that controls the amount of regularization. $\partial\Omega_\lambda(\beta)$ is the subdifferential of Ω_λ, i.e.,

the set of all subgradients of $\Omega_\lambda(\cdot)$ at β. For $\Omega_\lambda(\cdot)$, the subgradient of $\Omega_\lambda(\cdot)$ at β is defined as a vector $g \in \mathbb{R}^p$ satisfying $\Omega_\lambda(\beta') \geq \Omega_\lambda(\beta) + g^\top(\beta' - \beta)$ for all β'. Note that the REE is defined as $\mathbf{0} \in \boldsymbol{U}_1(\boldsymbol{\psi}) + \partial\Omega_\lambda(\boldsymbol{\psi})$ instead of $\mathbf{0} = \boldsymbol{U}_1(\boldsymbol{\psi}) + \partial\Omega_\lambda(\boldsymbol{\psi})$, since the subgradient might not be unique.

Regularized A-learning via REE has not been explored in previous work on ITRs, largely due to the computational burden in solving a REE, particularly in high dimensions (Johnson *and others*, 2008; Wang *and others*, 2012). Furthermore, selecting the tuning parameter in an ITR context can be difficult because of the potential for model misspecification.

An alternative approach to introducing sparsity was given by Bian *and others* (2022), who proposed another doubly robust estimating function for binary outcomes:

$$\boldsymbol{U}_2(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \begin{pmatrix} a_i \boldsymbol{x}_i \\ \boldsymbol{x}_i \end{pmatrix} |a_i - \widehat{\pi}_i^*| \left(y_i - \text{expit}(\boldsymbol{x}_i^T \boldsymbol{\beta} + \gamma(\boldsymbol{x}_i, a; \boldsymbol{\psi}))\right).$$

This estimating function was inspired by the form of $\boldsymbol{U}_1(\boldsymbol{\psi})$ and the dynamic weighted ordinary least squares estimator (Wallace and Moodie, 2015). This method is doubly robust if the variables in the blip function are contained in the baseline function, with the advantage that this approach more easily accommodates variable selection, as we describe below. Bian *and others* (2022) demonstrate that the problem of solving the REE $\boldsymbol{U}_2(\boldsymbol{\beta}, \boldsymbol{\psi})$ can be reformulated into a minimization framework. In fact, solving the following ITR REE using $\boldsymbol{U}_2$ via

$$\mathbf{0} \in \boldsymbol{U}_2(\boldsymbol{\beta}, \boldsymbol{\psi}) + \partial\Omega_\lambda(\boldsymbol{\beta}, \boldsymbol{\psi}) \tag{2.2}$$

is asymptotically equivalent to solving a penalized weighted general linear model given an appropriate initial estimator $\widehat{\boldsymbol{\psi}}_{ini}$; we refer to this as the penalized doubly robust (PDR) approach. The use of $\in$ indicates that the solution to this estimating equation is unique only if the subgradient is unique. A solution to (2.2) can be approximated by the following minimization problem:

$$(\widehat{\boldsymbol{\beta}}^{\text{PDR}}, \widehat{\boldsymbol{\psi}}^{\text{PDR}}) = \underset{\boldsymbol{\beta}, \boldsymbol{\psi}}{\arg\min} \left(\sum_{i=1}^{n} |a_i - \widehat{\pi}_i^*| \, L(y_i, \boldsymbol{x}_i, a_i; \boldsymbol{\beta}, \boldsymbol{\psi}) + \Omega_\lambda(\boldsymbol{\beta}, \boldsymbol{\psi}) \right), \tag{2.3}$$

where $L(\cdot\,; \boldsymbol{\beta}, \boldsymbol{\psi})$ is the loss function of the logistic regression. Tuning parameter selection for ITRs can be based directly on this objective function, i.e., the best model is selected so that $L_n^\lambda(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\psi}}) + \kappa_n s_\lambda$ is the smallest among the candidate models, where $L_n^\lambda(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\psi}}) = \sum_{i=1}^n |a_i - \widehat{\pi}_i^*| \, L(y_i, \boldsymbol{x}_i, a_i; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\psi}})$, κ_n is some positive sequence, and s_λ is the number of nonzero components in the model for a given λ. We follow Fan and Tang (2013) and set $\kappa_n = \big(\log(\log n)\big) \log p$, as this can achieve model selection consistency in a high-dimensional setting. Hereafter, we refer to this tuning parameter selection approach as weighted information criterion (WIC).

This promising new approach has a limitation: It relies on a well-behaved initial blip parameter estimate, $\widehat{\boldsymbol{\psi}}_{ini}$, which may be difficult to obtain in small sample sizes. Indeed, in the case study of Bian *and others* (2022), the set of candidate tailoring rules was restricted to only eight variables due to difficulty in achieving convergence. Reducing the number of variables prior to regularization to ensure convergence of the estimators is not ideal when the goal is data-driven selection.

2.2. Computational approach

We propose a new approach for estimating a sparse blip function or $\boldsymbol{\psi}$. The approach is based on a formulation of regularized A-learning (RAL) as a fixed-point problem inspired by Yang *and others* (2021) and Lian (2022), which we refer to as the RALF estimator (i.e., the F indicates a fixed-point RAL). This formulation can be solved using a fixed-point algorithm with Type-I Anderson Acceleration (AA) (Anderson, 1965). The Type-I variant (Fang and Saad, 2009) outperforms

the original AA (Anderson, 1965) but can be very unstable. To address this, Zhang *and others* (2020) recently proposed a stabilized Type-I AA and established that it possesses the global convergence property. We employ a Type-I AA algorithm to solve the REE by leveraging the framework introduced by Yang *and others* (2021), which transforms REE problems into fixed-point problems. The algorithm is computationally very efficient, scales to high-dimensional data, and the performance of RALF does not rely on a good initial estimators as required by PDR.

We choose smoothly clipped absolute deviation (SCAD) (Fan and Li, 2001) as the penalty function for RALF since the resulting coefficients are nearly unbiased (Fan and Li, 2001; Fan and Peng, 2004) and, unlike LASSO, model selection consistency of SCAD does not require the restricted irrepresentable or the mutual incoherence condition (Fan and Lv, 2011). The RALF estimator $\widehat{\boldsymbol{\psi}}^{\text{RALF}}$ using SCAD is defined as the solution to the following REE:

$$\mathbf{0} \in \boldsymbol{U}_1(\widehat{\boldsymbol{\psi}}^{\text{RALF}}) + \partial_C \Omega_\lambda(\widehat{\boldsymbol{\psi}}^{\text{RALF}}), \tag{2.4}$$

where $\boldsymbol{U}_1(\cdot)$ is the unregularized A-learning equation (2.1). The SCAD penalty $\Omega_\lambda(\cdot)$ with regularization parameter λ has the following form:

$$\Omega_\lambda(\boldsymbol{\psi}) = \sum_{j=1}^{p} \rho_\lambda(\psi_j), \text{ where } \rho_\lambda(\psi_j) = \begin{cases} \lambda|\psi_j|, & \text{if } |\psi_j| \le \lambda, \\ \frac{-(\psi_j^2 - 2b\lambda|\psi_j| + \lambda^2)}{2(b-1)}, & \text{if } \lambda < |\psi_j| \le b\lambda, \\ (b+1)\lambda^2/2, & \text{if } |\psi_j| > b\lambda, \end{cases}$$

for each $\psi_j \in \boldsymbol{\psi}$ with $j = 1, \ldots, p$ being the indices, and some $b > 2$. The unknown parameter b can be chosen using cross-validation, but computation can be expensive. We set $b = 3.7$ as suggested by Fan and Li (2001), who showed that this choice gives good practical performance. Note that the regular subdifferential does not exist for the nonconvex SCAD penalty function; thus, we adopt the Clarke (1990) subdifferential of the SCAD penalty, denoted by $\partial_C \Omega_\lambda(\cdot)$

$$(\partial_C \Omega(\boldsymbol{\psi}))_j = \partial_C \rho_\lambda(\psi_j) = \begin{cases} \rho'_\lambda(\psi_j)\text{sgn}(\psi_j), & \psi_j \neq 0, \\ [-\lambda, \lambda], & \psi_j = 0 \end{cases}, \tag{2.5}$$

where $\rho'_\lambda(t) = \lambda I(|t| < \lambda) + \frac{(b\lambda - |t|)_+}{(b-1)} I(|t| \ge \lambda)$. By the derivation in Yang *and others* (2021), solving the REE in (2.4) is equivalent to solving the following fixed-point problem

$$\widehat{\boldsymbol{\psi}}^{\text{RALF}} = f(\widehat{\boldsymbol{\psi}}^{\text{RALF}}), \text{ where } f(\boldsymbol{\psi}) = T_{\tau\lambda}(\boldsymbol{\psi} - \tau \boldsymbol{U}_1(\boldsymbol{\psi})) \tag{2.6}$$

for sufficiently small $\tau > 0$, with the SCAD thresholding operator $T_{\tau\lambda}(\cdot)$ defined as

$$T_{\tau\lambda}(t) = \begin{cases} S_{\tau\lambda}(t), & |t| \le 2\tau\lambda, \\ \frac{b-1}{b-2} S_{\tau \frac{b\lambda}{b-1}}(t) & 2\tau\lambda < |t| \le \tau b\lambda, \\ t & |t| > \tau b\lambda, \end{cases}$$

where $S_c(t) = \text{sgn}(t)\max(|t| - c, 0)$ is the soft-thresholding operator for some $c > 0$. The algorithmic parameter τ plays a role similar to the step sizes in optimization algorithms, although it is not formally defined as such by Yang *and others* (2021). A larger τ usually leads to more aggressive updates, but increases the change that the algorithm diverges. The best choice of τ is highly dependent on the REE, but there is currently no effective data-driven method (e.g., linesearch) to determine an adaptive "step size." For the RALF algorithm, we start with $\tau = 1$ and decrease its value until the algorithm converges.

RALF is summarized in Algorithm 1. As detailed in the Supplementary Materials, our approach avoids expensive computation of the estimating function gradient and its inverse, significantly

Algorithm 1: The RALF algorithm.

Input: Estimating function $\boldsymbol{U}_1(\boldsymbol{\psi})$ in (2.1), maximum memory m, parameter τ.
Output: $\widehat{\boldsymbol{\psi}}^{\text{RALF}}$

Initialize $k = 0$, $\boldsymbol{\psi}^{(0)} = \mathbf{0}^p$;
Set fixed-point problem $f(\boldsymbol{\psi})$ according to (2.6);
repeat
  Choose the memory size $m_k = \min(m, k)$;
  Compute $\boldsymbol{g}^{(k-m_k+r)} = \boldsymbol{\psi}^{(k-m_k+r)} - f(\boldsymbol{\psi}^{(k-m_k+r)})$ for $r = 0, \ldots, m_k$;
  Compute $\boldsymbol{V}_k$ and $\boldsymbol{S}_k$ using (A.3);
  Update $\boldsymbol{\psi}^{(k+1)}$ using (A.2);
  $k := k + 1$;
until *convergence*;
$\widehat{\boldsymbol{\psi}}^{\text{RALF}} = \boldsymbol{\psi}^{(k)}$.

reducing the computation burden of the method. Thus, RALF's fixed-point-based algorithm for solving general REEs provides higher estimation accuracy, computational efficiency, and scalability over existing algorithms.

3. SIMULATIONS

The aim of this numerical study was to evaluate variable selection in the context of ITR estimation for a binary outcome that has many potential tailoring variables but only a few that genuinely assist in assigning treatment. As the double robustness property under different model specification settings of PDR and A-learning was examined in Bian *and others* (2022) with a moderate sample size (500 and 1000), here we focus on a small sample size of 200 with $p = 40$, 80, or 120 tailoring variables. Throughout, we consider the challenging scenario in which the baseline model is misspecified while the treatment model is correctly specified. The blip function is also correctly specified in terms of functional form but includes many unnecessary covariates.

Previous work showed that bias in parameter estimates following selection via penalization is reduced by refitting (e.g., Bian *and others*, 2021; Wu *and others*, 2022): That is, in a "refitting procedure," after the variable selection is performed, the blip parameters are re-estimated by solving the unpenalized estimating equation using only the selected tailoring variables. Hence, the refitted RALF estimator can be a natural choice of the initial estimator. Nevertheless, as noted above, in small sample sizes, an unpenalized estimator is difficult to obtain; indeed, in the simulations, we found that even the calculation of a refitted estimator could fail to converge using the **R** package `rootSolve`. Therefore, we simply used the RALF as the initial estimator for PDR.

We focus on the performance of ITR parameter estimators found via RALF and PDR (Bian *and others*, 2022) as measured by estimation error, the estimated decision rules, and the resulting population outcomes (value function). In our simulations, we consider two approaches for selecting the tuning parameter λ: BIC and the WIC proposed in Bian *and others* (2022).

The data generation procedure is as follows: (Step 1) Generate p independent multivariate normal covariates $X_1, ..., X_p$ with mean 0.6 and unit variance. (Step 2) Set the baseline model to be $f(\boldsymbol{x}; \boldsymbol{\beta}) = -1.2 - 1.5x_1 - 1.2x_2 + 2\sin(x_1) + 2\sin(2x_2) + 2\cos(x_1 - x_2)$. (Step 3) Set the blip function to be $\gamma(x, a; \boldsymbol{\psi}) = a(\psi_0 + \psi_1 x_1)$ for $\psi_0 = 1.2$, and $\psi_1 = -2.2$. (Step 4) Set the nuisance treatment model to be $\mathbb{E}(A|Y = 0, \boldsymbol{X} = \boldsymbol{x}) = \text{expit}(-1.2 - 1.5x_1 - 1.2x_2 + 2\sin(x_1) + 2\sin(2x_2) + 2\cos(x_1 - x_2))$ and marginalize it over Y to obtain the propensity score model $\mathbb{E}(A|\boldsymbol{X} = \boldsymbol{x})$ and generate the observed treatment according to $\mathbb{E}(A|\boldsymbol{X} = \boldsymbol{x})$. (Step 5) Generate the outcome $Y \sim \text{Bernoulli}(\text{expit}(f(\boldsymbol{x}; \boldsymbol{\beta}) + \gamma(\boldsymbol{x}, a; \boldsymbol{\psi})))$. Under this data generation procedure,

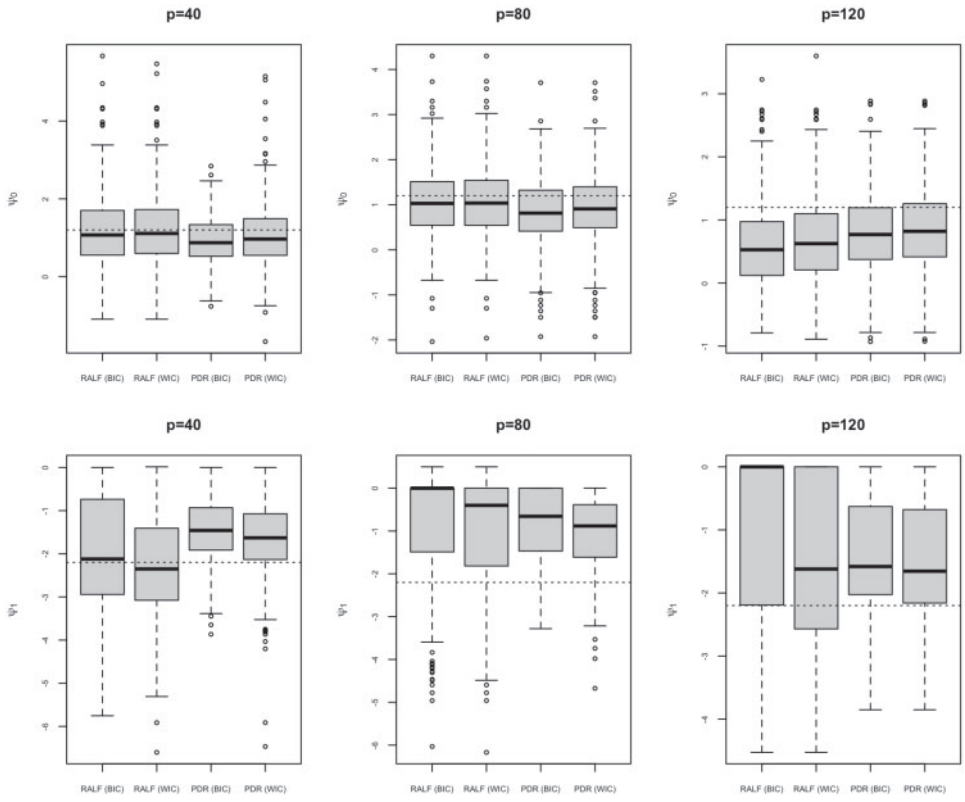


Figure 1. Estimates of blip parameters using RALF and PDR with sample size 200 (400 simulations) for $p = 40$, 80, and 120 as indicated. The true value is represented by the dashed horizontal line.

approximately 50% of patients receive treatment $A = 1$ and the marginal mean of the outcome Y is 0.51.

In all scenarios, we posited a linear baseline model. Hence, the baseline model was greatly misspecified (sin and cos functions omitted). Figure 1 and Table 1 summarize the estimators of the blip parameters ψ_0 and ψ_1, the error rate in the estimated treatment rule, the value (expected outcome under the estimated regime), and the false negative rate and false positive rate using RALF and PDR with tuning parameter selected by BIC and WIC. The value was estimated using a testing set of size 10 000.

In general, PDR outperformed RALF with respect to bias in blip parameter estimation, variable selection results, and ITR performance (error rate and value function), especially in the large p setting. This result is expected as PDR used the RALF estimator as an initial estimator. Both RALF and PDR achieved better performance when the tuning parameter was selected by WIC, as this led to lower bias in estimating blip parameters, a lower error rate, and a higher value function. Additionally, with WIC, both approaches achieved better variable selection results as demonstrated by a much lower false-negative rate at the cost of only a slight increase in the false-positive rate. This result is especially evident in the high-dimensional setting. For $p = 120$, the false-negative rate for RALF using WIC was only 13.50%, just over half the rate observed when using BIC, while the false-positive rate stayed low at 1.44%. Having a selection method with a lower false-negative rate is of particular importance in medical research, to avoid missing important sources of heterogeneity that might improve patient outcomes if considered during treatment decision-making. The best performance was attained by PDR combined with WIC.

This simulation illustrates that WIC outperforms BIC for selecting tuning parameters for ITRs. The simulation also demonstrated that when the initial estimator is obtained from RALF, PDR

Table 1. Simulation results comparing RALF and PDR for a sample size of 200 with $p = 40, 80$, and 120 using 400 iterations

		RALF (BIC)	RALF (WIC)	PDR (BIC)	PDR (WIC)
p=40					
	Value	54.06	54.46	55.32	55.28
	ER (%)	25.31	23.76	19.11	19.41
	FN (%)	11.62	7.37	5.00	3.37
	FP (%)	5.02	9.21	0.60	2.17
p=80					
	Value	51.15	51.71	52.80	53.28
	ER (%)	41.23	38.38	32.03	29.54
	FN (%)	30.25	23.12	15.62	8.25
	FP (%)	2.03	3.82	0.16	0.48
p=120					
	Value	52.06	53.88	54.70	54.94
	ER (%)	35.24	26.17	21.56	20.52
	FN (%)	25.00	13.50	8.88	5.37
	FP (%)	0.46	1.44	0.34	0.57

Notes: We report the estimated value and error rate (ER), i.e., the proportion of times the estimated optimal recommended treatment is different from the true optimal treatment. Additionally, we report the false-negative (FN) rate, i.e., the proportion of times a true tailoring variable is not selected, and false-positive (FP) rate, i.e., the proportion of times a spurious variable is selected as a tailoring variable. For comparison, the value function of the true optimal ITR is 68; the value function of the strategies "always treat" and "never treat" is 49 and 48, respectively.

performance is still excellent in the challenging scenario of a small sample size and severely misspecified baseline model.

4. INDIVIDUALIZING TREATMENT TO REDUCE THE RISK OF SEVERE DEPRESSION SYMPTOMS

Treatment individualization can be used to optimize outcomes and be applied to treatment classes in addition to specific medications. For unipolar depression, SSRIs are currently the recommended first-line treatment (Bauer *and others*, 2013), although this medication class contains several different drugs and other medication classes are also commonly used. Individuals often vary in their response to different antidepressant drug classes, and modest evidence supports using patient characteristics to tailor the choice of an antidepressant drug class to improve patient mental health outcomes (Green *and others*, 2017). Our main analysis focuses on the choice between SSRIs and SNRIs to minimize the risk of a patient having severe depression symptoms in the 1 year following medication initiation, defined as a response of 15 or greater on the PHQ, a patient-reported outcome that measures depression symptom severity with a maximum of 27, with higher scores corresponding to greater severity (Kroenke *and others*, 2001). In secondary analyses, we compare other drug classes. A brief outline on the cohort creation, covariates considered, is provided below, with further details in the Supplementary Materials.

4.1. Case study: Methods

4.1.1. Cohort construction and setting

Our study included all KPWA patients aged 13 years and older initiating antidepressant medication treatment between 2008 and 2018 who had a diagnosis of depression in the year prior to or within 15 days of treatment initiation, at least 12 months of health system enrollment prior to initiation, and no antidepressant medication fills in the 12 months prior to the initiating prescription. Some patients had multiple treatment initiation episodes over the study years meeting our criteria; for our analyses, we retained the first observed treatment episode for each patient.

4.1.2. Covariates, potential tailoring variables, outcome definitions

Demographic information was extracted from health records data on all patients. Data included age, sex (male or female), self-reported race and ethnicity, insurance type, and neighborhood educational attainment, level of poverty, and rurality score. General medication and mental health information at the time of treatment initiation was extracted from health records. We gathered information on the mental health diagnoses in the past year. We collected the number of suicide attempts and the number of psychiatric hospitalizations in the 6 months prior to treatment initiation. Baseline depression symptom severity was measured using the PHQ recorded closest to treatment initiation looking back up to 15 days and forward up to 3 days, to allow for data lags. All these covariates were considered potential tailoring variables and were used in the propensity score to account for potential confounding. We also added the calendar year of treatment initiation as a potential confounder in the treatment model.

At KPWA, the PHQ, a patient-reported outcome, is recommended for diagnosing depression and monitoring depression symptoms while on treatment. Longitudinal PHQ information and mental heath outcomes, i.e., psychiatric hospitalization and in- or outpatient visits for self-harm, were collected during the year following medication initiation. We defined the binary outcome of severe depression symptoms as a PHQ score in the health record that was greater than or equal to 15 (Kroenke *and others*, 2010) at any point in the first 12 months of follow-up, excluding the PHQ measured at treatment initiation.

4.1.3. Case study: Statistical analyses

Using the first imputed dataset, we summarized baseline characteristics of the study cohort stratified by the initiating treatment. The propensity score model was estimated using logistic regression using all confounding variables listed in Section 4.1.2. Overlap weights were computed as a function of the propensity score. The standardized mean differences (SMDs) were computed before and after weighting to assess balance across treatment groups in potential confounders at initiation.

We implemented both PDR and RALF approaches to variable selection to estimate an ITR to minimize severe depression symptoms in the year after treatment initiation. As WIC led to better performance in the simulation study, it was used as the criterion to select the tuning parameter. We report the proportion of nonzero coefficients in the blip function using PDR and RALF.

4.1.4. Addressing missing data

We used multiple imputation to replace missing values in all baseline and longitudinal variables. In addition to the covariates listed above, we used and imputed missing values for longitudinal height and weight and PHQ after the original antidepressant initiation. We used the `mice` package in **R** version 4.0.4 to create 25 completed datasets, with all missing values imputed.

For imputation, we treated the first eight items of the PHQ score (PHQ8), as a depressive symptom severity measure ranging from 0 to 24 with 24 indicating worse symptoms, and the ninth item (referred to as PHQ item 9), ranging from 0 to 3, as a measure of suicidal ideation. We used the norm option in `mice` to impute missing values in the variable PHQ8 and polytomous logistic regression to impute values in PHQ item 9. We used predictive mean matching to impute missing weight values, which generally provided better results than the norm option. For the remaining variables, we used logistic regression for binary and unordered categorical variables and proportional odds model regression for ordinal variables.

4.2. Case study: Results

Our study included 73 103 patients with 82 129 episodes of antidepressant initiation. Of the 73 103 episodes, 56 876 (78%) corresponded to initiation of an SSRI, 4056 (5.5%) to an SNRI, and the remainder to other drugs not considered in the primary analysis. We observed considerable differences in several baseline covariates between patients who initiated SSRIs or SNRIs (Table 2). The most notable differences were found in age, use of tobacco products, insurance type, neighborhood

Table 2. Baseline characteristics of the cohort of new users of antidepressants stratified by drug class before and after re-weighting data with overlap weights for confounding, KPWA, 2008–2018

	No weighting			Overlap weighting		
	Treatment			Treatment		
Covariates[†]	SNRI	SSRI	SMD	SNRI	SSRI	SMD
Age[‡]	49.1 (16.1)	43.6 (20.0)	0.30	48.8 (16.2)	48.8 (19.6)	< 0.01
Male	1215 (30.0)	18687 (32.9)	0.06	1108 (30.3)	1108 (30.3)	< 0.01
Race and ethnicity			0.13			0.06
American Indian/Alaska Native	112 (2.8)	1296 (2.3)		101 (2.8)	92 (2.5)	
Asian	131 (3.2)	3224 (5.7)		122 (3.3)	138 (3.8)	
Black/African American	147 (3.6)	2417 (4.2)		134 (3.6)	145 (3.9)	
Native Hawaiian/Pacific Islander	38 (0.9)	734 (1.3)		35 (0.9)	42 (1.1)	
Hispanic	257 (6.3)	3519 (6.2)		234 (6.4)	198 (5.4)	
White	3313 (81.7)	44947 (79.0)		2986 (81.5)	3006 (82.0)	
Other race	58 (1.4)	739 (1.3)		53 (1.4)	44 (1.2)	
Height[‡] (inches)	66.4 (3.9)	66.0 (4.1)	0.11	66.4 (3.9)	66.4 (4.0)	< 0.01
Weight[‡] (pounds)	192.6 (54.9)	180.2 (51.5)	0.23	191.5 (54.3)	191.5 (56.3)	< 0.01
Use of tobacco products[§]	591 (14.6)	6595 (11.6)	0.09	524 (14.3)	524 (14.3)	< 0.01
Charlson score[‡]	0.8 (1.5)	0.6 (1.3)	0.15	0.8 (1.5)	0.8 (1.6)	< 0.01
Rurality score[‡]	2.6 (1.2)	2.4 (1.2)	0.17	2.6 (1.2)	2.6 (1.3)	< 0.01
Insurance type			0.14			0.11
Commercial	2900 (71.5)	42633 (75.0)		2632 (71.8)	2593 (70.8)	
Medicaid	13 (0.3)	592 (1.0)		12 (0.3)	33 (0.9)	
Medicare	867 (21.4)	9743 (17.1)		766 (20.9)	840 (22.9)	
Private	276 (6.8)	3908 (6.9)		255 (6.9)	198 (5.4)	
Nbhd educational attainment	1508 (37.2)	18406 (32.4)	0.10	1349 (36.8)	1349 (36.8)	< 0.01
Nbhd poverty level	180 (4.4)	2410 (4.2)	0.01	161 (4.4)	161 (4.4)	< 0.01
Nbhd income level	201 (5.0)	2634 (4.6)	0.02	179 (4.9)	179 (4.9)	< 0.01
Anxiety disorder	1211 (29.9)	13400 (23.6)	0.14	1062 (29.0)	1062 (29.0)	< 0.01
Mental health/substance use[¶]	407 (10.0)	3572 (6.3)	0.14	346 (9.5)	346 (9.5)	< 0.01
Prior mental health inpatient stay[‡,‖]	0.0 (0.1)	0.0 (0.1)	0.02	0.0 (0.1)	0.0 (0.1)	< 0.01
Prior suicide attempts[‡,‖]	0.1 (0.3)	0.0 (0.3)	0.07	0.1 (0.3)	0.1 (0.3)	< 0.01
Prior antidepressant med use[‡,#]	0.9 (1.3)	0.4 (0.8)	0.45	0.8 (1.2)	0.8 (1.2)	< 0.01
Any prior psychotherapy visit[§]	980 (24.2)	13672 (24.0)	< 0.01	873 (23.8)	873 (23.8)	< 0.01
PHQ measured at intiation[‡‡]	1422 (35.1)	25086 (44.1)	0.19	1300 (35.5)	1300 (35.5)	< 0.01
Prior PHQ measurements[§]	0.6 (1.8)	0.4 (1.3)	0.11	0.5 (1.6)	0.5 (1.8)	< 0.01
Total score PHQ item 1–8[‡]	14.3 (5.3)	14.6 (5.1)	0.05	14.3 (5.3)	14.3 (5.2)	< 0.01
PHQ item 9[‡]	0.5 (0.9)	0.6 (0.9)	0.06	0.5 (0.9)	0.5 (0.9)	< 0.01

[†] Frequencies (%) presented, unless otherwise stated.
[‡] Mean (standard deviation). Rural score defined between 1 and 6, with 6 indicating most rural. PHQ items 1–8 total score ranges from 0 to 24, with 24 indicating most severe symptoms. PHQ item 9 ranges between 0 and 3, with 3 the highest level of recurrent suicidal ideation (nearly every day) in the past two weeks.
[§] In the year prior to cohort entry.
[¶] Includes diagnosis for autism spectrum disorder, obsessive compulsive disorder, personality disorder, post-traumatic stress disorder, alcohol-use disorder, opioid-use disorder, or sedative-use disorder.
[‖] In the 6 months prior to cohort entry.
[#] In the 5 years prior to cohort entry.
[‡‡] Within cohort entry minus 15 days and cohort entry plus 3 days. Not imputed.

education, prevalence of anxiety and other psychiatric disorders, and having a PHQ measured at treatment initiation. The overlap weights were highly effective in balancing covariates across the SSRI and SNRI initiators (Table 2). The only two variables for which a difference remained after re-weighting were race and ethnicity and insurance type, with SMDs of 0.06 and 0.11, respectively.

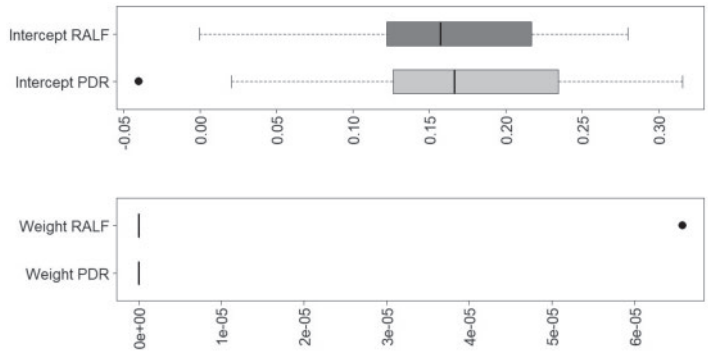


Figure 2. Distribution across the 25 imputed datasets of the coefficient estimates for the intercept and patients' weight. These were the only nonzero blip coefficients among the 23 potential effect modifiers in comparisons of SSRI and SNRI to minimize the risk of severe depression symptoms as measured by PHQ greater than 15. RALF, fixed-point regularized A-learning; PDR, penalized doubly robust. The first term of the *Y*-axis labels corresponds to the effect modifier for which nonnull coefficient(s) were found.

Over the 25 imputed datasets, 46 814 of the SSRI initiation episodes were followed by severe depressive symptoms (PHQ $\geq$ 15) in the following year, on average, for a crude risk of 82.3%. Among SSRI initiators with severe depression symptoms recorded in the following year, the mean time from treatment initiation until the first recorded severe symptoms was 4.4 months (standard deviation [SD] 2.8) and the mean time from treatment initiation until the last recorded severe symptoms was 8.4 months (SD 2.6).

Across the 25 imputed datasets, among the patients who initiated SNRIs, 3441 were followed by severe depressive symptoms, for a crude risk of 84.8%. Among patients who initiated SNRIs and had severe depression symptoms recorded in the following year, the mean time from treatment initiation until the first recorded severe symptoms was 4.1 months (SD 2.8) and the mean time from treatment initiation until the last recorded severe symptoms was 8.5 months (SD 2.6).

The RALF variable selection approach consistently retained more tailoring variables than the PDR approach. RALF consistently retained the intercept coefficient in the blip, and retained both age and weight in 2 of the 25 imputed datasets. The PDR approach, in contrast, retained the blip intercept coefficient most of the time but never retained tailoring variables for any of the 25 datasets (Figure 2). Given that none of the tailoring variables were statistically significant more than twice out of the 25 imputed datasets, the coefficients of tailoring variables were not statistically significant when combined across the 25 imputed datasets (except for the intercept, results not shown). A comparison of the optimal treatment decisions across some of the largest treatment groups, e.g., SSRI versus SNRI, revealed that the models fit under both the RALF and PDR approaches led to a 100% agreement in optimal treatment decisions; for this reason, we do not show the comparison of optimal treatment decisions across fitting methods in these instances. In general, cross-validation could be used to compare the optimal treatment decisions under both approaches in a testing set after fitting the models on a training set.

Results for the other pairwise comparisons of drug classes and the outcome of more severe symptoms similarly indicated no evidence that treatment tailoring improves outcomes (Supplementary Materials). Results were similar across all pairwise comparisons, with the PDR approach returning more conservative results and RALF retaining one or two tailoring variables in a small number of datasets. Results for other types of outcomes (severe depression outcomes, weight gain of 10% between months 1 and 12, treatment failure defined as the addition of a second antidepressant medication or the use of antipsychotic drugs, and remission defined as a baseline PHQ8 $\geq$ 10 and a 9-item PHQ $\leq$ 10 after 1 year) were similar (results not shown).

5. DISCUSSION

In this work, we combined two recent advancements, a novel computational approach (Yang *and others*, 2021) and a PDR method of estimating an optimal individualized treatment strategy, to address the challenging question of how to select tailoring variables in the context of optimizing a treatment choice for binary outcomes of interest (Tchetgen Tchetgen *and others*, 2010; Bian *and others*, 2022). This methodology was deployed to investigate the pressing question of whether treatment response heterogeneity in unipolar depression may be explained by demographic and clinical characteristics of individuals starting antidepressant therapy.

Given the large KPWA sample and the many comparisons we explored, our findings do not support tailoring initial antidepressant therapy by medication class to minimize the risk of severe depression symptoms. Tailoring at the level of specific antidepressant medications may be useful, but tailoring by medication class did not lead to better management of severe symptoms. Moreover, exploratory analyses of outcomes including severe events (hospitalizations and suicide attempts) and side effects (significant weight gain) failed to show any evidence of benefit to treatment tailoring.

The approach we used relies on several assumptions to ensure consistent estimators of the ITR (blip) parameters. Perhaps most fundamentally, the results of our data analysis depend on the assumptions we have made about the missing data. Up to 90% of PHQ measurements were missing for some months and had to be imputed. We addressed that by including many variables in the imputation models including all the covariates, PHQ measurements, and potential tailoring variables considered in the analytic model. We must also assume that all potential confounders have been measured without error. We included several measures of sociodemographic factors, physical and mental health and, importantly, depression symptom severity at treatment initiation. Thus, the assumption of no unmeasured confounders is strong but plausible, given our knowledge of KPWA prescribing patterns and the low imbalance observed between the two treatment groups in our main analysis, we are confident our analyses were not dramatically impacted by unmeasured confounding. We also assumed positivity. The medication classes considered in our primary analysis, SSRIs and SNRIs, are prescribed fairly interchangeably for depression treatment, making it likely that treatment positivity holds. Additionally, no estimated propensity scores of 0 or 1, or close to these extreme values were observed.

Finally, we made no attempt to perform variable selection or complex model fitting in the propensity score. All variables included in the propensity score are known to be related to mental health and symptoms of depression. Given that the number of potential confounders is small relative to the sample size, there was no need to perform selection. We acknowledge that high-dimensional settings may require carefully selecting confounders for inclusion in the propensity score. This can be done in a variety of ways, for example, using outcome-adaptive LASSO (Shortreed and Ertefaie, 2017), an approach that has been successfully implemented in the context of ITRs (Bian *and others*, 2023). Other machine-learning methods have also been suggested as approaches to estimating the propensity score, however they are not without controversy (Alam *and others*, 2019).

In this work, we focused on a setting with only two treatment alternatives. It would be straightforward to extend these methods to allow for multiple treatments. A multinomial logistic model rather than binomial logistic regression can be used to fit the nuisance treatment model; although, as in the algorithm outlined in Section 2.1, the treatment model used conditions on those for whom the outcome takes value 0 (i.e., it is not a typical "propensity score" across the whole population). A generalized propensity score for categorical treatments has previously been incorporated into a related, weighted regression framework to estimating tailored treatment strategies (Schulz and Moodie, 2021), although to date it has only been implemented in a continuous outcome setting. However, there is no conceptual or theoretical barrier to using a similar approach with suitable balancing weights in the binary outcome estimating function $\boldsymbol{U}_2(\boldsymbol{\beta}, \boldsymbol{\psi})$. Further, Bian *and others* (2022) demonstrated a doubly robust estimation procedure with selection (not including the fixed-point

approach described here) for count as well as binary outcomes. Extending the current work to accommodate multi-valued (count) outcomes also would likely be straightforward.

An important but challenging consideration for future work is the extension of the proposed estimation approach to treatment sequences, i.e., estimating adaptive treatment strategies. This methodology is particularly difficult in the context of a binary outcome, as the typical backward-induction approach to estimation requires using pseudo-outcomes computed by "scaling up" the observed outcome by the estimated blip function for individuals whose observed later-stage treatment is not concordant with the estimated optimal treatment. However, binary outcomes have no obvious, valid ways to scale up a zero-outcome to create a meaningful pseudo-outcome for recursive estimation, as this scaling is done on the multiplicative scale; and thus, zero outcomes remain unchanged. This problem has been discussed previously—though no formal or theoretically validated solution has been proposed—by Wallace *and others* (2019). A possible alternative may be to pursue a dynamic marginal structural-based approach to estimation (Murphy *and others*, 2001; van der Laan and Petersen, 2007; Orellana *and others*, 2010; Shortreed and Moodie, 2012) for the longitudinal treatment sequence context rather than the regression-based backward induction used in methods such as dWOLS and g-estimation. Thus, although extending the estimation approach to treatment sequences has statistical and computation hurdles that are not trivial, overcoming these hurdles is critical for providing better, evidence-backed support for treatment decision-making. Given that this study did not find evidence that tailoring initial antidepressant therapies by medication class reduces risk of severe depression symptoms, the ability to estimate the optimal second-line and subsequent treatments for depression becomes more important. The approach in this study provides a foundation for investigating if factors such as timing and reasons for treatment failure (e.g., side effects, lack of therapeutic effect) are useful for guiding decisions about the sequence of depression treatment.

SOFTWARE

Software in the form of R code is available online at https://github.com/ZeyuBian/Biostatistics_Moodie_2023.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R01 MH114873. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. E.E.M.M. is a Canada Research Chair (Tier 1) and acknowledges the support of a chercheur de mérite career award from the Fonds de Recherche du Québec, Santé.

SUPPLEMENTARY MATERIAL

Supplementary material is available at http://biostatistics.oxfordjournals.org.

CONFLICT OF INTEREST

S.M.S. has worked on grants awarded to Kaiser Permanente Washington Health Research Institute (KPWHRI) by Bristol Meyers Squibb and by Pfizer. She was also a co-Investigator on grants awarded to KPWHRI from Syneos Health, which represented a consortium of pharmaceutical companies carrying out U.S. Food and Drug Administration-mandated studies on the safety of extended-release opioids. The study protocol was approved by KPWA's Institutional Review board.

REFERENCES

Alam, S., Moodie, E. E. M. and Stephens, D. A. (2019). Should a propensity score model be super? The utility of ensemble procedures for causal adjustment. *Journal of the American Statistical Association* **38**, 1690–1702.

Anderson, D. G. (1965). Iterative procedures for nonlinear integral equations. *Journal of the ACM* **12**, 547–560.

Bauer, M., Pfennig, A., Severus, E., Whybrow, P. C., Angst, J., and Möller, H.; and World Federation of Societies of Biological Psychiatry Task Force on Unipolar Depressive Disorders. (2013). World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, Part 1: Update 2013 on the acute and continuation treatment of unipolar depressive disorders. *The World Journal of Biological Psychiatry* **14**, 334–385.

Bian, Z., Moodie, E. E. M., Shortreed, S. M. and Bhatnagar, S. (2021). Variable selection in regression-based estimation of dynamic treatment regimes. *Biometrics*.

Bian, Z., Moodie, E. E. M., Shortreed, S. M., Lambert, S. and Bhatnagar, S. (2023). Tailoring variable selection and ranking for optimal treatment decisions. In: Laber, E., Chakraborty, B., Moodie, E. E. M., Cai, T. and van der Laan, M. (editors), *Handbook of Statistical Methods for Precision Medicine*. CRC Press.

Bian, Z., Moodie, E. E. M., Shortreed, S. M., Lambert, S. D. and Bhatnagar, S. (2022). Variable selection for individualized treatment rules with discrete outcomes. *arXiv preprint arXiv:2205.13609* **1**, 1–32.

Chakraborty, B. (2011). Dynamic treatment regimes for managing chronic health conditions: A statistical perspective. *American Journal of Public Health* **101**, 40–45.

Chakraborty, B. and Moodie, E. E. M. (2013). *Statistical Methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine*. New York, NY: Springer-Verlag.

Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, R. (1987). Charlson comorbidity index. *Journal of Crohn's and Colitis* **40**, 373–383.

Clarke, F. H. (1990). *Optimization and Nonsmooth Analysis*. SIAM.

Coulombe, J., Moodie, E. E. M., Shortreed, S. M. and Renoux, C. (2021). Can the risk of severe depression-related outcomes be reduced by tailoring the antidepressant therapy to patient characteristics? *American Journal of Epidemiology* **190**, 1210–1219.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association* **96**, 1348–1360.

Fan, J. and Lv, J. (2011). Nonconcave penalized likelihood with NP-dimensionality. *IEEE Transactions on Information Theory* **57**, 5467–5484.

Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *The Annals of Statistics* **32**, 928–961.

Fan, Y. and Tang, C. Y. (2013). Tuning parameter selection in high dimensional penalized likelihood. *Journal of the Royal Statistical Society, Series B* **75**, 531–552.

Fang, H. and Saad, Y. (2009). Two classes of multisecant methods for nonlinear acceleration. *Numerical Linear Algebra with Applications* **16**, 197–221.

Fu, W. J. (2003). Penalized estimating equations. *Biometrics* **59**, 126–132.

Goldberg, Y., Song, R. and Kosorok, M. R. (2013). Adaptive Q-learning. In: Banerjee, M., Huang, J., Koltchinskii, V., Bunea, F. and Maathuis, M. H. (editors), *From Probability to Statistics and Back: High-Dimensional Models and Processes—A Festschrift in Honor of Jon Wellner*. Beachwood, OH: Institute of Mathematical Statistics, pp. 150–162.

Green, E., Goldstein-Piekarski, A. N., Schatzberg, A. F., Rush, A. J., Ma, J. and Williams, L. (2017). Personalizing antidepressant choice by sex, body mass index, and symptom profile: An iSPOT-D report. *Personalized Medicine in Psychiatry* **1**, 65–73.

Jeng, X. J., Lu, W. and Peng, H. (2018). High-dimensional inference for personalized treatment decision. *Electronic Journal of Statistics* **12**, 2074.

Johnson, B. A., Lin, D. Y. and Zeng, D. (2008). Penalized estimating functions and variable selection in semiparametric regression models. *Journal of the American Statistical Association* **103**, 672–680.

Kroenke, K., Spitzer, R. L. and Williams, J. B. W. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine* **16**, 606–613.

Kroenke, K., Spitzer, R. L., Williams, J. B. and Lowe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *General Hospital Psychiatry* **32**, 345–359.

Lian, Y. (2022). Some new computational methods in high-dimensional statistical learning in biostatistics, [PhD. Thesis]. McGill University.

Lu, W., Zhang, H. H. and Zeng, D. (2013). Variable selection for optimal treatment decision. *Statistical Methods in Medical Research* **96**, 1348–1360.

Murphy, S. A. (2003). Optimal dynamic treatment regimes (with discussion). *Journal of the Royal Statistical Society, Series B* **65**, 331–366.

Murphy, S. A., van der Laan, M. J. and Robins, J. M.; Conduct Problems Prevention Research Group. (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association* **96**, 1410–1423.

Orellana, L., Rotnitzky, A. and Robins, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: main content. *The International Journal of Biostatistics* **6**.

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

Robins, J. M. (2000). Robust estimation in sequentially ignorable missing data and causal inference models. In: *Proceedings of the American Statistical Association*, Vol. 1999. pp. 6–10.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In: Lin, D. Y. and Heagerty, P. (editors), *Proceedings of the Second Seattle Symposium on Biostatistics*. New York: Springer, pp. 189–326.

Schulz, J. and Moodie, E. E. M. (2021). Doubly robust estimation of optimal dosing strategies. *Journal of the American Statistical Association* **116**, 256–268.

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

Shortreed, S. M. and Ertefaie, A. (2017). Outcome-adaptive lasso: Variable selection for causal inference. *Biometrics* **73**, 1111–1122.

Shortreed, S. M. and Moodie, E. E. E. M. (2012). Estimating the optimal dynamic antipsychotic treatment regime: Evidence from the sequential multiple-assignment randomized clinical antipsychotic trials of intervention and effectiveness schizophrenia study. *Journal of the Royal Statistical Society, Series C* **61**, 577–599.

Simon, G. (2001). Choosing a first-line antidepressant: Equal on average does not mean equal for everyone. *Journal of the American Medical Association* **286**, 3003–3004.

Song, R., Wang, W., Zeng, D. and Kosorok, M. R. (2015). Penalized Q-learning for dynamic treatment regimens. *Statistica Sinica* **25**, 901–920.

Tchetgen Tchetgen, E. J., Robins, J. M. and Rotnitzky, A. (2010). On doubly robust estimation in a semiparametric odds ratio model. *Biometrika* **97**, 171–180.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B* **58**, 267–288.

van der Laan, M. J. and Petersen, M. L. (2007). Causal effect models for realistic individualized treatment and intention to treat rules. *The International Journal of Biostatistics* **3**.

Wallace, M. P. and Moodie, E. E. M. (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* **71**, 636–644.

Wallace, M. P., Moodie, E. E. M. and Stephens, D. A. (2019). Model selection for G-estimation of dynamic treatment regimes. *Biometrics* **75**, 1205–1215.

Wang, L., Zhou, J. and Qu, A. (2012). Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics* **68**, 353–360.

Wu, J., Galanter, N., Shortreed, S. M. and Moodie, E. E. M. (2022). Ranking tailoring variables for constructing individualized treatment rules: An application to schizophrenia. *Journal of the Royal Statistical Society, Series C* **71**, 309–330.

Yang, A. Y., Zhao, Y., Gu, Y., Lian, Y. and Fan, J. (2021). Regularized estimating equations: some new perspectives. *arXiv* doi:10.48550/ARXIV.2110.11074.

Zetterqvist, J. and Sjölander, A. (2015). Doubly robust estimation with the `R` package `drgee`. *Epidemiologic Methods* **4**, 69–86.

Zhang, B. and Zhang, M. (2018). C-learning: A new classification framework to estimate optimal dynamic treatment regimes. *Biometrics* **74**, 891–899.

Zhang, J., O'Donoghue, B. and Boyd, S. (2020). Globally convergent Type-I Anderson acceleration for nonsmooth fixed-point iterations. *SIAM Journal on Optimization* **30**, 3170–3197.

APPENDIX

Supplementary Materials to "Variable selection in high dimensions for discrete-outcome individualized treatment rules: A case study in reducing severity of depression symptoms"

by

EEM MOODIE, Z BIAN, J COULOMBE, Y LIAN, AY YANG, and SM SHORTREED

These Supplementary Materials contain details of the missing data procedures and results of secondary analyses including additional tailored treatment analyses that compare results of initiation with the antidepressant medication classes SSRIs and SNRIs, MOIs, mirtazapine, TCAs, and bupropion. These materials also report results on secondary analyses defining severe depression symptoms as a PHQ score of 20 or greater.

A. Additional Computational Details

As an alternative to the use of the Clarke subdifferential of the SCAD penalty as in Equation (2.5), we can make a local quadratic approximation (LQA, Fan and Li, 2001; Johnson *and others*, 2008) or a local linear approximation (LLA, Zou and Li, 2008) to the penalty. As LLA has been shown to be computationally superior to LQA (Zou and Li, 2008), we only show how the LLA approximation can be incorporated into our method. The LLA approximates the penalty using

$$\rho_\lambda(\psi_j) \approx \rho_\lambda^{\mathrm{LLA}}(\psi_j)\rho_\lambda(\tilde{\psi}_j) + \rho'_\lambda(\tilde{\psi}_j)(|\psi_j| - |\tilde{\psi}_j|),$$

where $\tilde{\psi}_j$ is an initial estimate satisfying certain conditions (Zou and Li, 2008); this leads to the following fixed-point problem with a weighted LASSO soft-thresholding operator

$$\widehat{\boldsymbol{\psi}}^{\mathrm{RALF}} = f(\widehat{\boldsymbol{\psi}}^{\mathrm{RALF}}), \text{ where } f(\boldsymbol{\psi}) = S_{\tau\rho'_\lambda(\tilde{\psi}_j)\lambda}(\boldsymbol{\psi} - \tau \boldsymbol{U}_1(\boldsymbol{\psi})). \tag{A.1}$$

Accordingly, the formulations of the REE in Equations (2.6) and (A.1) can both be very efficiently solved by a fixed-point algorithm (Yang *and others*, 2021). Since these two approaches (exact and approximation) are both computationally efficient and statistically similar, we focus on the exact approach throughout this paper. Following Zhang *and others* (2020) and Lian (2022), we incorporate a Type-I AA technique during the iterative update to accelerate the computation. Specifically, the update at the k-th iteration has the form

$$\boldsymbol{\psi}^{(k+1)} = \sum_{r=0}^{m_k} \alpha_r^{(k)} f\left(\boldsymbol{\psi}^{(k-m_k+r)}\right),$$

where $m_k = \min(m, k)$ for some maximum memory size, $m > 0$, and coefficients $\boldsymbol{\alpha}^{(k)} = (\alpha_0^{(k)}, \alpha_1^{(k)}, \ldots, \alpha_{m_k}^{(k)})^\top$ sum to one. A memory size, m, ranging from 2 to 50 is found to be a reasonable choice. In this paper, we choose $m = 10$ as suggested by Zhang *and others* (2020), which shows that the influence of m is mostly on the convergence rate rather than solutions. The coefficients are designed to minimize the residuals $g(\boldsymbol{\psi}) \equiv \boldsymbol{\psi} - f(\boldsymbol{\psi})$ of the previous m_k iterations,

$$\begin{aligned} &\underset{\boldsymbol{\alpha}^{(k)}, \mathbf{1}^\top \boldsymbol{\alpha}^{(k)}=1}{\arg\min} \left\| \sum_{r=0}^{m_k} \alpha_r^{(k)} \left(\psi^{(k-m_k+r)} - f(\psi^{(k-m_k+r)}) \right) \right\|_2, \\ = &\underset{\boldsymbol{\alpha}^{(k)}, \mathbf{1}^\top \boldsymbol{\alpha}^{(k)}=1}{\arg\min} \left\| \sum_{r=0}^{m_k} \alpha_r^{(k)} \boldsymbol{g}^{(k-m_k+r)} \right\|_2. \end{aligned}$$

Zhang *and others* (2020) showed that the optimal coefficients can be acquired in closed form so the AA update is

$$\boldsymbol{\psi}^{(k+1)} = \boldsymbol{\psi}^{(k)} - \left(I + (\boldsymbol{S}_k - \boldsymbol{V}_k)(\boldsymbol{S}_k^\top \boldsymbol{V}_k)^{-1} \boldsymbol{S}_k^\top\right) \boldsymbol{g}^{(k)}, \tag{A.2}$$

where

$$\begin{aligned} \boldsymbol{V}_k &= [(\boldsymbol{g}^{(k-m_k+1)} - \boldsymbol{g}^{(k-m_k)}), \ldots, (\boldsymbol{g}^{(k)} - \boldsymbol{g}^{(k-1)})], \\ \boldsymbol{S}_k &= [(\boldsymbol{\psi}^{(k-m_k+1)} - \boldsymbol{\psi}^{(k-m_k)}), \ldots, (\boldsymbol{\psi}^{(k)} - \boldsymbol{\psi}^{(k-1)})]. \end{aligned} \tag{A.3}$$

We see that update Equation (A.2) avoids expensive computation of the estimating function gradient and its inverse and requires only computation of the estimating function itself, which significantly reduces the computation cost per iteration, especially in the high-dimensional case. Overall, RALF is built on a fixed-point based algorithm for solving general REEs (Yang *and*

others, 2021) that was shown to provide higher estimation accuracy, computational efficiency, and scalability (Yang *and others*, 2021; Lian, 2022) over existing REE algorithms (Fu, 2003; Johnson *and others*, 2008; Wang *and others*, 2012). As discussed in Section 2.1, when the sample size is small, it may be difficult to obtain a well-behaved initial estimator for PDR. If a well-behaved estimator cannot be identified, performance may be poor. However, with the help of RALF, a reasonable initial estimator can be obtained. Specifically, after $\widehat{\boldsymbol{\psi}}^{RALF}$ is estimated as in Expression (2.4), it can be plugged into the PDR minimization objective function in Expression (2.3) to further obtain the blip estimator.

B. Additional information on the KPWA data

Cohort construction and setting

KPWA provides care and coverage to patients in Washington state, U.S.A. Kaiser Permanente Washington Heath Research Institute uses a virtual data warehouse created for research that brings together electronic health records and insurance billing information, including demographics, prescription fills, patient-reported outcomes, and health care utilization including for serious outcomes such as hospitalizations and deaths. The KPWA Institutional Review Board approved waivers of consent for use of records data in this research.

In addition to the inclusion criteria listed in the main paper, cohort construction excluded some individuals. Individuals were excluded if they had a diagnosis of personality, bipolar, or psychotic disorder in the year prior to treatment initiation. Also excluded from analyses were individuals who initiated treatment with more than one antidepressant medication, determined by extracting data on all antidepressant medications that were filled at the time of medication initiation.

Covariates, potential tailoring variables, outcome definitions

Demographic information was extracted from health records data on all patients. Age in years at treatment initiation was calculated from date of birth. At the time of the data pull, information on patient sex (male or female) most likely represented sex assigned at birth. Health records data on race and ethnicity information are usually self-reported at the time of the first outpatient medical appointment within the health system. Race and ethnicity information was combined to categorize all individuals who self-reported Hispanic ethnicity, while all other individuals were classified using the following race categories: American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian/Pacific Islander, or White. Additional demographic information included insurance type (commercial, Medicaid, Medicare, or private) and information obtained from patient addresses and the 2010 Census, including neighborhood educational attainment (less than 25% college degrees), income (median lower than 40,000 USD), and level of poverty (20% of households below federal poverty level). We also scored if patients lived in an urban or rural area (1 to 6, with 1 the most urban and 6 the most rural). General medication and mental health information at the time of treatment initiation was extracted from health records, including the Charlson score, a general measure of comorbidity (Charlson *and others*, 1987), and tobacco use in the year prior. Height in inches and weight in pounds were collected from the visit closest in time to treatment initiation, looking back up to 5 years for height information and up to 2 years for weight information. We gathered information on the following mental health diagnosis in the past year: anxiety, alcohol use disorder, autism spectrum disorder, obsessive compulsive disorder, opioid use disorder, personality disorder, post-traumatic stress disorder, and sedative use disorder. In this population, mental health conditions other than anxiety were very rare, so all diagnoses other than anxiety were combined into a single indicator of a mental health or substance use disorder. We collected the number of suicide attempts and the number of psychiatric hospitalizations in the 6 months prior to treatment initiation. Additionally, we collected the number

of different antidepressants taken in the 5 years prior, if the patient had received psychotherapy in the 5 years prior, and the number of PHQ measurements recorded in the medical record in the year prior to treatment initiation. Baseline depression symptom severity was measured using the PHQ recorded closest to treatment initiation looking back up to 15 days and forward up to 3 days, to allow for data lags. All these covariates were considered potential tailoring variables and were used in the propensity score to account for potential confounding. We also added the calendar year of treatment initiation as a potential confounder in the treatment model.

Covariates considered in secondary analyses

In secondary analyses, we extended the study to other pairwise comparisons of treatments across SSRI, SNRI, monoamine oxidase inhibitors (MOIs), mirtazapine, tricyclic antidepressants (TCA), and bupropion, always with the aim of reducing the risk of severe depression symptoms. Of the 73,103 episodes, 56,876 (78%) corresponded to initiation of an SSRI, 4,056 (5.5%) to an SNRI; 22 (<1%) to an MOI; 1,747 (2%) to mirtazapine; 2,011 (3%) to a TCA; 8,330 (11%) to bupropion; and 61 (<1%) corresponding to initiation of an antidepressant in a medication class not included in any analyses

In secondary analyses, we also considered more severe symptoms using the threshold of PHQ greater than or equal to 20, along with other outcomes associated with severe depression, including self-harm, hospitalization for depression, and treatment failure, as well as remission of depression symptoms and the potential side effect of weight gain.

References

Charlson, ME, Pompei, P, Ales, KL and MacKenzie, R. (1987). Charlson comorbidity index. *Journal of Crohn's and Colitis* **40**(5), 373–383.

Fan, Jianqing and Li, Runze. (2001). Variable selection via nonconcave penalized likelihood

and its oracle properties. *Journal of the American statistical Association* **96**(456), 1348–1360.

Fu, Wenjiang J. (2003). Penalized estimating equations. *Biometrics* **59**(1), 126–132.

Johnson, Brent A, Lin, DY and Zeng, Donglin. (2008). Penalized estimating functions and variable selection in semiparametric regression models. *Journal of the American Statistical Association* **103**(482), 672–680.

Lian, Yi. (2022). Some new computational methods in high-dimensional statistical learning in biostatistics [Ph.D. Thesis]. McGill University.

Wang, Lan, Zhou, Jianhui and Qu, Annie. (2012). Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics* **68**(2), 353–360.

Yang, Archer Yi, Zhao, Yu, Gu, Yuwen, Lian, Yi and Fan, Jun. (2021). Regularized estimating equations: Some new perspectives. *arXiv* **doi:10.48550/ARXIV.2110.11074**.

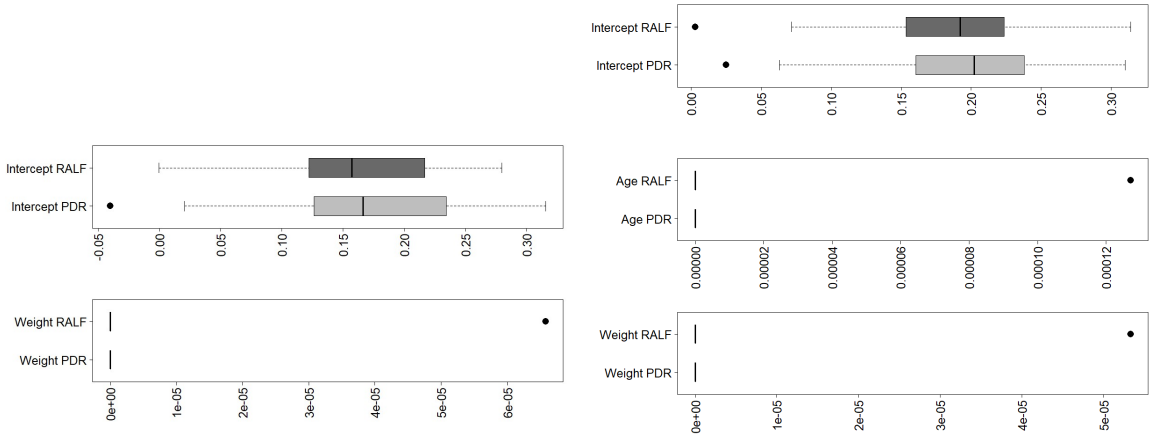
Zhang, Junzi, O'Donoghue, Brendan and Boyd, Stephen. (2020). Globally convergent type-I anderson acceleration for nonsmooth fixed-point iterations. *SIAM Journal on Optimization* **30**(4), 3170–3197.

Zou, Hui and Li, Runze. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *The Annals of Statistics* **36**(4), 1509 – 1533.

C. Additional results from the KPWA analysis

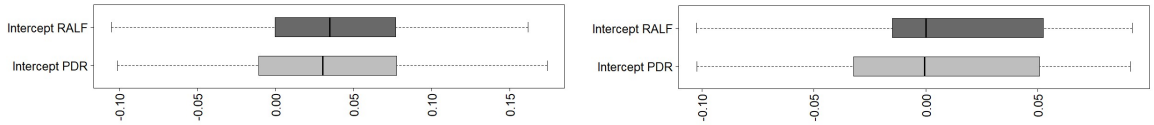
Table S1: Crude risk of severe depression symptoms, average time to the first severe depression symptoms observed (in months), and average time to the last severe depression symptoms observed (in months) by drug class in the first year of follow-up, computed using Rubin's rule, KPWA health data, 2008-2018

Drug class	Crude risk PHQ $\geqslant 15$	Crude risk PHQ $\geqslant 20$	Time to PHQ $\geqslant 15$ (SD)	Time to PHQ $\geqslant 20$ (SD)	Time to the last PHQ $\geqslant 15$ (SD)	Time to the last PHQ $\geqslant 20$ (SD)
SSRI	82.3 %	43.4 %	4.4 (2.8)	5.7 (3.0)	8.4 (2.6)	7.6 (2.8)
SNRI	84.8 %	49.3 %	4.1 (2.8)	5.4 (3.0)	8.5 (2.6)	7.6 (2.9)
Mirtazapine	77.0 %	39.6 %	4.3 (2.8)	5.5 (2.9)	8.1 (2.7)	7.3 (2.8)
Bupropion	82.3 %	41.8 %	4.5 (2.9)	5.9 (3.0)	8.5 (2.6)	7.7 (2.8)
TCA	84.8 %	45.7 %	4.7 (2.9)	6.2 (2.9)	8.8 (2.4)	8.1 (2.6)
MOI	66.9 %	26.1 %	5.2 (3.1)	6.4 (2.8)	7.9 (2.9)	7.7 (2.8)



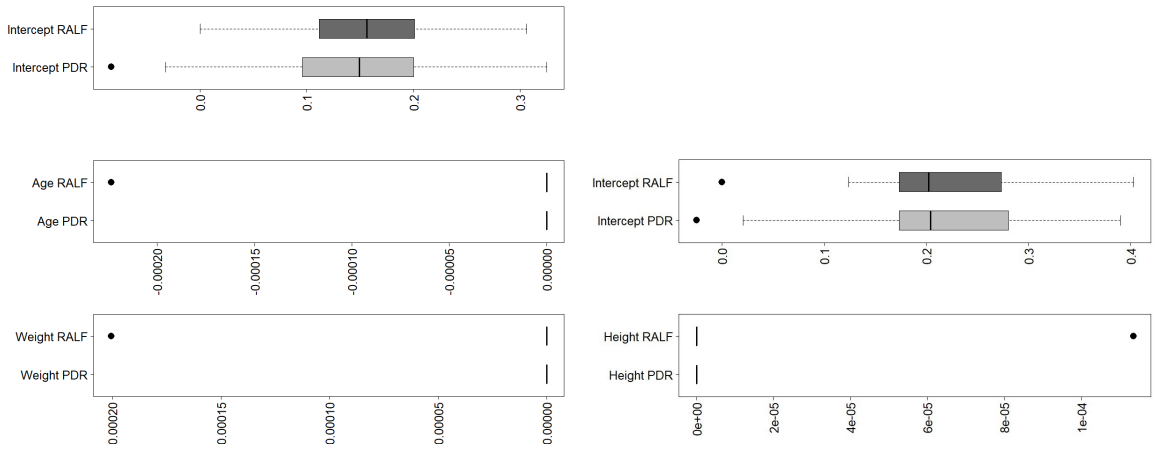
(a) (b)

Fig. S1: Distribution across the 25 imputed datasets of the coefficients on age and weight, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SSRI and SNRI** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found).



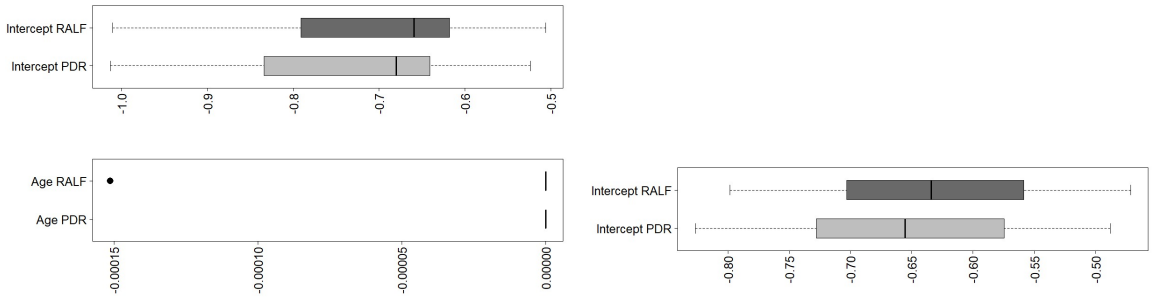
(a) (b)

Fig. S2: Distribution across the 25 imputed datasets of the intercept coefficients, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SSRI and bupropion** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). Here, the intercept is positive, meaning that the average treatment effect of SSRI is preferable to that of bupropion and that SSRI would always be the recommended treatment.



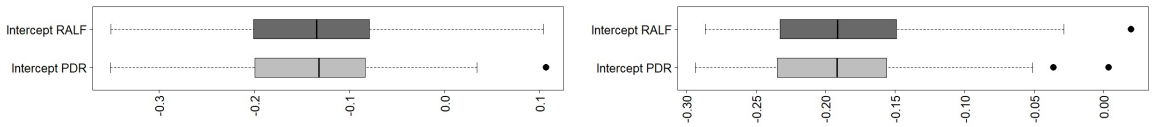
(a) (b)

Fig. S3: Distribution across the 25 imputed datasets of the coefficients on age, weight and height, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SSRI and mirtazapine** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found).



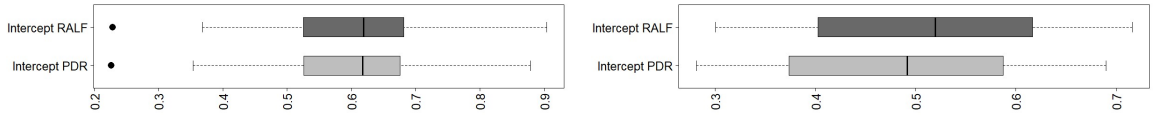
(a) (b)

Fig. S4: Distribution across the 25 imputed datasets of the coefficients on age, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **bupropion and TCA** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). The distribution in b) contains only 19 coefficients due to lack of convergence in 6 of the imputed datasets.



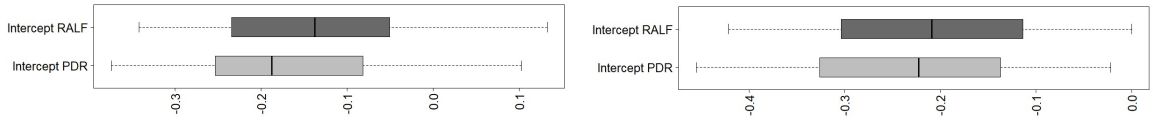
(a) (b)

Fig. S5: Distribution across the 25 imputed datasets of the intercept coefficients, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SNRI and bupropion** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). Here, the intercept is negative, meaning that the average treatment effect of bupropion is preferable to that of SNRI and that bupropion would always be the recommended treatment.



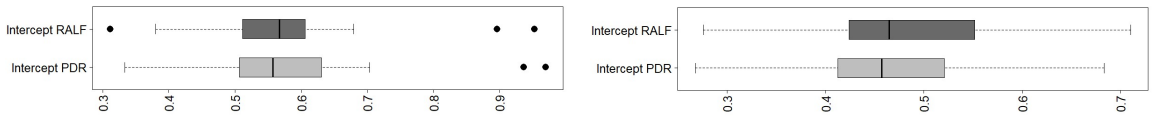
(a) (b)

Fig. S6: Distribution across the 25 imputed datasets of the intercept coefficients, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SNRI and TCA** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). The distribution in b) contains only 20 coefficients due to lack of convergence in 5 of the imputed datasets. Here, the intercept is positive, meaning that the average treatment effect of SNRI is preferable to that of TCA and that SNRI would always be the recommended treatment.



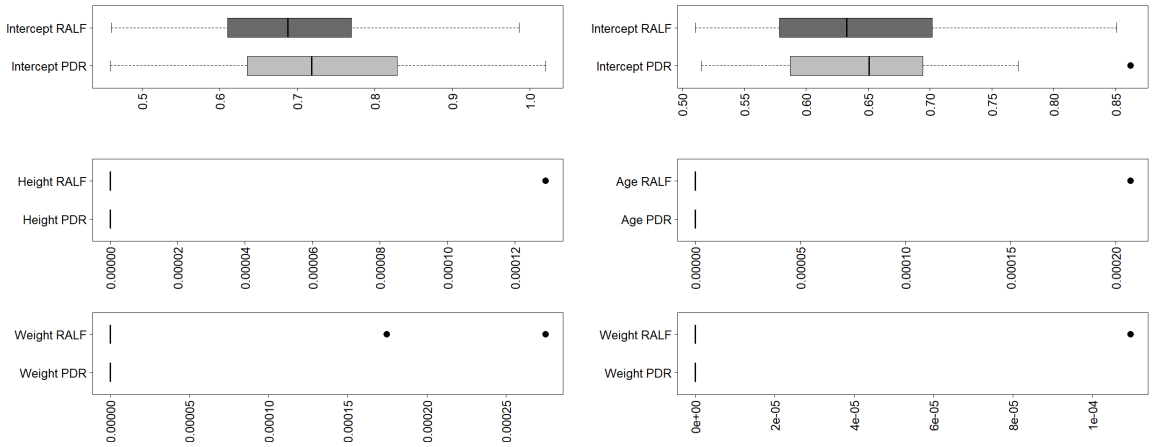
(a) (b)

Fig. S7: Distribution across the 25 imputed datasets of the intercept coefficients, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **mirtazapine and bupropion** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). Here, the intercept is negative, meaning that the average treatment effect of bupropion is preferable to that of mirtazapine and that bupropion would always be the recommended treatment.



(a) (b)

Fig. S8: Distribution across the 25 imputed datasets of the intercept coefficients, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **mirtazapine and TCA** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). The distribution in b) contains only 19 coefficients due to lack of convergence in 6 of the imputed datasets. Here, the intercept is positive, meaning that the average treatment effect of mirtazapine is preferable to that of TCA and that mirtazapine would always be the recommended treatment.



(a) (b)

Fig. S9: Distribution across the 25 imputed datasets of the coefficients on age, height and weight, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SSRI and TCA** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). The distribution in b) contains only 20 coefficients due to lack of convergence in 5 of the imputed datasets.

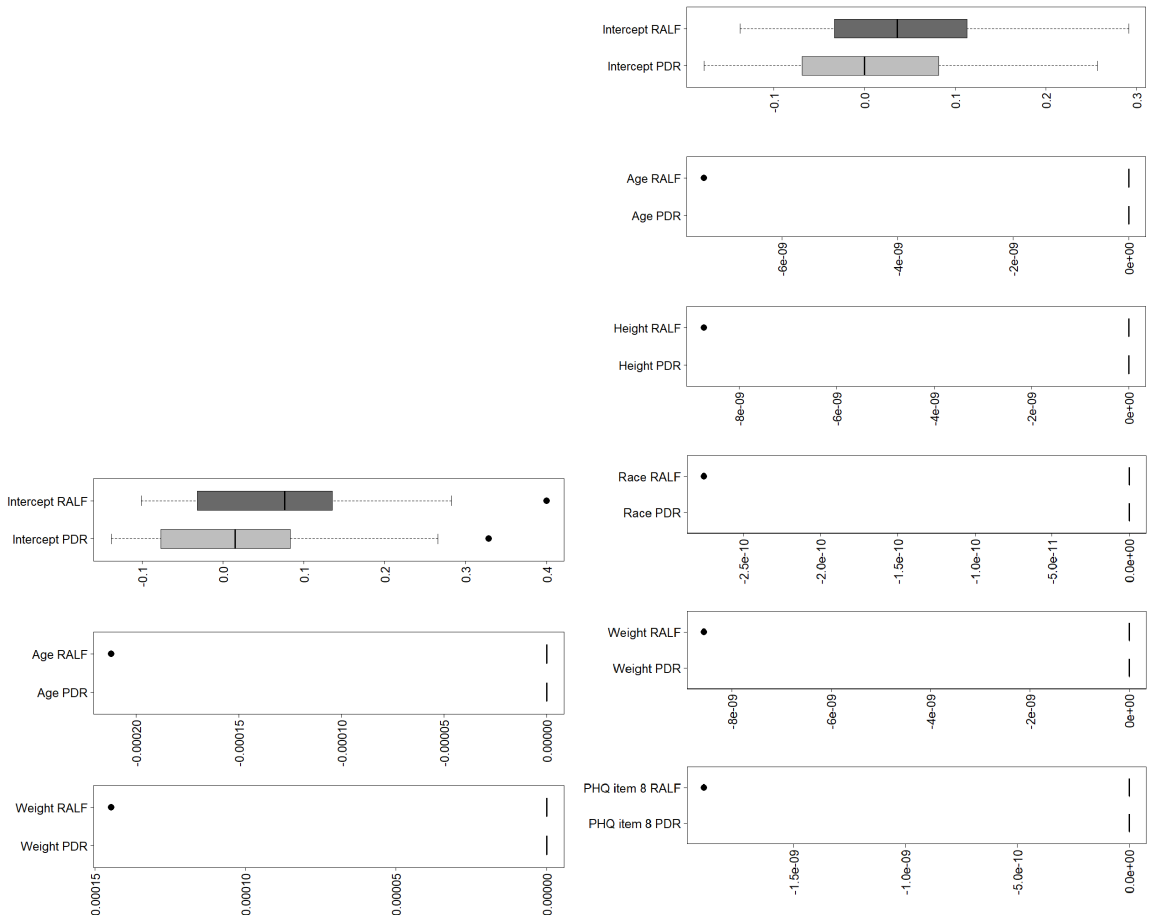


Fig. S10: Distribution across the 25 imputed datasets of the coefficients on age, race, height, weight and PHQ8, the only nonzero blip coefficients among the 23 potential effect modifiers in the comparison of **SNRI and mirtazapine** to minimize the risk of **a) a PHQ greater than 15**; and **b) a PHQ greater than 20** (RALF, fixed-point regularized A-learning; PDR, penalized doubly robust; the first term in the labels on the Y-axis corresponds to the effect modifier for which non-null coefficient(s) were found). The distribution in b) contains only 24 coefficients due to lack of convergence in 1 of the imputed datasets.