Received: 22 December 2017 | Revised: 13 July 2019 | Accepted: 18 July 2019

DOI: 10.1002/sim.8342

RESEARCH ARTICLE

WILEY Statistics in Medicine

Estimating the quality of optimal treatment regimes

Aniek Sies | **Iven Van Mechelen**

Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Correspondence
Aniek Sies, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, box 3713, 3000 Leuven, Belgium.
Email: aniek.sies@kuleuven.be

Funding information
Research Foundation-Flanders, Grant/Award Number: G080219N; Research Fund of KU Leuven, Grant/Award Number: GOA/15/003

When multiple treatment alternatives are available for a disease, an obvious question is which alternative is most effective for which patient. One may address this question by searching for optimal treatment regimes that specify for each individual the preferable treatment alternative based on that individual's baseline characteristics. When such a regime has been estimated, its quality (in terms of the expected outcome if it was used for treatment assignment of all patients in the population under study) is of obvious interest. Obtaining a good and reliable estimate of this quantity is a key challenge for which so far no satisfactory solution is available. In this paper, we consider for this purpose several estimators of the expected outcome in conjunction with several resampling methods. The latter have been evaluated before within the context of statistical learning to estimate the prediction error of estimated prediction rules. Yet, the results of these evaluations were equivocal, with different best performing methods in different studies, and with near-zero and even negative correlations between true and estimated prediction errors. Moreover, for different reasons, it is not straightforward to extrapolate the findings of these studies to the context of optimal treatment regimes. To address these issues, we set up a new and comprehensive simulation study. In this study, combinations of different estimators with $.632^+$ and out-of-bag bootstrap resampling methods performed best. In addition, the study shed a surprising new light on the previously reported problematic correlations between true and estimated prediction errors in the area of statistical learning.

KEYWORDS

bootstrap, cross-validation, optimal treatment regime, resampling, statistical learning, treatment-subgroup interactions

1 | INTRODUCTION

Often, multiple treatment alternatives are available for a certain medical or psychological problem. When this is the case, an obvious question is which of these treatment alternatives is most effective. One way to address this question is by setting up a comparative treatment evaluation study and using the resulting data to find the overall most effective treatment alternative (ie, the one with the highest expected outcome across all patients under study).

However, the best treatment alternative often differs between patients due to treatment effect heterogeneity.[1-3] Hence, there is a clear need for personalized treatment assignment,[4,5] which can be accomplished by looking for a decision rule (ie, a treatment regime) that specifies the treatment alternative that patients should preferably receive given their

 © 2019 John Wiley & Sons, Ltd.

pretreatment characteristics. Treatment regimes can take different forms, on the basis of which they can be organized in different classes. Examples include the class of regimes based on hyperplanes, which may be represented by a weighted linear combination of predictor variables[6-9] and regimes based on rectangular regions that may be represented by a decision tree (with binary splits based on dichotomized pretreatment characteristics and with the subgroups underlying the regime corresponding to the leaves of the tree).[10-16] The optimal treatment regime within a class then is the one that leads to the greatest benefit (ie, the highest expected potential outcome*) if it would be used to assign treatment to all patients in the population.

Recently, quite a few methods have been proposed to estimate optimal treatment regimes within various classes of regimes (for a comprehensive review, see the work of Lipkovich et al (2016)). These include both policy search methods that directly maximize an estimator of the expected outcome within the class of regimes of interest and indirect methods that first estimate some (eg, regression-based) outcome model, from which subsequently an optimal regime is derived.[18] Moreover, the methods in question are increasingly being used.[19-25]

Once a regime has been estimated, its quality is of obvious interest. The most straightforward measure to quantify this quality is the regime's expected potential outcome, as this captures how the regime would perform on average (in terms of the outcome variable of interest), if it would be used to assign treatment to the entire population under study. Obtaining a good and reliable estimate of the expected potential outcome is a major challenge for precision medicine[26,27]; it will be the focus of this paper.

To estimate the expected potential outcome of an optimal treatment regime, several estimators have been proposed. Examples include a simple estimator, where the mean outcome is calculated over all so-called lucky patients (ie, those patients for whom the treatment alternative received in the comparative treatment evaluation study and the treatment alternative implied by the treatment regime agree), and more sophisticated estimators including ones that augment the scores of the lucky patients based on an outcome model and/or that weight them based on a propensity model.[28,29]

When an optimal treatment regime has been estimated, one may further wonder which part of the data is to be used for the estimation of its expected outcome. The simplest data handling approach would be to use a resubstitution method, meaning that the full data set is used first to estimate the optimal treatment regime and next to estimate its expected potential outcome. Yet, this approach will usually lead to positively biased estimates, as a treatment regime (similar to a prediction model) will tend to perform better on the data used for estimating it compared to new data.[30] As a way out, one could split up the data into a training and a test set, where the training set is used to estimate the regime and the test set to estimate its expected potential outcome. However, reliably estimating optimal treatment regimes typically requires large sample sizes, and hence, it is often not desirable to split up the data.

In the domain of statistical learning, a similar problem has been at play, namely, the estimation of the prediction error of an estimated prediction rule. Within this domain, too, on the one hand, the resubstitution method has been found to lead to positively biased estimates, and on the other hand, researchers are often reluctant to split the sample into a training and test sample because of the associated loss of statistical precision. As a way out, resampling methods have been proposed to estimate the prediction error of a prediction rule (including, eg, *k*-fold cross-validation[30] and different variants of the bootstrap[31,32]). Given the similarities between estimating the prediction error of a prediction rule and estimating the expected potential outcome of an optimal treatment regime, it seems natural to consider the use of resampling methods within the latter context as well.

A number of comparative studies have been conducted to evaluate resampling methods in a statistical learning context.[31-35] However, there are several problems with them. First, the results of the studies in question are equivocal, with different best performing methods in different studies.[32-34,36] Second, negative correlations between true and estimated prediction errors have been reported.[30,31,37] Such correlations are worrisome, as they make one wonder whether it makes any sense to use any estimation procedure to estimate the quality of an estimated prediction rule or optimal treatment regime.[30] Third, it is not straightforward to extrapolate the findings from the domain of statistical learning to the context of estimating the expected potential outcome of an optimal treatment regime, for at least two reasons: (a) The estimands in both contexts are clearly different (ie, prediction error versus expected potential outcome); (b) A problem of fundamental missingness or missingness by design underlies the estimation of the expected outcome of an optimal treatment regime, unlike that of the prediction error of a prediction rule in statistical learning. For each experimental unit in a comparative treatment evaluation study, the outcome under only *one* of the treatment alternatives is typically known

*The potential outcome of a patient under a certain treatment alternative can be defined as the outcome that would have been observed if the patient would have received that treatment, regardless of the actually received treatment[17] Throughout this paper we assume, without loss of generality, that higher values of the outcome variable are better.

(as each experimental unit will typically receive only one of the treatment alternatives); the outcome under the other treatment alternative then is missing by design; this further implies that the true optimal treatment alternative for each experimental unit is not known and is to be estimated (unlike in a standard statistical learning problem, where the true class label for each experimental unit is typically known).

Because of the problems listed above, it is unclear which combination of an estimator of the expected potential outcome and a data handling approach will lead to the best estimates of the expected outcome of an estimated optimal treatment regime. This paper will address this question by setting up a new and comprehensive simulation study that involves several estimators of the expected potential outcome and several resampling methods.

2 | SIMULATION STUDY

2.1 | Overall setup and aims

2.1.1 | Data

The type of data that we will focus on is RCT data with n patients receiving one out of two available treatment alternatives, denoted by $A = 0$ and $A = 1$. The data include pretreatment characteristics denoted by $\mathbf{X} = (X_1, \ldots, X_j, \ldots, X_J)$. The observed outcome of interest is Y; we assume that larger values are better. As such, the observed data set is (X_i, A_i, Y_i), $i = 1, \ldots, n$.

In our Monte Carlo simulation study, we will manipulate three characteristics of the data and the data-generating mechanism. Subsequently, we will successively discuss these.

A first characteristic will be the sample size because, in practice, this may vary considerably between comparative treatment evaluation studies. At this point, we will include small, medium, and large sample sizes.

Second, we will manipulate the data-generating scenario in terms of the nature of the underlying treatment effect heterogeneity, where we aim to cover a broad and representative spectrum of types of heterogeneity. Specifically, we will include three scenarios in which the treatment effect heterogeneity is qualitative in nature, meaning that the optimal treatment alternative differs across subgroups of patients. Such a qualitative treatment-subgroup interaction implies that a nontrivial optimal treatment regime, assigning at least two different subgroups of patients to different treatment alternatives, is called for. The difference between the first scenario on the one hand and scenarios two and three on the other hand is that, for the first scenario, the subgroups are separated by a hyperplane (which can be represented by a linear combination of pretreatment characteristics), whereas, for the second and the third scenario, the subgroups are based on rectangular regions (which can be represented by a tree based on conjunctions of dichotomized pretreatment characteristics; see also the first and second panels of Figure 1in Section 2.2.1). The difference between the latter two scenarios further is that in the third scenario, unlike in the second, there is one leaf in which the two treatment alternatives perform equally well (which makes the estimation of the expected outcome of the estimated regimes more challenging, see the works of Luedtke and van der Laan[38] and Chakraborty et al [39]). Finally, the fourth scenario (which can be represented by a tree, too; see also the rightmost panel of Figure 1 in Section 2.2.1) involves a quantitative treatment effect heterogeneity, meaning that, although the difference in mean outcomes between the treatment alternatives differs across subgroups, the same treatment alternative is optimal in all subgroups; this implies that a trivial regime is called for, which is a regime that assigns all patients to the same, marginally best, treatment alternative.

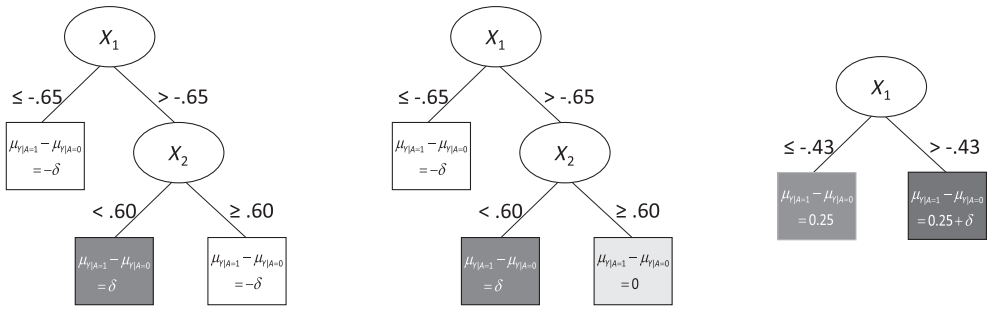


FIGURE 1 Graphical representation of data-generating scenarios 2, 3, and 4

Third, whereas within (most) subgroups at the basis of the true regimes, the size of the difference in outcome between the two treatment alternatives will differ across patients; we will manipulate the mean size of this difference. The reason for this is that the mean size in question may influence the difficulty of finding the true optimal treatment regime. Specifically, we will include a small, medium, and large mean size.

2.1.2 | Estimation of optimal treatment regime

From each simulated data set, we will estimate an optimal treatment regime within a prespecified class of treatment regimes τ, g_τ^{opt}, with g_τ^{opt} being a mapping from the value set of $\mathbf{X}$ to the value set of A, meaning that a patient with a pattern of pretreatment characteristics $\mathbf{X}$ will receive $A = 0$ if $g_\tau^{\text{opt}}(\mathbf{X}) = 0$ and $A = 1$ if $g_\tau^{\text{opt}}(\mathbf{X}) = 1$. The resulting estimated optimal treatment regime will further be denoted by $\widehat{g_\tau^{\text{opt}}}$.

We will use two different policy search methods for estimating optimal treatment regimes, each of which targets a different class of optimal treatment regimes. We will do this, on the one hand, to cover a sufficiently broad scope of different kinds of estimation methods, and on the other hand, to create both situations in which the true optimal regime underlying the data generation corresponds to the type of treatment regime that the estimation method targets and situations in which this is not the case.

2.1.3 | Estimand of interest

The estimand of interest is the true expected potential outcome of each estimated optimal treatment regime that was estimated on the basis of the specific data set under study, $\widehat{g_\tau^{\text{opt}}}$. Formally speaking, if we denote the potential outcome for treatment $A = 0$ (respectively, $A = 1$) by Y_0^* (respectively Y_1^*), the estimand of interest comes down to $E[Y_0^*(1 - \widehat{g_\tau^{\text{opt}}}) + Y_1^*\widehat{g_\tau^{\text{opt}}}]$.

2.1.4 | Estimation strategies to be compared

The specification of a strategy to estimate the target estimand includes the choice of an estimator and that of a data handling strategy. In this study, we will orthogonally combine two estimators with seven data handling approaches. In addition, we will also include an estimation strategy in which estimation and data handling are more intertwined.

As estimator of the expected outcome, we will use, on the one hand, a simple estimator, with the mean outcome being calculated over all lucky patients. On the other hand, we will use the more sophisticated AIPW estimator,[28] which implies that the outcomes of the lucky patients are weighted based on a propensity model and augmented based on an outcome model. In an RCT, the simple estimator (with an appropriate data handling technique) is unbiased. From its part, the choice of the AIPWE can be motivated on the basis of its solid theoretical properties (ie, double robustness, implying that it remains consistent for the expected potential outcome if either the outcome model or the propensity model is specified correctly) and of the fact that it has proven to perform well empirically.[6,10]

With regard to the data that are used to estimate the expected outcome, we will compare the (presumably severely upward biased) resubstitution method with six resampling methods. A first resampling method that we will include is k-fold cross-validation, where the data are divided in k equal-sized folds, where in each of k rounds, all folds except the kth one are used to estimate the regime, and where the kth fold is used to estimate its expected outcome; finally, the estimates are averaged across the folds. Taking into account previous results, we will include fivefold and tenfold cross-validation in our comparison. Second, we will include several variants of the bootstrap,[31] where data sets of the same size as the original sample are drawn with replacement from the original data, and a regime is estimated on each bootstrap sample. In our comparison, we will include: (a) the ordinary bootstrap,[31] where the overoptimism implied by using the data twice is first estimated (by estimating the expected outcome of each regime on both the bootstrap sample and the original data set) and subsequently subtracted from the resubstitution estimate of the expected outcome; (b) the out-of-bag bootstrap,[31] where only the observations that were not used for estimation of the treatment regime (ie, the observations that were not in the bootstrap sample) are used for estimating its expected outcome; (c) the .632 bootstrap,[31] which pulls the out-of-bag estimator (which often overcorrects for the upward bias) toward the resubstitution estimate by weighting the out-of-bag estimate by the probability of an observation being in a bootstrap sample (.632) and the resubstitution estimate by its complement (.368); and (d) the $.632^+$ estimator,[32] which also relies on a weighted combination of the out-of-bag estimate and the resubstitution estimate, yet uses weights based on the estimated amount of overfitting.

In addition to the fourteen combinations of the two estimators with the seven data handling approaches outlined above, we will also include in the comparison an instance of cross-validated targeted minimum loss-based estimation

(CV-TMLE).[40] For the key estimand at issue here, this approach includes an estimation of working outcome and propensity models, followed by a targeted bias reduction step (with the latter meaning that the approach aims at reducing the bias of the estimate of the key estimand, at the expense of a possibly larger bias for other model aspects that are dealt with as nuisance parameters). To make the approach further maximally robust, in the bias reduction step, an appeal was made to *k*-fold cross-validation. The CV-TMLE approach was added to the current comparison because of its theoretical properties (which include asymptotically valid inference under minimal conditions).

2.1.5 | Aims

Our primary aim is to compare the performance of the different estimation strategies under study. The key performance measure for this purpose will be the size of the absolute difference between the true value of the estimand of interest and the estimate produced by each estimation strategy (which may be considered a natural measure of estimation error). Beyond a performance comparison of the methods, we further also want to know whether possible performance differences depend on characteristics of the data and/or the targeted type of optimal treatment regime.

As a secondary aim, we want to investigate whether each of the estimation strategies is on average too strict or too liberal and, if yes, to what extent. For this purpose, we will rely on the arithmetic difference between the true value of the estimand of interest and the estimate produced by the estimation strategy under study.

Last but not least, we want to investigate the correlations between the true and the estimated values of the estimand of interest. Key concern at this point is whether the highly worrisome near-zero and even negative values of analogous correlations found in the domain of statistical learning would also be found within the present context.

2.2 | Method

2.2.1 | Data

Each simulated data set comprised n observations (Y_i, A_i, X_i), $i = 1, \ldots, n$, where $X_i = (X_{i1}, \ldots, X_{i5})^T$, was generated from a multivariate normal distribution with $\mu_{X_j} = 0$, $\sigma_{X_j} = 1$ and $\rho_{jj'} = 0.2$. A was generated from a Bernoulli distribution with $\theta = 0.50$. Outcomes were generated as $Y_i = W_i(A, X) + \epsilon_i$ with ϵ_i standard normal, and with $W_i(A, X)$ as to be specified below.

Three factors were manipulated in the data generation process. First, we manipulated the sample size with levels $n = 150$, $n = 300$, and $n = 1000$.

Second, we manipulated the nature of the underlying treatment effect heterogeneity in terms of the following four data-generating scenarios (the last three of which are graphically represented in Figure 1):

1. *Hyperplane-based, qualitative*: $W_i(A, X) = 1.0 + 0.25X_{i1} + 0.25X_{i2} - 0.25X_{i5} - D_i[A_i - g^{\text{opt}}(X_i)]^2$; g^{opt} denotes the true optimal treatment regime with $g^{\text{opt}}(X) = I(X_1 + X_2 > 0)$; furthermore, D_i is the difference in outcome for patient i between the optimal and nonoptimal treatment alternatives (as to be further specified below).
2. *Tree-based, qualitative*: $W_i(A, X) = 1.0 + 0.25X_{i1} + 0.25X_{i2} - 0.25X_{i5} - D_i[A_i - g^{\text{opt}}(X_i)]^2$; $g^{\text{opt}}(X) = I(X_1 > -0.65)I(X_2 < 0.60)$.
3. *Tree-based, qualitative with one leaf in which two treatment alternatives perform equally well*: $W_i(A, X) = 1.0 + 0.25X_{i1} + 0.25X_{i2} - 0.25X_{i5} - D_i[1 - I(X_1 > -0.65)I(X_2 \geq 0.60)][A - g^{\text{opt}}(X_i)]^2$; $g^{\text{opt}}(X) = I(X_1 > -0.65)I(X_2 < 0.60)$. Note that in this scenario, the optimal treatment regime is not uniquely defined because of the existence of a subgroup for which the two treatment alternatives perform equally well. Hence, the $g^{\text{opt}}(X)$ proposed above is merely one of two possible optimal treatment regimes.
4. *Tree-based, quantitative*: $W_i(A, X) = 1.0 + 0.25X_{i1} + 0.25X_{i2} - 0.25X_{i5} - 0.25\{I(X_{i1} \leq -0.43)[A_i - I(X_{i1} \leq -0.43)]^2\} - (0.25 + D_i)\{I(X_{i1} > -0.43)[A_i - I(X_1 > -0.43)]^2\}$, with 0.25 now being the loss in expected outcome implied by assigning patient i to the nonoptimal treatment alternative when $X_1 \leq -0.43$ and $0.25 + D_i$ to the loss implied when assigning that patient to the nonoptimal treatment alternative when $X_{i1} > -0.43$.

Third, in all scenarios, D_i denotes the size of the difference in outcome between the two treatment alternatives for patient i (except in the indifference subgroup under scenario (3) and in the subgroups under scenario (4)). We put $D = \delta U$, with U following a uniform distribution with support $[0, 2]$, and δ denoting the mean difference. We manipulated the size of the mean difference, with levels small ($\delta = 0.25$), medium ($\delta = 0.50$), and large ($\delta = 1$).

The three design factors were fully crossed, leading to 3 (sample size) × 4 (scenario) × 3 (mean size difference) = 36 conditions. For each condition, we simulated 100 Monte Carlo data sets, yielding 3600 data sets in total.

2.2.2 | Estimation of optimal treatment regime

We included two methods to estimate optimal treatment regimes g_τ^{opt} for each data set, namely, the method of Zhang et al (2012), which estimates optimal linear regimes that imply subgroups separated by hyperplanes ($\tau = h$), g_h^{opt}, and the method of Zhang et al (2012), which estimates tree-based treatment regimes that imply subgroups represented by rectangular regions ($\tau = r$), g_r^{opt}. This implies a total of $3600 \times 2 = 7200$ estimated treatment regimes.

For the method of Zhang et al (2012), we used the package DynTxRegime[41] with default settings, except that we used $\mu(A, X, \hat{\beta}) = \beta_0 + \sum_{j=1}^{5} \beta_j X_j + A(\beta_0^* + \sum_{j=1}^{5} \beta_{j+5} X_j)$ as outcome model and empirical proportions as propensity scores to account for possible chance imbalance of treatment assignment. We used the class of regimes (`regimes`) $\beta_0^* + \beta_6 X_1 + \beta_7 X_2 + \beta_8 X_3 + \beta_9 X_4 + \beta_{10} X_5 > 0$. As domains (`Domains`), we used $(-1, 1)$; as starting values (`starting.values`), we used 0; and as solution tolerance (`solution.tolerance`), we used .0001.

For the method of Zhang et al (2012), we used self-built R-code for the contrast calculation. As outcome model, we used $\mu(A, X, \hat{\beta}) = \beta_0 + \sum_{j=1}^{5} \beta_j X_j + A(\beta_0^* + \sum_{j=1}^{5} \beta_{j+5} X_j)$. With regard to the propensity model, we used empirical proportions. We estimated the contrast using $\hat{C}_{AIPWE}(X_i, A_i, Y_i, \hat{\beta})$. Subsequently, we defined binary responses $\hat{Z}_i = I(\hat{C}_{AIPWE}(X_i, A_i, Y_i; \hat{\beta}) > 0)$ and weights $\hat{W}_i = |\hat{C}_{AIPWE}(X_i, A_i, Y_i; \hat{\beta})|$. The new data set $(\hat{Z}_i, X_i, \hat{W}_i)$ was analyzed using `rpart`[42] with default setting, except that $\hat{W}$ was used as individual misclassification costs (`weights=` $\hat{W}$), the minimum number of persons in a node to split was set at 40 (`minsplit=40`), the minimum number of persons in a terminal node at 20 (`minbucket=20`), and the maximum depth of the tree at six (`maxdepth=6`). For the complexity parameter used in the pruning procedure (`cp`), we used the complexity parameter maximally one standard error above the one with the smallest cross-validated error.[30]

2.2.3 | Estimand of interest

The key estimand is the true expected potential outcome of each estimated optimal treatment regime. Importantly, this estimand varies across the 7200 estimated optimal treatment regimes under study.

Usually, it is not possible to analytically derive the value of the estimand in question. Therefore, we will approximate this value making use of a Monte Carlo procedure. This procedure implies that, given a simulated data set **D** and an estimated optimal treatment regime $\widehat{g_\tau^{\text{opt}}}$ derived from **D**, full pretreatment characteristic and potential outcome data are simulated for a pseudo-population of 10^6 patients based on the true data-generating model underlying **D**. Next, this pseudo-population is subjected to $\widehat{g_\tau^{\text{opt}}}$; the resulting potential outcome values can then be simply averaged, yielding a suitable approximation of the estimand of interest.[6,10,33]

2.2.4 | Estimation strategies to be compared

After having estimated the optimal treatment regimes, we will estimate their expected outcomes. At this point, we will use two estimators, namely, a simple one and the more involved AIPWE,[28] in combination with seven data handling approaches (viz, the resubstitution method, fivefold cross-validation, tenfold cross-validation, ordinary bootstrap, out-of-bag bootstrap, .632 bootstrap, and .632$^+$ bootstrap). In addition, we also included a cross-validated targeted minimum loss-based estimator (CV-TMLE)[40] in the comparison.

A simple estimate of the expected potential outcome was obtained as follows:

$$\frac{1}{\sum_{i=1}^{n} I\left(A_i = \widehat{g_\tau^{\text{opt}}}(X_i)\right)} \sum_{i=1}^{n} Y_i \, A_i \, \widehat{g_\tau^{\text{opt}}}(X_i) + Y_i \, (1 - A_i) \left(1 - \widehat{g_\tau^{\text{opt}}}(X_i)\right). \tag{1}$$

From its part, the AIPW estimate was obtained as

$$n^{-1} \sum_{i=1}^{n} \left[\frac{C_{\widehat{g_\tau^{\text{opt}}},i} \, Y_i}{\pi_c\left(X_i; \widehat{g_\tau^{\text{opt}}}, \hat{\gamma}\right)} - \frac{C_{\widehat{g_\tau^{\text{opt}}},i} - \pi_c\left(X_i; \widehat{g_\tau^{\text{opt}}}, \hat{\gamma}\right)}{\pi_c\left(X_i; \widehat{g_\tau^{\text{opt}}}, \hat{\gamma}\right)} m\left(X_i; \widehat{g_\tau^{\text{opt}}}, \hat{\beta}\right) \right], \tag{2}$$

where

- $C_{\widehat{g_\tau^{\text{opt}}}} = A\widehat{g_\tau^{\text{opt}}}(X) + (1 - A)(1 - \widehat{g_\tau^{\text{opt}}}(X))$;

- $\pi_c(X;\widehat{g_\tau^{\text{opt}}},\hat{\gamma}) = P(C_{\widehat{g_\tau^{\text{opt}}}} = 1|X,\hat{\gamma}) = \pi(X;\hat{\gamma})\,\widehat{g_\tau^{\text{opt}}}(X) + (1-\pi(X;\hat{\gamma}))\,(1-\widehat{g_\tau^{\text{opt}}}(X))$, with $\pi(X;\gamma)$ denoting a posited model for the propensity score $P(A=1|X)$, for which we used empirical proportions, whereas normally $\hat{\gamma}$ is the maximum likelihood estimate for γ;
- $m(X;\widehat{g_\tau^{\text{opt}}},\boldsymbol{\beta}) = \mu(1,X,\boldsymbol{\beta})\widehat{g_\tau^{\text{opt}}}(X) + \mu(0,X,\boldsymbol{\beta})(1-\widehat{g_\tau^{\text{opt}}}(X))$, where $\mu(A,X;\beta)$ is a model for $\text{E}(Y|A,X)$ and $\hat{\beta}$ is an estimator for β; for $\mu(A,X;\beta)$, we reused the model that was used for estimating the regimes (based on the full data set) for computational efficiency.

Regarding the data handling approaches:

- *resubstitution*: We obtained the resubstitution estimate, $\widehat{RES}$, by calculating the estimator from Equation (1) (respectively (2)) on the full data set.
- *k-fold cross-validation*: We assigned each observation to one of k folds using the `sample` function in `R`. We obtained the k-fold cross-validation estimate by (a) estimating g_τ^{opt} on the data while leaving out one of the folds, and subsequently, calculating the estimator from Equation (1) (respectively (2)) on the data from the fold that was left out, and (b) averaging the estimate from the previous step across all folds.
- *ordinary bootstrap*: We created 50 bootstrap samples of the same size as the original data set by sampling with replacement from it using the `sample` function in `R`. Next we (a) estimated g_τ^{opt} on each bootstrap sample, subsequently calculated the estimator from Equation (1) (respectively (2)) on both the bootstrap sample and the full data set and calculated the difference between the former and latter estimate (as a measure of overoptimism), and (b) averaged the difference resulting from the previous step across all 50 bootstrap samples and subtracted this from $\widehat{RES}$.
- *out-of-bag bootstrap*: We used again the 50 bootstrap samples and the estimated g_τ^{opt} for each sample. To obtain the out-of-bag bootstrap estimate, $\widehat{OOB}$, we calculated the estimator from Equation (1) (respectively (2)) based on the patients outside each bootstrap sample and, finally, averaged the result across all 50 bootstrap samples.
- *.632 bootstrap*: This estimate, further denoted by $\widehat{.632}$, was obtained as $.632\,\widehat{OOB} + .368\,\widehat{RES}$.
- *.632⁺ bootstrap*: In a simplified form, the .632⁺ bootstrap estimator $\widehat{.632^+}$ may be defined as follows:

$$\widehat{.632^+} = w\widehat{OOB} + (1-w)\widehat{RES}, \tag{3}$$

with w ranging from .632 to 1 and representing the estimated amount of overfitting. In fact, w is defined as $w = \frac{.632}{1-.368\hat{R}}$, with the normalized estimated overfitting rate $\hat{R}$ being defined as $\hat{R} = \frac{\widehat{RES}-\widehat{OOB}}{\widehat{RES}-\hat{\imath}}$; $\hat{\imath}$ further denotes the estimated expected outcome in case of independence between the treatment regime under study and outcome; it can be calculated as $\hat{\imath} = \overline{Y|A=1}\cdot\widehat{p(g_\tau^{\text{opt}})} + \overline{Y|A=0}\cdot(1-\widehat{p(g_\tau^{\text{opt}})})$, where $\widehat{p(g_\tau^{\text{opt}})}$ is the proportion of patients assigned to treatment 1 according to the estimated treatment regime, $\overline{Y|A=1} = \frac{1}{\sum A_i}\sum_{i=1}^n Y_i A_i$ and $\overline{Y|A=0} = \frac{1}{n-\sum A_i}\sum_{i=1}^n Y_i(1-A_i)$. While one may wish $\hat{R}$ to take values between 0 and 1, this may not always be the case as it is possible that $\widehat{RES} \leq \hat{\imath}$ or $\widehat{OOB} < \hat{\imath} \leq \widehat{RES}$. To account for these possibilities in line with the work of Efron and Tibshirani,[32] we first observe that, taking into account all definitions given above, (3) can be rewritten as

$$\widehat{.632^+} = \widehat{.632} + \frac{.368\cdot.632\cdot\hat{R}}{1-.368\cdot\hat{R}}(\widehat{OOB}-\widehat{RES}). \tag{4}$$

Furthermore, in line with the aforementioned work,[32] we define $\widehat{OOB}^* = \max(\widehat{OOB},\hat{\imath})$ and

$$\hat{R}^* = \begin{cases} \frac{\widehat{OOB}^*-\widehat{RES}}{\hat{\imath}-\widehat{RES}} & \text{if } \widehat{OOB} < \widehat{RES} \text{ and } \hat{\imath} < \widehat{RES} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

We then finally may replace (4) by

$$\widehat{.632^+} = \widehat{.632} + \frac{.368\cdot.632\cdot\hat{R}^*}{1-.368\cdot\hat{R}^*}(\widehat{OOB}^*-\widehat{RES}). \tag{6}$$

For the *CV-TMLE*, we used code provided by Luedtke in the following way: First, to estimate the expected outcome under both treatments for each patient with tenfold cross-validation, we used the outcome model $\mu(A, X, \hat{\beta}) = \beta_0 + \sum_{j=1}^{5} \beta_j X_j + A(\beta_0^* + \sum_{j=1}^{5} \beta_{j+5} X_j)$. Second, to estimate the probability of receiving each treatment alternative for each patient, we used the empirical proportions for both treatment alternatives. Third, for each fold the linear (respectively tree-based), regime was estimated on the full data set minus that fold and, subsequently, the blip (expected difference in expected outcome when obtaining treatment 1 compared to treatment 0, conditional on covariates, when assigning treatment based on the optimal treatment regime) was estimated on that fold. Using these three elements, and after a bias reduction step based on a fluctuation test, the expected outcome of the estimated treatment regime was estimated. The full procedure (including the tenfold cross-validation) was subsequently repeated ten times, and as our final estimate of the expected outcome, we used the average of those ten estimates.

2.3 | Results

2.3.1 | Absolute difference

Each of the 7200 estimated optimal treatment regimes was subjected to each of the $(2 \times 7 + 1 =)$ 15 estimation strategies. Primary dependent variable was the absolute difference between the (approximated) estimand of interest and the estimate resulting from each strategy.

Prior to the subsequent analyses, we aligned the results of CV-TMLE with those of the other strategies. For this purpose, we took into account that the CV-TMLE, through its use of working outcome and propensity models, is similar in construction to the AIPWE-based strategies. We therefore subjected the results of the first $(2 \times 7 =)$ 14 estimation strategies to a factorial ANOVA with sample size, scenario, mean size difference, and regime estimation method as between-blocks factors and estimator and data-handling strategy as within-blocks factors. We used η^2 as effect size, where $\eta^2 = \frac{SS_{\text{effect}}}{SS_{\text{total}}}$ (with SS_{effect} denoting the sum of squares of the effect and SS_{total} the total sum of squares). We calculated SS_{total} separately for the within- and between-subject effects, which implies that the η^2-values of all between-subject effects and their interactions sum up to 1; similarly, the η^2-values of the within-subject effects (including their interactions) and their interactions with between-subject effects sum up to 1. As a cutoff, we used $\eta^2 = .01$, implying that an effect will be reported if it explains at least 1% of the corresponding total sum of squares. The sizes of the main effect of estimator and the interaction effects involving estimator appeared to be negligible, $\eta^2 < .01$. As a consequence, for the subsequent analyses, we only retained the results for the AIPWE (paired with the seven data handling strategies) and the CV-TMLE.

For the retained conditions, the overall mean absolute difference appeared to be 0.087 (sd = 0.078). To investigate whether possible performance differences depended on the estimation strategy, characteristics of the data and/or the targeted type of optimal treatment regime, we subjected the primary dependent variable to a factorial ANOVA similar to the one outlined above, with now estimation strategy as the only within-blocks factor, and retaining again only effects with $\eta^2 \geq .01$. This ANOVA revealed a sizeable main effect ($\eta^2 = .36$) of sample size with mean values (standard deviations) of 0.123 (0.095) for $n = 150$, 0.090 (0.071) for $n = 300$, and 0.047 (0.037) for $n = 1000$. A second sizeable main effect was found for estimation strategy ($\eta^2 = .18$). The means (standard deviations) per strategy are presented in the second column of Table 1. The performance of the resubstitution strategy appeared to much worse than that of the resampling methods. Within the group of resampling methods, the best results were obtained for the out-of-bag and .632$^+$ bootstrap estimators.

The ANOVA further revealed three nonnegligible interactions. Firstly, we observed a two-way interaction ($\eta^2 = .07$) between estimation strategy and size of the mean difference in outcome between the two treatment alternatives (see Figure 2). From this Figure, it appears that resubstitution performs much better in case of more sizeable differences

TABLE 1 Mean absolute differences and standard deviations, mean arithmetic differences and standard deviations, and correlations averaged over all replications and data-generating conditions of the simulation study, the two methods for estimating treatment regimes, separately for each of the seven data handling approaches (coupled with the AIPWE) and the cross-validated targeted minimum loss-based estimation (CV-TMLE)

Estimation strategy	Mean absolute difference (sd)	Mean arithmetic difference (sd)	Correlation
Resubstitution	0.152 (0.112)	0.143 (0.120)	.21
Fivefold cross-validation	0.081 (0.074)	−0.004 (0.109)	.48
Tenfold cross-validation	0.081 (0.072)	−0.003 (0.109)	.49
Ordinary bootstrap	0.084 (0.071)	0.031 (0.105)	.40
Out-of-bag bootstrap	0.072 (0.063)	−0.025 (0.094)	.52
.632 bootstrap	0.077 (0.065)	0.037 (0.093)	.44
.632$^+$ bootstrap	0.072 (0.062)	0.012 (0.094)	.47
CV-TMLE	0.076 (0.067)	−0.003 (0.101)	.52

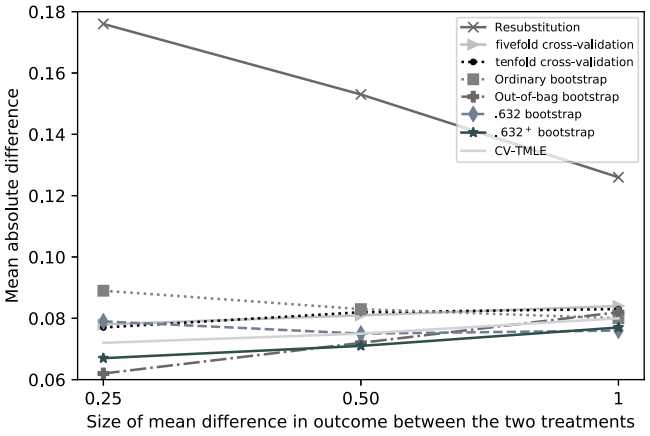


FIGURE 2 Interaction effect between estimation strategy and size of the mean difference in outcome between the two treatment alternatives on the mean absolute difference between the estimated expected outcome and the true expected outcome of a regime. CV-TMLE, cross-validated targeted minimum loss-based estimation [Colour figure can be viewed at wileyonlinelibrary.com]

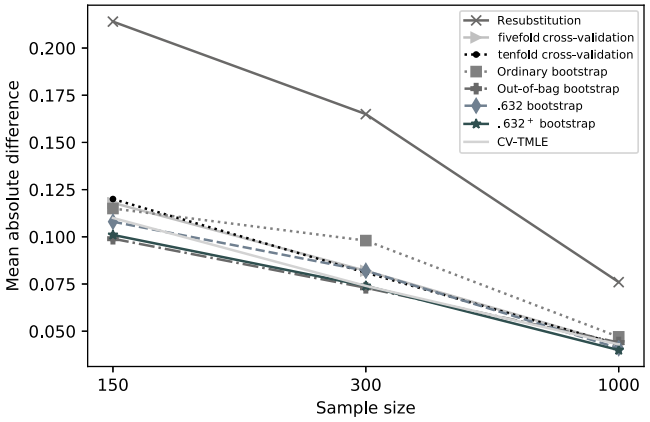


FIGURE 3 Interaction effect between estimation strategy and sample size on the mean absolute difference between the estimated expected outcome and the true expected outcome of a regime. CV-TMLE, cross-validated targeted minimum loss-based estimation [Colour figure can be viewed at wileyonlinelibrary.com]

between the two treatment alternatives; moreover, this is also the case, albeit to a lesser extent, for the other methods that overrely on the resubstitution estimate (ie, bootstrap and .632 bootstrap).

Secondly, we also observed a two-way interaction ($\eta^2 = .03$) between estimation strategy and sample size (see Figure 3). From this Figure, it appears that, whereas all methods perform better with increasing sample sizes, this is especially the case for resubstitution.

Thirdly, the latter two-way interaction appeared to be qualified by a three-way interaction ($\eta^2 = .03$) between sample size, estimation strategy, and class of treatment regimes, with for resubstitution the impact of increasing sample size on mean absolute difference being much larger for linear treatment regimes (see Figure 4).

2.3.2 | Arithmetic difference

To investigate whether each of the estimation strategies is on average too strict or too liberal, we further investigated the arithmetic difference between the true value of the estimand of interest and the estimate produced by the estimation strategy under study. We calculated the mean and standard deviation per strategy. Importantly, one should note that, because the estimand of interest varies across all 7200 estimated optimal treatment regimes, the resulting means cannot be readily interpreted in terms of a bias of the estimation strategies under study, and especially also that the standard deviations do not relate in a straightforward way to standard errors (for a formal account of this issue, see equation 2.1 in the work of Kim[33]).

The means and standard deviations are shown in the third column of Table 1. It appears that the arithmetic difference is closest to zero for the tenfold cross-validation and the CV-TMLE and farthest away from zero for the resubstitution method.

2.3.3 | Correlation

Finally, we calculated the correlations between the true and the estimated values of the estimand of interest for each of the data handling approaches separately. Except for the resubstitution estimator, for all approaches, the correlations between

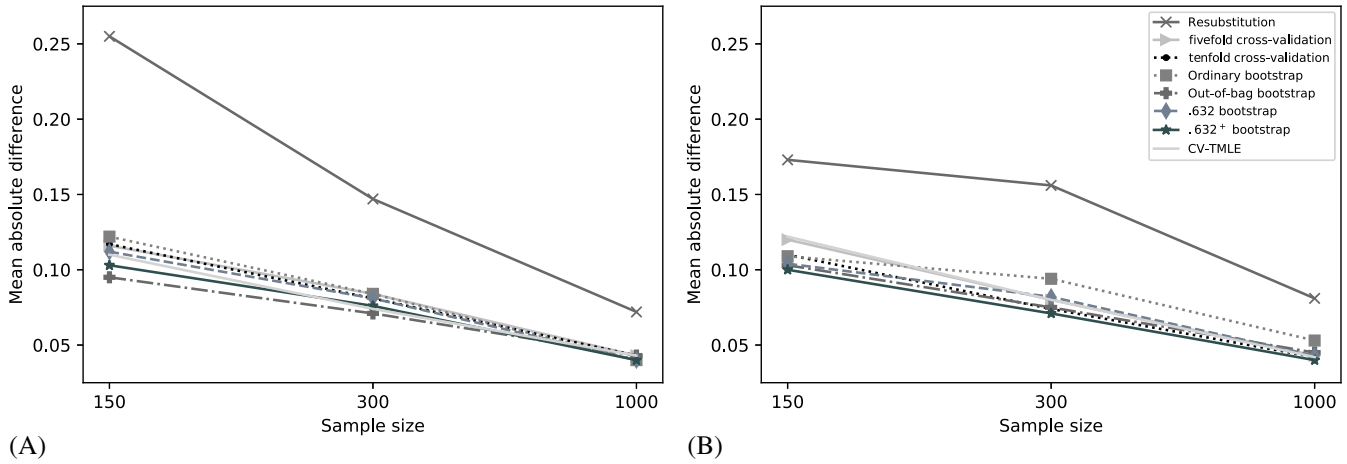


FIGURE 4 Three-way interaction effect between sample size, estimation strategy, and class of treatment regimes on the mean absolute difference between the estimated expected outcome and the true expected outcome of a regime A, Linear treatment regimes; B, Tree-based treatment regimes. CV-TMLE, cross-validated targeted minimum loss-based estimation [Colour figure can be viewed at wileyonlinelibrary.com]

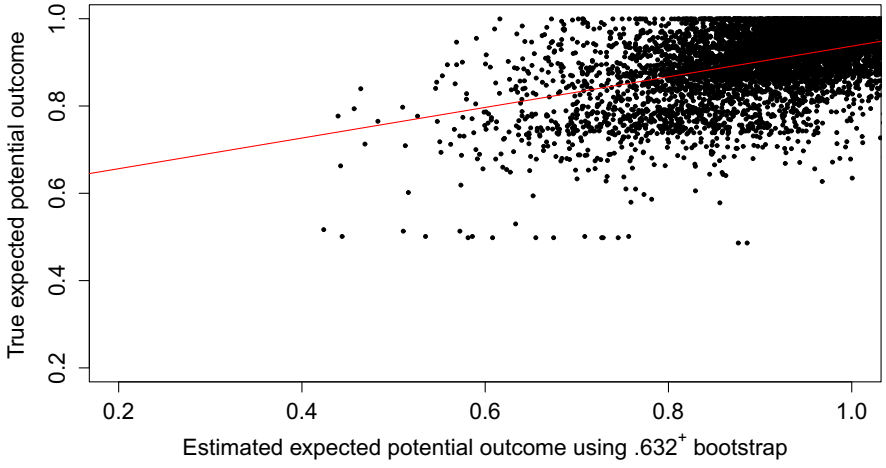


FIGURE 5 Scatter diagram of true vs estimated values of estimand of interest for 7200 estimated treatment strategies with AIPWE coupled with .632+ bootstrap estimator [Colour figure can be viewed at wileyonlinelibrary.com]

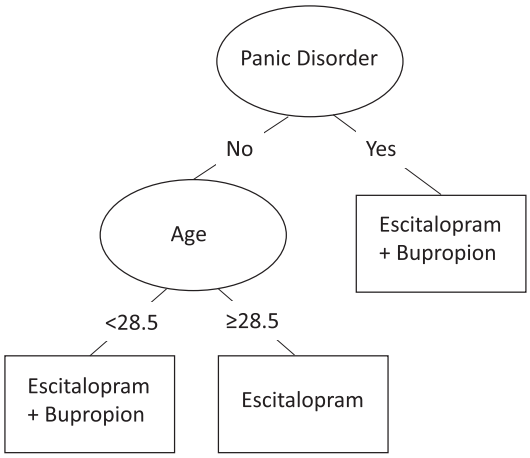
the true and estimated expected outcomes were between .40 and .52 (see rightmost column of Table 1). To exemplify this result, one may inspect the scatter diagram for the .632+ bootstrap estimator (see Figure 5).

3 | ILLUSTRATIVE APPLICATION

We will illustrate using the "Combining Medication to Enhance Depression Outcomes" (CO-MED) data set.[43] In this RCT data set (on 665 participants with moderately severe nonpsychotic chronic and/or recurrent major depressive disorder), the efficacy of two combination treatments (sustained-release bupropion + escitalopram and extended-release venlafaxine + mirtazapine) was compared to the efficacy of a mono-treatment (escitalopram + placebo) in terms of the QIDS-SR score (a self-report measure of depressive symptom severity[44]). In addition, for each patient, a number of pretreatment characteristics was measured. At this point, we will focus on the bupropion + escitalopram combination treatment arm and the escitalopram + placebo mono-treatment arm and on a selection of pretreatment characteristics, which includes a number of demographic characteristics, characteristics relating to disease severity and characteristics that indicate whether a certain comorbid psychiatric disorder is present. From these data, we estimated a linear treatment regime using the method of Zhang et al and a tree-based treatment regime using the method of Zhang et al.[10] In the linear regime, depicted in Panel (A) of Figure 6, positive weights indicate that higher values on the characteristic in question imply a benefit from the combination treatment, whereas negative weights indicate that higher values on the characteristic in question imply a benefit from the mono-treatment. Interestingly, the tree-based treatment regime, depicted in

Characteristic	Weight
Intercept	-0.264
Age	-0.551
Gender	-0.921
Length of episode	-0.103
QIDS baseline	-0.411
PTSD	0.537
Bulimia	-0.295
OCD	-0.372
Panic disorder	0.640
Agoraphobia	-0.539
Social phobia	0.439
Alcohol addiction	0.166
Drug addiction	0.355
Generalized anxiety	0.042
Somatisation	0.143
Hypochondriasis	0.169

(A)



(B)

FIGURE 6 Optimal treatment regimes estimated from the Combining Medication to Enhance Depression Outcomes data set: (A) Linear treatment regime depending on a weighted linear combination of predictor variables, with the (standardized) weights being specified in the second column. If this combination is larger than zero, the optimal treatment alternative is the combination treatment (escitalopram + bupropion); otherwise, it is the mono-treatment (escitalopram). (b) Tree-based treatment regime with the optimal treatment alternative depending on a conjunction of dichotomized pretreatment characteristics.

	Linear		Tree	
	Simple	AIPWE	Simple	AIPWE
Resubstitution	9.83	10.14	9.73	9.50
Fivefold cross-validation	9.42	9.13	9.17	8.92
Tenfold cross-validation	9.04	9.03	9.10	8.83
Ordinary bootstrap	9.56	9.83	9.33	8.85
Out-of-bag bootstrap	9.05	8.88	9.13	8.88
.632 bootstrap	9.34	9.34	9.35	9.11
$.632^{+}$ bootstrap	9.09	8.91	9.19	8.92
CV-TMLE	–	8.34	–	8.48

TABLE 2 Estimated expected reduction in QIDS-SR score if all patients under study would receive treatment according to the linear treatment regime shown in Figure 4A, and according to the tree-based treatment regime shown in Figure 4B (both estimated from the Combining Medication to Enhance Depression Outcomes data), based on the simple estimator of the expected outcome (the mean of the lucky patients) and based on the more sophisticated AIPWE, each of which being coupled with the seven different data handling approaches, and based on the cross-validated targeted minimum loss-based estimation (CV-TMLE)

Panel (B) of Figure 6, partly links up with the linear one: It indicates that patients with panic disorder and patients younger than 28.5 years old should receive the combination treatment, whereas patients without panic disorder who are at least 28.5 years old should receive the mono-treatment. In the linear regime, leaving aside gender, panic disorder, and age also has the largest standardized weights (in absolute value), with a positive and negative sign, respectively.

Next, we used all possible combinations of the two estimators of the expected outcome with the seven resampling methods, in addition to the CV-TMLE, to evaluate the quality of both treatment regimes. Table 2 shows the results. By and large, the estimates for the linear and tree-based regimes appeared to be fairly similar, with the expected reduction in QIDS-SR score after 12 weeks of treatment (if the whole population would be assigned to treatment based on one of the estimated treatment regimes) being between 8.4 and 9.4.

Interestingly, the rank order of the estimates of the expected outcomes of the different data handling approaches at least partially links up with the rank order of the mean arithmetic differences that was found in the simulation study (given the standard deviations found in the simulation study, it is unlikely that the rank order of the methods in the application would fully match the rank order in the simulation results). In the simulation study, the resubstitution estimate was the highest, followed by the .632 bootstrap and ordinary bootstrap estimates. In the application, too, the highest estimate is consistently produced by resubstitution, whereas in most columns, the .632 and ordinary bootstrap estimates also figure among the three highest estimated values.

4 | DISCUSSION

On average, in terms of our primary criterion, the best data handling approaches for estimating the expected potential outcome of a treatment regime appeared to be the out-of-bag and $.632^{+}$ bootstrap. Compared to the other approaches,

these two methods turned out to have optimal or at least near-optimal performance across all data types, methods for estimating treatment regimes, and estimators of expected outcome. Not surprisingly, resubstitution performed worse, with the difference between resubstitution and the other approaches being especially large in "more difficult" conditions (smaller sample sizes and smaller mean differences between the two treatment alternatives). With regard to the estimators of the expected outcome, it can be concluded that, for the conditions in our simulation study, their estimation performance was very similar (as the η^2-values of the main effect and the interactions with data generating characteristics, resampling methods, and methods for estimating a treatment regime were smaller than .01).

As a secondary criterion, we investigated the arithmetic difference between the true value of the estimand of interest and the estimate produced by each of the estimation strategies under study. Interestingly, the rank order of the estimates of the expected outcomes that resulted from the application of the different data handling approaches to the treatment regimes that were obtained for the empirical data set reflected, at least in part, the amount of overoptimism/underoptimism that was found in the simulation study. When focusing more in detail on the size of the mean arithmetic difference in the latter study (and leaving aside the obviously highly overoptimistic resubstitution), the cross-validation–based approaches (including the CV-TMLE) appeared to outperform their bootstrap counterparts. Although, in this study, the link between the mean arithmetic difference and the concept of bias is somewhat more involved (as the estimand of interest varies across all estimated optimal treatment regimes), this result is fully in line with the classical finding that cross-validation in general outperforms bootstrap methods in terms of amount of bias.[30-32,34,36] That being said, within the group of bootstrap methods, the $.632^+$ bootstrap has the lowest mean arithmetic difference, which links up with the rationale behind the construction of this estimator to optimally reduce bias through a fine-tuning of weights based on an estimated amount of overfitting.[32]

Otherwise, bootstrap estimators are also known to have a lower variance than methods based on cross-validation.[31-34,36] Although we do not have readily interpretable information on the variance of the estimation strategies under study (again, as the target estimand varies across all estimated optimal treatment regimes), we may reasonably conjecture that the combination of a lower variance with a relatively lower bias might account for the good result for the $.632^+$ bootstrap (and to some extent also for the out-of-bag bootstrap) on the primary criterion variable.

Finally, we found correlations between the true and estimated expected outcomes that were consistently between .40 and .52. Fortunately, these values are substantially larger than the corresponding correlations reported in studies in the domain of statistical learning.[30-32,37] One might be tempted to attribute these differences in results to the fact that, in the domain of statistical learning, correlations between true and estimated prediction errors of (estimated) prediction rules have been investigated, whereas our focus was on correlations between the true and estimated expected outcomes of (estimated) treatment regimes. Alternatively, one might conjecture that the studies in the statistical learning domain mainly focused on the prediction errors of linear methods, whereas we addressed a broader scope of methods. However, the most critical difference between earlier studies within the domain of statistical learning and the present one is that the earlier studies exclusively focused on correlations between true and estimated prediction errors across samples drawn from a single population, whereas this study focused on correlations between true and estimated expected outcomes across samples from varying populations. At this point, focusing on a single population only can be expected to lead to a considerable restriction of range for the true prediction errors, which may explain the low and even negative correlations. By focusing on varying populations, this problem may disappear. Moreover, such a setting may also be much more representative for data-analytic practice.

ACKNOWLEDGEMENTS

Data used in the preparation of this manuscript were obtained and analyzed from the controlled access data sets distributed from the NIMH-supported National Database for Clinical Trials (NDCT). NDCT is a collaborative informatics system created by the National Institute of Mental Health to provide a national resource to support and accelerate discovery related to clinical trial research in mental health. Data set identifiers: Study ID: N01 MH090003-02; clinical trial ID: NCT00590863. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIMH or of the submitters submitting original data to NDCT.

DATA AVAILABILITY STATEMENT

The data can be requested through the NDCT.

ORCID

Aniek Sies https://orcid.org/0000-0002-3166-1413

REFERENCES

1. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q*. 2004;82(4):661-687. https://doi.org/10.1111/j.0887-378x.2004.00327.x
2. Tunis SR, Benner J, McClellan M. Comparative effectiveness research: policy context, methods development and research infrastructure. *Statist Med*. 2010;29(19):1963-1976. https://doi.org/10.1002/sim.3818
3. Heckman JJ. Micro data, heterogeneity, and the evaluation of public policy: nobel lecture. *J Political Econ*. 2001;109(4):673-748. https://doi.org/10.1086/322086
4. Hamburg MA, Collins FS. The path to personalized medicine. *N Engl J Med*. 2010;363(4):301-304. https://doi.org/10.1056/nejmp1006304
5. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793-795. https://doi.org/10.1056/nejmp1500523
6. Zhang B, Tsiatis AA, Laber EB, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012;68(4):1010-1018. https://doi.org/10.1111/j.1541-0420.2012.01763.x
7. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *Ann Stat*. 2011;39(2):1180-1210. https://doi.org/10.1214/10-aos864
8. Brinkley J, Tsiatis A, Anstrom KJ. A generalized estimator of the attributable benefit of an optimal treatment regime. *Biometrics*. 2010;66(2):512-522. https://doi.org/10.1111/j.1541-0420.2009.01282.x
9. Xu Y, Yu M, Zhao Y-Q, Li Q, Wang S, Shao J. Regularized outcome weighted subgroup identification for differential treatment effects. *Biometrics*. 2015;71(3):645-653. https://doi.org/10.1111/biom.12322
10. Zhang B, Tsiatis AA, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. *Statistics*. 2012;1(1):103-114. https://doi.org/10.1002/sta.411
11. Laber EB, Zhao YQ. Tree-based methods for individualized treatment regimes. *Biometrika*. 2015;102(3):501-514. https://doi.org/10.1093/biomet/asv028
12. Su X, Tsai C-L, Wang H, Nickerson DM, Li B. Subgroup analysis via recursive partitioning. *J Mach Learn Res*. 2009;10:141-158. https://doi.org/10.2139/ssrn.1341380
13. Zeileis A, Hothorn T, Hornik K. Model-based recursive partitioning. *J Comput Graph Stat*. 2008;17(2):492-514. https://doi.org/10.1198/106186008x319331
14. Loh W-Y, He X, Man M. A regression tree approach to identifying subgroups with differential treatment effects. *Statist Med*. 2015;34(11):1818-1833. https://doi.org/10.1002/sim.6454
15. Dusseldorp E, Van Mechelen I. Qualitative interaction trees: a tool to identify qualitative treatment–subgroup interactions. *Statist Med*. 2014;33(2):219-237. https://doi.org/10.1002/sim.5933
16. Foster JC, Taylor JMG, Ruberg SJ. Subgroup identification from randomized clinical trial data. *Stat Med*. 2011;30(24):2867-2880. https://doi.org/10.1002/sim.4322
17. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66(5):688-701. https://doi.org/10.1037/h0037350
18. Laber EB, Lizotte DJ, Qian M, Pelham WE, Murphy SA. Dynamic treatment regimes: Technical challenges and applications. *Electron J Stat*. 2014;8(1):1225-1272.
19. Boyer BE, Doove LL, Geurts HM, Prins PJM, Van Mechelen I, Van der Oord S. Qualitative treatment-subgroup interactions in a randomized clinical trial of treatments for adolescents with ADHD: exploring what cognitive-behavioral treatment works for whom. *PLOS ONE*. 2016;11(3):e0150698. https://doi.org/10.1371/journal.pone.0150698
20. Gueorguieva R, Wu R, Tsai W-M, et al. An analysis of moderators in the combine study: identifying subgroups of patients who benefit from acamprosate. *European Neuropsychopharmacology*. 2015;25(10):1586-1599. https://doi.org/10.1016/j.euroneuro.2015.06.006
21. Hou J, Seneviratne C, Su X, et al. Subgroup identification in personalized treatment of alcohol dependence. *Alcohol Clin Exp Res*. 2015;39(7):1253-1259. https://doi.org/10.1111/acer.12759
22. Liu W, Zhang Z, Nie L, Soon G. A case study in personalized medicine: rilpivirine versus efavirenz for treatment-naive HIV patients. *J Am Stat Assoc*. 2014;112(520):1381-1392. https://doi.org/10.1080/01621459.2017.1280404
23. Schneider LS, Frangakis C, Drye LT, et al. Heterogeneity of treatment response to citalopram for patients with alzheimer's disease with aggression or agitation: the CitAd randomized clinical trial. *Am J Psychiat*. 2016;173(5):465-472. https://doi.org/10.1176/appi.ajp.2015.15050648
24. Seibold H, Zeileis A, Hothorn T. Individual treatment effect prediction for amyotrophic lateral sclerosis patients. *Stat Methods Med Res*. 2018;27(10):3104-3125.
25. Setyawan J, Yang H, Cheng D, et al. Developing a risk score to guide individualized treatment selection in attention deficit/hyperactivity disorder. *Value Health*. 2015;18(6):824-831. https://doi.org/10.1016/j.jval.2015.06.005
26. Lipkovich I, Dmitrienko A, D'Agostino RB Sr. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. *Stat Med*. 2017;36(1):136-196. https://doi.org/10.1002/sim.7064
27. Laber EB, Qian M. Evaluating personalized treatment regimes. In: Gatsonis C, Morton SC, eds. *Methods in Comparative Effectiveness Research*. Boca Raton, FL: CRC Press; 2017:867-871.

28. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962-973. https://doi.org/10.1111/j.1541-0420.2005.00377.x
29. Cao W, Tsiatis AA, Davidian M. Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika*. 2009;96(3):723-734. https://doi.org/10.1093/biomet/asp033
30. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer; 2009.
31. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc*. 1983;78(382):316-331. https://doi.org/10.1080/01621459.1983.10477973
32. Efron B, Tibshirani R. Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc*. 1997;92(438):548-560. https://doi.org/10.1080/01621459.1997.10474007
33. Kim J-H. Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap. *Comput Stat Data Anal*. 2009;53(11):3735-3745. https://doi.org/10.1016/j.csda.2009.04.009
34. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*. 2005;21(15):3301-3307. https://doi.org/10.1093/bioinformatics/bti499
35. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95); 1995; Montreal, Canada.
36. Sun L, Bull SB. Reduction of selection bias in genomewide studies by resampling. *Genetic Epidemiology*. 2005;28(4):352-367. https://doi.org/10.1002/gepi.20068
37. Zhang P. On the estimation of prediction errors in linear regression models. *Ann Inst Stat Math*. 1993;45(1):105-111. https://doi.org/10.1007/bf00773671
38. Luedtke AR, van der Laan MJ. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Ann Stat*. 2016;44(2):713-742.
39. Chakraborty B, Laber EB, Zhao YQ. Inference about the expected performance of a data-driven dynamic treatment regime. *Clinical Trials*. 2014;11(4):408-417.
40. van der Laan MJ, Luedtke AR. Targeted learning of the mean outcome under an optimal dynamic treatment rule. *J Causal Inference*. 2015;3(1):61-95.
41. Holloway S, Laber E, Linn K, Zhang B, Davidian M, Tsiatis A. Package 'DynTxRegime'. 2015.
42. Therneau T, Atkinson B, Ripley B. Package 'rpart'. 2015.
43. Rush AJ, Trivedi MH, Stewart JW, et al. Combining medications to enhance depression outcomes (CO-MED): acute and long-term outcomes of a single-blind randomized study. *Am J Psychiat*. 2011;168(7):689-701.
44. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*. 2003;54(5):573-583.

How to cite this article: Sies A, Van Mechelen I. Estimating the quality of optimal treatment regimes. *Statistics in Medicine*. 2019;38:4925–4938. https://doi.org/10.1002/sim.8342