Received: 5 March 2018 | Revised: 11 September 2018 | Accepted: 9 November 2018

DOI: 10.1002/sim.8054

RESEARCH ARTICLE

WILEY Statistics in Medicine

Machine learning methods for leveraging baseline covariate information to improve the efficiency of clinical trials

Zhiwei Zhang | Shujie Ma

Department of Statistics, University of California, Riverside, California

Correspondence
Zhiwei Zhang, Department of Statistics, University of California, Riverside, CA.
Email: zhiwei.zhang@ucr.edu

Clinical trials are widely considered the gold standard for treatment evaluation, and they can be highly expensive in terms of time and money. The efficiency of clinical trials can be improved by incorporating information from baseline covariates that are related to clinical outcomes. This can be done by modifying an unadjusted treatment effect estimator with an augmentation term that involves a function of covariates. The optimal augmentation is well characterized in theory but must be estimated in practice. In this article, we investigate the use of machine learning methods to estimate the optimal augmentation. We consider and compare an indirect approach based on an estimated regression function and a direct approach that aims directly to minimize the asymptotic variance of the treatment effect estimator. Theoretical considerations and simulation results indicate that the direct approach is generally preferable over the indirect approach. The direct approach can be implemented using any existing prediction algorithm that can minimize a weighted sum of squared prediction errors. Many such prediction algorithms are available, and the super learning principle can be used to combine multiple algorithms into a super learner under the direct approach. The resulting direct super learner has a desirable oracle property, is easy to implement, and performs well in realistic settings. The proposed methodology is illustrated with real data from a stroke trial.

KEYWORDS

augmentation, asymptotic efficiency, cross-validation, influence function, semiparametric theory, super learner

1 | INTRODUCTION

Randomized clinical trials are widely considered the gold standard for comparing two or more medical treatments, one of which may be placebo or no treatment. Analysis of clinical trial data is essentially a comparison of different treatments with respect to clinically important outcome variables. Depending on the scientific context, the effect of one treatment versus another may be measured by a difference in mean or proportion, a risk ratio or odds ratio for a binary outcome, or a difference in survival probability or mean restricted survival time for a survival outcome subject to censoring. Traditionally, such an effect measure is estimated using its empirical counterpart (eg, difference in sample mean or proportion), although special techniques may be required to deal with censoring and/or missingness.

In addition to treatment and outcome data, clinical trials usually collect information on a large number of baseline covariates, such as demographics, clinical measurements, and biomarkers. Some of these covariates may be related to

outcome variables in one or both treatment groups, and it makes sense to consider how to use such covariate data to improve the efficiency of treatment comparisons. In practice, this is commonly done under a regression model that relates the outcome of interest to treatment and relevant baseline covariates. Such a regression model describes the conditional treatment effect (ie, the treatment effect in each subpopulation defined by the covariates), which may differ from the marginal treatment effect due to possible noncollapsibility.[1] For example, with a binary outcome, the conditional odds ratio estimated from a logistic regression model is generally different from the marginal odds ratio, unless the former is identically equal to one. A notable exception here is the analysis of covariance model, which implies that the treatment difference in mean outcome is the same with or without conditioning on the baseline covariates in the model. In reality, an unsaturated regression model is likely to be misspecified, which may lead to inconsistent estimation of the marginal or conditional effect, although some exceptions have been noted,[2] and some simple modifications proposed,[3,4] for robust estimation of the marginal effect.

Suppose we are interested in estimating a marginal treatment effect, as is the case in most clinical trials, and would like to incorporate baseline covariate data without changing the estimand or relying on correct specification of a regression model for consistent estimation. This problem has been considered by a number of authors.[2-9] Tsiatis et al[2] and Zhang et al[5] proposed an augmentation approach based on semiparametric theory. Their approach is to augment an unadjusted estimator or estimating function, which does not involve covariate data, with a term based on treatment and an arbitrary function of the covariates. The optimal augmentation, which leads to the smallest variance of the augmented estimator, has also been characterized. To take advantage of the optimality result, one may use a working regression model to estimate the optimal augmentation and use the estimated augmentation to construct an augmented estimator. The resulting estimator is consistent and asymptotically normal regardless of model (in)correctness. If the working model is correct, the estimator attains the minimum variance in the class of augmented estimators. Rubin and van der Laan[6] provided an improvement based on empirical efficiency maximization in estimating the optimal augmentation. Tian et al[7] considered the use of a lasso penalty in estimating the optimal augmentation and proposed a cross-validation procedure for choosing the tuning parameter and for making valid inference. Other approaches to this problem include targeted maximum likelihood estimation (TMLE)[3,9] and inverse probability weighting.[8]

In this article, we focus on the augmentation approach and consider how to use machine learning methods to estimate or approximate the optimal augmentation. Compared to parametric regression models, machine learning methods offer more flexibility and hence a better chance to approximate the optimal augmentation well, and their slower-than-parametric rate of convergence does not incur a cost in terms of the asymptotic variance of the treatment effect estimator. We consider two ways to use machine learning methods in the present context, ie, an indirect approach based on an estimated regression function and a direct approach that aims directly to minimize the asymptotic variance of the treatment effect estimator. The direct approach is better aligned with our analytical objective, performs better in simulation experiments, and therefore is recommended over the indirect approach. The optimization in the direct approach can be formulated as minimizing a weighted sum of squared prediction errors, and thus can be easily implemented using any existing prediction algorithm that can accommodate such a loss function. With a large variety of machine learning methods available,[10] it may be difficult to choose the best-performing method for a given application, especially in the design stage. This is unnecessary, however, according to the super learning principle, which allows us to combine a collection of learning methods into a super learner with a desirable oracle property.[11,12] We apply the super learning principle to the present problem and develop a simple and easy-to-implement super learner under the direct approach. Simulation results show that the direct super learner performs well in realistic settings.

The rest of this paper is organized as follows. In the next section, we describe the various methods. In Section 3, we present numerical results (simulation and data analysis) based on a stroke study. Concluding remarks are given in Section 4. Some technical details are relegated to an appendix.

2 | METHODOLOGY

2.1 | Preliminaries

Consider a randomized clinical trial comparing an experimental treatment ($T = 1$) with a control treatment ($T = 0$). Let Y denote the outcome variable of interest and $\boldsymbol{X}$ a vector of baseline covariates that may be related to Y in one or both treatment groups. Note that T and $\boldsymbol{X}$ are independent of each other because of randomization. Let θ be a marginal effect measure of $T = 1$ versus $T = 0$, ie, some contrast of the conditional distributions of Y given $T = 1$ versus $T = 0$. A

common example of θ is the difference in mean

$$\theta = \mu_1 - \mu_0, \tag{1}$$

where $\mu_t = \mathrm{E}(Y|T = t)$, $t = 0, 1$. Assuming that θ has been chosen appropriately based on clinical considerations, we consider the statistical problem of how to estimate θ consistently and efficiently.

Let $(\boldsymbol{X}_i, T_i, Y_i)$, $i = 1, \ldots, n$, denote the observed data based on n independent subjects in the trial; these are conceptualized as independent copies of $(\boldsymbol{X}, T, Y)$. Suppose an unadjusted estimator of θ, say $\widetilde{\theta}$, can be obtained from (T_i, Y_i), $i = 1, \ldots, n$. For example, if θ is given by (1), then a natural choice of $\widetilde{\theta}$ is the difference in sample mean

$$\widetilde{\theta} = \overline{Y}_1 - \overline{Y}_0, \tag{2}$$

where $\overline{Y}_t = \{\sum_{i=1}^n I(T_i = t)\}^{-1} \sum_{i=1}^n I(T_i = t) Y_i$ and $I(\cdot)$ is the indicator function. We do not assume specific choices of $\widetilde{\theta}$ and only require that $\widetilde{\theta}$ be consistent, regular, and asymptotically linear, ie,

$$\sqrt{n}(\widetilde{\theta} - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(T_i, Y_i) + o_p(1), \tag{3}$$

where ψ is known as the influence function. This requirement is met by many estimators that are commonly used in clinical trials. For example, with $\widetilde{\theta}$ defined by (2), expression (3) holds with

$$\psi(T, Y) = \frac{T(Y - \mu_1)}{\pi} - \frac{(1 - T)(Y - \mu_0)}{1 - \pi}, \tag{4}$$

where $\pi = \mathrm{P}(T = 1)$. Section 3 of Tian et al[7] gives more examples. Expression (3) implies that $\widetilde{\theta}$ is asymptotically normal and justifies inference procedures based on asymptotic normality. However, $\widetilde{\theta}$ may be inefficient because it does not make use of the available information in $\boldsymbol{X}$.

Motivated by the semiparametric theory,[13,14] an augmentation approach has been developed to incorporate covariate information into the estimation of θ.[2,5,7] Following Tsiatis et al[2] and Tian et al,[7] an augmented estimator of θ may be obtained as

$$\widehat{\theta}(a) = \widetilde{\theta} - \frac{1}{n} \sum_{i=1}^n (T_i - \pi) a(\boldsymbol{X}_i), \tag{5}$$

where $a(\boldsymbol{X})$ is an arbitrary function of $\boldsymbol{X}$ such that $\mathrm{E}\{a(\boldsymbol{X})^2\} < \infty$. For example, $a(\boldsymbol{X})$ may be taken to be a linear function of $\boldsymbol{X}$. For any fixed function a, $\widehat{\theta}(a)$ is consistent for θ and asymptotically normal with asymptotic variance $\mathrm{var}\{\psi(T, Y) - (T - \pi)a(\boldsymbol{X})\}$. The asymptotic variance is minimized by taking $a(\boldsymbol{x})$ equal to

$$a_{\mathrm{opt}}(\boldsymbol{x}) = \eta(1, \boldsymbol{x}) - \eta(0, \boldsymbol{x}), \tag{6}$$

where $\eta(t, \boldsymbol{x}) = \mathrm{E}\{\psi(T, Y)|T = t, \boldsymbol{X} = \boldsymbol{x}\}$. The corresponding estimator, ie, $\widehat{\theta}(a_{\mathrm{opt}})$, is the most efficient among all estimators of form (5). In a nonparametric model where the only constraint on the joint distribution of $(\boldsymbol{X}, T, Y)$ is $\mathrm{P}(T = 1|\boldsymbol{X}) = \pi$ (due to randomization), if $\widetilde{\theta}$ is a nonparametric estimator, then $\widehat{\theta}(a_{\mathrm{opt}})$ is the most efficient nonparametric estimator that is regular and asymptotically linear.[13,14] For nonparametric estimation of (1) with $\widetilde{\theta}$ given by (2), the optimal augmentation is

$$a_{\mathrm{opt}}(\boldsymbol{x}) = \frac{m_1(\boldsymbol{x}) - \mu_1}{\pi} + \frac{m_0(\boldsymbol{x}) - \mu_0}{1 - \pi},$$

where $m_t(\boldsymbol{x}) = \mathrm{E}(Y|T = t, \boldsymbol{X} = \boldsymbol{x})$, $t = 0, 1$, and the influence function of $\widehat{\theta}(a_{\mathrm{opt}})$ is equal to the efficient influence function for nonparametric estimation of θ, ie,

$$\psi_{\mathrm{eff}}(\boldsymbol{X}, T, Y) = \frac{TY}{\pi} - \frac{(1 - T)Y}{1 - \pi} - (T - \pi) \left\{ \frac{m_1(\boldsymbol{x})}{\pi} + \frac{m_0(\boldsymbol{x})}{1 - \pi} \right\} - \theta.$$

In this case, both the augmentation approach and the TMLE approach[3,9] aim to solve the efficient influence function equation, ie, $\sum_{i=1}^n \psi_{\mathrm{eff}}(\boldsymbol{X}_i, T_i, Y_i) = 0$.

2.2 | Learning the optimal augmentation

In reality, the optimal function a_{opt} is unknown and must be estimated from data. Let $\widehat{a}$ be a generic estimator with probability limit a^*. We recognize that, in general, a^* is not equal to a_{opt}. Under mild regularity conditions, we show in Appendix that

$$\sqrt{n} \left\{ \widehat{\theta}(\widehat{a}) - \theta \right\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{\psi(T_i, Y_i) - (T_i - \pi) a^*(\boldsymbol{X}_i)\} + o_p(1), \tag{7}$$

so that $\widehat{\theta}(\widehat{a})$ is asymptotically equivalent to $\widehat{\theta}(a^*)$. In particular, the asymptotic variance of $\widehat{\theta}(\widehat{a})$ depends on the limit but not on the variability or convergence rate of $\widehat{a}$. Motivated by this observation, we consider estimating a_{opt} using a flexible machine learning method in the hope that $\widehat{a}$ will converge to a limit close to a_{opt} (in a sense that will be defined later), possibly at the expense of a slower-than-parametric convergence rate. At the same time, since (7) is an asymptotic result, we should also pay attention to the finite-sample performance of $\widehat{\theta}(\widehat{a})$ and optimize $\widehat{a}$ accordingly. We will consider different ways of applying machine learning methods to the present setting, starting with an obvious and straightforward approach.

2.2.1 | An indirect approach

Equation (6) suggests that an estimate of a_{opt} may be obtained from an estimate of the regression function $\eta(t,\boldsymbol{x}) = \mathrm{E}\{\psi(T,Y)|T = t,\boldsymbol{X} = \boldsymbol{x}\}$. This latter regression problem is complicated by the fact that $\psi(T,Y)$, the "response variable", typically depends on unknown parameters and is not fully observed. For instance, with $\widetilde{\theta}$ defined by (2), $\psi(T,Y)$ given by (4) involves μ_1 and μ_0, which are unknown but easily estimated. An estimate of $\psi(T,Y)$, say $\widehat{\psi}(T,Y)$, may be obtained by replacing the unknown parameters in $\psi(T,Y)$ with empirical estimates (see section 3 in the work of Tian et al[7] for specific examples). Once $\widehat{\psi}$ is available, we can apply any existing machine learning method to the "data" $(\widehat{\psi}(T_i,Y_i), T_i, \boldsymbol{X}_i)$, $i = 1,\ldots,n$, treating $(T_i,\boldsymbol{X}_i)$ as input and $\widehat{\psi}(T_i,Y_i)$ as output, to obtain an estimate of $\eta(t,\boldsymbol{x})$, say $\widehat{\eta}(t,\boldsymbol{x})$. In the special case of (4), where $\psi(T,Y)$ is linear in Y, we could estimate $\mathrm{E}(Y \mid T,\boldsymbol{X})$ using a machine learning method and substitute the estimate into (4) (together with $\overline{Y}_t$ for μ_t, $t = 0,1$) to obtain $\widehat{\eta}(t,\boldsymbol{x})$. Once $\widehat{\eta}(t,\boldsymbol{x})$ is available, the corresponding estimate of $a_{\text{opt}}(\boldsymbol{x})$ is

$$\widehat{a}_{\text{ind}}(\boldsymbol{x}) = \widehat{\eta}(1,\boldsymbol{x}) - \widehat{\eta}(0,\boldsymbol{x}),$$

where the subscript "ind" emphasizes that this is an indirect estimate (through $\widehat{\eta}$).

There is a great variety of machine learning methods,[10] some of which may be more appropriate than others for a given application. Without knowing or assuming which methods perform best, we can simply combine all candidate methods into a super learner with a desirable oracle property.[11,12] Let $\{\widehat{\eta}_k, k = 1,\ldots,K\}$ be a collection of K candidate learners, where K may increase with n in a polynomial fashion. The super learner is a linear combination of the $\widehat{\eta}_k$ with coefficients determined using a cross-validation procedure. Let the sample be partitioned into J subsamples that are roughly equal in size. For each $j \in \{1,\ldots,J\}$, we use the jth subsample as a validation sample and combine the other subsamples into a training sample. Let $\widehat{\psi}^{(-j)}$ and $\widehat{\eta}_k^{(-j)}$ be obtained from the training sample using the same methods for obtaining $\widehat{\psi}$ and $\widehat{\eta}_k$. Then, we can find the coefficients in the super learner as

$$(\widehat{\alpha}_1,\ldots,\widehat{\alpha}_K) = \arg\min_{(\alpha_1,\ldots,\alpha_K)} \sum_{i=1}^{n} \left\{ \widehat{\psi}^{(-j_i)}(T_i,Y_i) - \sum_{k=1}^{K} \alpha_k \widehat{\eta}_k^{(-j_i)}(T_i,\boldsymbol{X}_i) \right\}^2,$$

where j_i is the index of the subsample that contains subject i. Theoretical considerations suggest that α_k be constrained to a bounded set in the aforementioned minimization, and practical considerations lead to the following constraints: $\sum_{k=1}^{K} \alpha_k = 1$, $\alpha_k \geq 0\ \forall k$ (see chapter 3 in the work of van der Laan and Rose[12]). In other words, we perform a linear regression of $\widehat{\psi}^{(-j_i)}(T_i,Y_i)$ on $\{(\widehat{\eta}_k^{(-j_i)}(T_i,\boldsymbol{X}_i), k = 1,\ldots,K\}$ without an intercept using constrained least squares. The super learner estimate of η is just $\widehat{\eta}_{\text{sl}} = \sum_{k=1}^{K} \widehat{\alpha}_k \widehat{\eta}_k$, which can be substituted into (6) to obtain $\widehat{a}_{\text{sl.ind}}(\boldsymbol{x}) = \widehat{\eta}_{\text{sl}}(0,\boldsymbol{x}) + \widehat{\eta}_{\text{sl}}(1,\boldsymbol{x})$, where the subscript "sl.ind" indicates that this is an indirect super learner.

The indirect nature of this approach is unsatisfactory. The risk function that $\widehat{\eta}_{\text{sl}}$ attempts to minimize is

$$h \mapsto \mathrm{E}\{\psi(T,Y) - h(T,\boldsymbol{X})\}^2, \tag{8}$$

where the expectation is with respect to the joint distribution of $(\boldsymbol{X},T,Y)$. The global minimizer of (8) is just the true regression function η, which corresponds to a_{opt} through (6). If $\widehat{\eta}$ converges to η, say under a correctly specified model $\mathcal{H}$, the corresponding $\widehat{a}_{\text{ind}} = \widehat{\eta}(1,\cdot) - \widehat{\eta}(0,\cdot)$ converges to a_{opt}. If the model $\mathcal{H}$ is misspecified, then $\widehat{\eta}$ is expected to converge to

$$h^* = \arg\min_{h\in\mathcal{H}} \mathrm{E}\{\psi(T,Y) - h(T,\boldsymbol{X})\}^2,$$

and $\widehat{a}_{\text{ind}}$ to $a^*_{\text{ind}} = h^*(1,\cdot) - h^*(0,\cdot)$, which does not have a clear interpretation in terms of estimating θ. Although we know that

$$a^*_{\text{ind}} \in \mathcal{A}_{\mathcal{H}} = \{h(1,\cdot) - h(0,\cdot) : h \in \mathcal{H}\},$$

we do not know (without further information) that $\widehat{\theta}(\widehat{a}_{\text{ind}})$ attains the minimum asymptotic variance among the estimators $\widehat{\theta}(a)$, $a \in \mathcal{A}_{\mathcal{H}}$. Therefore, there is a mismatch between the risk function (8) and the objective of estimating θ efficiently.

As a result, an indirect augmented estimator $\widehat{\theta}(\widehat{a}_{\text{ind}})$ is not necessarily more efficient than the unadjusted estimator $\widetilde{\theta}$,[6] and the estimator $\widehat{\theta}(\widehat{a}_{\text{sl.ind}})$ based on the indirect super learner is not necessarily more efficient that the estimators $\widehat{\theta}(\widehat{a}_k)$, $k = 1, \ldots, K$, based on individual learners. These problems do arise in realistic situations, as we shall see in a simulation study.

2.2.2 | A direct approach

For efficient estimation of θ, the relevant quantity to minimize is the asymptotic variance of $\widehat{\theta}(a)$, ie,

$$\mathrm{E}\{\psi(T, Y) - (T - \pi)a(\boldsymbol{X})\}^2. \tag{9}$$

Treating this as a risk function, the direct approach is to directly minimize a sample version of (9) with respect to a. For an individual learner, this means choosing $\widehat{a}$ to minimize the sum of squares

$$\sum_{i=1}^{n} \left\{\widehat{\psi}(T_i, Y_i) - (T_i - \pi)a(\boldsymbol{X}_i)\right\}^2 \tag{10}$$

or a penalized version of it. This has been suggested by Rubin and van der Laan[6] in the case where a is constrained to a parametric family. Another example of this, with a assumed linear and with a lasso penalty, has been considered by Tian et al.[7] More generally, we note that (10) can be rewritten as a weighted sum of squares

$$\sum_{i=1}^{n} (T_i - \pi)^2 \left\{ \frac{\widehat{\psi}(T_i, Y_i)}{(T_i - \pi)} - a(\boldsymbol{X}_i) \right\}^2,$$

where $(T_i - \pi)^2$ is regarded as a weight and $\widehat{\psi}(T_i, Y_i)/(T_i - \pi)$ as a response variable. This implies that (10) can be minimized by any prediction algorithm that is able to minimize a weighted sum of squared prediction errors. When $\pi = 1/2$ (a common case in clinical trials), the weight is the same for all subjects, and any prediction algorithm that can minimize an unweighted sum of squared prediction errors can be used to estimate a_{opt}.

An important advantage of this approach over the indirect approach is that the risk function (9) is better aligned than (8) is with our analytical objective of estimating θ efficiently. A local minimizer of (9), say

$$a^*_{\text{dir}} = \arg\min_{a \in \mathcal{A}} \mathrm{E}\{\psi(T, Y) - (T - \pi)a(\boldsymbol{X})\}^2,$$

corresponds directly to a locally optimal estimator of θ, ie,

$$\operatorname{avar}\left\{\widehat{\theta}\left(a^*_{\text{dir}}\right)\right\} = \min_{a \in \mathcal{A}} \operatorname{avar}\left\{\widehat{\theta}(a)\right\},$$

where avar denotes asymptotic variance. In particular, if $\widehat{a}_{\text{dir}}$ minimizes (10) over a class of functions that include $a(\boldsymbol{x}) \equiv 0$, then $\widehat{\theta}(\widehat{a}_{\text{dir}})$ is asymptotically at least as efficient as the unadjusted estimator $\widetilde{\theta} = \widehat{\theta}(0)$.

Here, again, we can combine multiple candidate learners into a super learner using the general super learning principle (see section 3.5 in the work of van der Laan and Rose[12]). Let $\{\widehat{a}_k, k = 1, \ldots, K\}$ be a collection of K candidate learners. Suppose the sample has been partitioned into J subsamples as before, and let $\widehat{\psi}^{(-j)}$ and $\widehat{a}_k^{(-j)}$ be obtained from the jth training sample (the whole sample minus the jth subsample) using the same methods for obtaining $\widehat{\psi}$ and $\widehat{a}_k$. The direct super learner of a_{opt} is given by $\widehat{a}_{\text{sl.dir}} = \sum_{k=1}^{K} \widehat{\beta}_k \widehat{a}_k$ with

$$(\widehat{\beta}_1, \ldots, \widehat{\beta}_K) = \arg\min_{(\beta_1, \ldots, \beta_K)} \sum_{i=1}^{n} \left\{ \widehat{\psi}^{(-j_i)}(T_i, Y_i) - (T_i - \pi) \sum_{k=1}^{K} \beta_k \widehat{a}_k^{(-j_i)}(\boldsymbol{X}_i) \right\}^2, \tag{11}$$

where j_i is the index of the subsample containing subject i. This minimization can be carried out as a linear regression of $\widehat{\psi}^{(-j_i)}(T_i, Y_i)$ on $\{(T_i - \pi)\widehat{a}_k^{(-j_i)}(\boldsymbol{X}_i), k = 1, \ldots, K\}$ without an intercept using constrained least squares. Here, we could apply the same constraints as in Section 2.2.1, requiring β_k to form a probability vector; however, based on our numerical experience, we choose to use a less stringent set of constraints, ie, $|\beta_k| \leq B \; \forall k$ for some constant $B > 1$. Under quite general conditions, $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ is asymptotically equivalent to an oracle "estimator" based on the best (in the sense of minimizing (9) on a sample of size $(J - 1)n/J$) linear combination of $\widehat{a}_k$ (subject to any constraints on the β_k) (see theorem 1 in the work of van der Laan and Dudoit[15] for precise statements of the results and the technical conditions required). In particular, $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ is asymptotically at least as efficient as $\widehat{\theta}(\widehat{a}_k)$ for any k, as long as the configuration $(\beta_k = 1, \beta_{k'} = 0 \; \forall k')$ is permissible in (11).

As an example, suppose a linear model is assumed for a and estimated by minimizing (10) with a lasso penalty, and each $\widehat{a}_k$ is obtained using a different tuning parameter, as in the work of Tian et al.[7] With a finite collection of tuning parameter values, the estimator of a_{opt} proposed by Tian et al[7] is $\widehat{a}_{\widehat{k}}$, where

$$\widehat{k} = \arg\min_k \sum_{i=1}^{n} \left\{ \widehat{\psi}^{(-j_i)}(T_i, Y_i) - (T_i - \pi)\widehat{a}_k^{(-j_i)}(\boldsymbol{X}_i) \right\}^2.$$

This $\widehat{a}_{\widehat{k}}$ is known as a discrete super learner based on the $\widehat{a}_k$ and is asymptotically equivalent to the best of the $\widehat{a}_k$ in the sense of minimizing (9)(see chapter 3 in the work of van der Laan and Rose[12]). However, the oracle property of $\widehat{a}_{\widehat{k}}$ is not as strong as that of the "true" super learner $\widehat{a}_{\text{sl.dir}}$ based on $\widehat{a}_k$ because the minimization in (11) is over a larger class of functions than the search for $\widehat{k}$. Consequently, $\widehat{\theta}(\widehat{a}_{\widehat{k}})$ may be expected to underperform $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ asymptotically.

In a similar spirit, Balzer et al[9] used cross-validation to choose a working model for outcome regression in the TMLE approach. Their approach is "direct" in the sense that model selection is based on the principle of empirical efficiency maximization.[6] Their estimator has the flavor of a discrete super learner in that cross-validation is only used to select one model and not to combine several candidate models.

2.3 | Variance estimation

For any estimator $\widehat{a}$ that satisfies (7), the asymptotic variance of $\widehat{\theta}(\widehat{a})$ is consistently estimated by

$$\frac{1}{n}\sum_{i=1}^{n} \left\{ \widehat{\psi}(T_i, Y_i) - (T_i - \pi)\widehat{a}(\boldsymbol{X}_i) \right\}^2. \tag{12}$$

Although this variance estimator is asymptotically valid, it may suffer from a downward resubstitution bias in finite samples, especially if $\widehat{a}$ is chosen to minimize (12). This may or may not be a serious issue, depending on how hard $\widehat{a}$ attempts to minimize (12). If necessary, a cross-validation procedure can be used to remove the resubstitution bias. Let the sample be partitioned into L subsamples that are roughly equal in size, and let $\widehat{\psi}^{(-l)}$ and $\widehat{a}^{(-l)}$ be obtained from the lth subsample using the same method for obtaining $\widehat{\psi}$ and $\widehat{a}$. Then, the cross-validated estimate of $\text{avar}\{\widehat{\theta}(\widehat{a})\}$ is given by

$$\frac{1}{n}\sum_{i=1}^{n} \left\{ \widehat{\psi}^{(-l_i)}(T_i, Y_i) - (T_i - \pi)\widehat{a}^{(-l_i)}(\boldsymbol{X}_i) \right\}^2,$$

where l_i is the index of the subsample that includes subject i. The partitioning used here may or may not be the same as the ones used in Section 2.2. When a super learner is developed from a collection of learning methods, the same partitioning used to find the coefficients for the super learner can be used to obtain cross-validated variance estimates for the individual learning methods. For example, in the direct approach, a cross-validated estimate of $\text{avar}\{\widehat{\theta}(\widehat{a}_k)\}$ may be obtained as

$$\frac{1}{n}\sum_{i=1}^{n} \left\{ \widehat{\psi}^{(-j_i)}(T_i, Y_i) - (T_i - \pi)\widehat{a}_k^{(-j_i)}(\boldsymbol{X}_i) \right\}^2,$$

because $\widehat{a}_k$ does not depend on the partitioning. This is not true for $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ because the super learner $\widehat{a}_{\text{sl.dir}}$ depends on the aforementioned partitioning. The minimum in (11) is not fully cross-validated for $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ as it does not account for the estimation of the coefficients β_k. To obtain a fully cross-validated estimate of $\text{avar}\{\widehat{\theta}(\widehat{a}_{\text{sl.dir}})\}$, one should perform an external cross-validation in which all steps in obtaining $\widehat{a}_{\text{sl.dir}}$ are repeated for each training sample. It is often difficult to decide whether or not the resubstitution bias is a serious issue for a given augmented estimator. To be safe, we recommend always using a fully cross-validated variance estimate to make inference about θ.

3 | NUMERIAL RESULTS

3.1 | Background

We now illustrate and evaluate the proposed methodology in the context of the National Institute of Neurological Disorders and Stroke (NINDS) Recombinant Tissue Plasminogen Activator (rt-PA) for Acute Ischemic Stroke Treatment Trial.[16]

In this randomized placebo-controlled double-blind trial, 624 patients with acute ischemic stroke were randomized 1 : 1 to rt-PA or placebo within three hours of stroke onset, and their clinical outcomes were evaluated at 24 hours and three months post-treatment. The study consisted of two (temporally ordered) parts, with 291 patients enrolled in part 1 and 333 in part 2. The main difference between the two parts was that the time from stroke onset to treatment (TOT) was within 90 minutes in part 1 and between 90 minutes and 180 minutes in part 2. Except for that difference, the patients in the two parts were treated and evaluated in the same manner, and therefore will be combined in our simulation and analysis.

As in previous analyses of the NINDS rt-PA trial data,[16-18] we will compare treatments with respect to clinical outcomes at three months, as measured by the Barthel Index, the Modified Rankin Scale, the Glasgow Outcome Scale, and the National Institutes of Health Stroke Scale (NIHSS). The Barthel Index measures the ability to perform activities of daily living and yields a score that ranges from 0 to 100 (complete independence). The Modified Rankin Scale is a simplified overall assessment of function, with a score that ranges from 0 (absence of symptoms) to 5 (severe disability). The Glasgow Outcome Scale is a global assessment of function, with a score that ranges from 1 (good recovery) to 5 (death). The NIHSS is a 42-point scale with 0 indicating no stroke symptoms and higher scores for more severe symptoms. Scores of 95 or 100 on the Barthel Index, 0 or 1 on the Modified Rankin Scale and the NIHSS, and 1 on the Glasgow Outcome Scale are considered to indicate treatment success.[16] For each instrument, the efficacy of rt-PA relative to placebo is measured by the between-group differences (rt-PA minus placebo) in the mean score and the probability of success. For ease of interpretation and comparison, the mean score analyses are performed on standardized scores in the unit interval, obtained by dividing the original scores by the maximum score for each instrument.

Our objective here is to improve the efficiency of these treatment comparisons by utilizing the following baseline covariates: clinical site, study part, continuous TOT, baseline NIHSS, age, preexisting disability, and history of diabetes. These covariates were chosen by Ingall et al[17] in their regression analysis and will be taken as given in applying the proposed methodology. Six patients with missing covariate data will be excluded in this exercise. The super learning principle will be used to combine the following prediction algorithms (identified as R packages): glm (generalized linear models), glmnet (glm with elastic net[19]), gam (generalized additive models[20]), rpart (recursive partitioning and regression trees[21]), and randomForest.[22] The super learner library includes two versions of glmnet, with $\alpha = 1$ for lasso and 0 for ridge, and two versions of randomForest, with $m = 2$ or 3 (number of covariates to be chosen randomly for each tree). In addition, the library will include a trivial predictor based on the sample mean to produce an "intercept" in the super learner. The super learner is implemented with five-fold cross-validation, and its variance estimate is obtained from another (external) five-fold cross-validation.

3.2 | Simulation

The different methods are compared in a simulation study, which can be described as follows. We consider a population of patients with potential outcomes $Y(0) \equiv Y(1)$ satisfying Fisher's sharp null hypothesis, which implies that $\theta = \mu_1 - \mu_0$ is 0. Let $\boldsymbol{X}$ be a vector of baseline covariates and let F denote the joint distribution of $\boldsymbol{X}$ and $Y = Y(0) = Y(1)$. Suppose F is discrete with support $\{(\boldsymbol{x}_l, y_l), l = 1, \ldots, N\}$, and write $F = \sum_{l=1}^{N} p_l \delta_l$, where $p_l \geq 0, \forall l, \sum_{l=1}^{N} p_l = 1$, and δ_l denotes a point mass of 1 at $(\boldsymbol{x}_l, y_l)$. Note that the population may be infinite even though the support for $(\boldsymbol{X}, Y)$ is assumed finite. Consider a randomized trial based on a random sample of size n from the population described earlier. Let $(\boldsymbol{X}_i, Y_i)$, $i = 1, \ldots, n$, denote the covariate and outcome values in the sample; thus, $(\boldsymbol{X}_i, Y_i)$ are independent and identically distributed according to F. Let T_i denote the randomized treatment for the ith subject in the sample, so T_i is Bernoulli with success probability π, independently of $(\boldsymbol{X}_i, Y_i)$ and across subjects. The aforementioned description is for an arbitrary support set $\{(\boldsymbol{x}_l, y_l), l = 1, \ldots, N\}$ and any probability vector $(p_l)_{l=1}^{N}$. To make the simulation study somewhat relevant to the NINDS rt-PA study, we take the support set to be the collection of observed $(\boldsymbol{X}, Y)$ values in the study, and set $p_l = 1/N, l = 1, \ldots, N$. For each outcome variable, 1000 samples of size $n = 200$ are simulated and analyzed using different methods.

Table 1 presents the simulation results for the standardized scores, ie, empirical bias, standard deviation, relative efficiency, median standard error, and coverage probability for estimating θ, a zero difference in mean score. Inference is based on both un–cross-validated and fully cross-validated standard errors and the associated 95% Wald confidence intervals. The relative efficiency of an estimator is calculated as the ratio of the variance of the unadjusted estimator $\widetilde{\theta}$ (defined by (2)) to the variance of the estimator under consideration. For example, compared to the unadjusted estimator, an estimator with relative efficiency 2 allows the sample size to be reduced by half while maintaining the same level of precision. The results are compared between $\widetilde{\theta}$ and various versions of $\widehat{\theta}(\widehat{a})$, where $\widehat{a}$ may be obtained using a super learner or one of the prediction algorithms described earlier, under the direct or indirect approach. In Table 1, all estimators are nearly

TABLE 1 Simulation results for standardized (to the unit interval) scores: empirical bias, standard deviation (SD), relative efficiency (RE), median standard error (SE), and coverage probability (CP) for estimating a zero difference in mean score ($\theta = 0$). The RE of an estimator is the ratio of the variance of the unadjusted estimator to that of the estimator in question. SE and CP results are obtained with and without cross-validation (cv)

Method	Indirect Approach							Direct Approach						
(Algorithm)	Bias	SD	RE	SE	SE_{cv}	CP	CP_{cv}	Bias	SD	RE	SE	SE_{cv}	CP	CP_{cv}
Barthel Index														
unadjusted	−0.001	0.061	1.00	0.059		0.937		−0.001	0.061	1.00	0.059		0.939	
glm	−0.001	0.061	1.00	0.059	0.060	0.936	0.940	−0.001	0.050	1.50	0.046	0.050	0.936	0.954
glmnet ($\alpha = 0$)	−0.001	0.061	1.00	0.059	0.060	0.937	0.940	−0.001	0.050	1.49	0.047	0.050	0.932	0.949
glmnet ($\alpha = 1$)	−0.001	0.061	1.00	0.059	0.060	0.937	0.939	−0.001	0.050	1.50	0.047	0.050	0.931	0.944
gam	−0.001	0.061	1.00	0.059	0.060	0.936	0.939	−0.001	0.050	1.49	0.045	0.050	0.930	0.952
rpart	−0.002	0.057	1.15	0.051	0.056	0.907	0.938	−0.001	0.049	1.53	0.040	0.054	0.880	0.957
randomForest ($m = 2$)	−0.002	0.056	1.18	0.051	0.055	0.922	0.943	−0.001	0.049	1.57	0.033	0.048	0.817	0.939
randomForest ($m = 3$)	−0.002	0.054	1.29	0.047	0.054	0.906	0.946	−0.001	0.048	1.62	0.024	0.047	0.679	0.936
super learner	−0.002	0.054	1.29	0.047	0.054	0.910	0.940	−0.002	0.048	1.64	0.027	0.048	0.729	0.946
Modified Rankin Scale														
unadjusted	0.002	0.057	1.00	0.059		0.953		0.002	0.057	1.00	0.059		0.954	
glm	0.002	0.057	1.00	0.059	0.059	0.952	0.956	0.002	0.049	1.39	0.046	0.050	0.946	0.957
glmnet ($\alpha = 0$)	0.002	0.057	1.00	0.059	0.059	0.953	0.955	0.002	0.048	1.43	0.046	0.049	0.946	0.963
glmnet ($\alpha = 1$)	0.002	0.057	1.00	0.059	0.059	0.953	0.955	0.002	0.048	1.45	0.046	0.049	0.949	0.962
gam	0.002	0.057	1.00	0.059	0.059	0.952	0.956	0.002	0.048	1.41	0.045	0.050	0.936	0.960
rpart	0.002	0.056	1.06	0.051	0.056	0.923	0.944	0.002	0.050	1.31	0.039	0.053	0.880	0.964
randomForest ($m = 2$)	0.002	0.053	1.18	0.051	0.055	0.945	0.957	0.002	0.047	1.47	0.033	0.047	0.825	0.949
randomForest ($m = 3$)	0.002	0.051	1.27	0.047	0.053	0.931	0.950	0.002	0.046	1.55	0.023	0.046	0.662	0.948
super learner	0.002	0.052	1.24	0.048	0.053	0.928	0.948	0.002	0.046	1.55	0.026	0.047	0.712	0.956
Glasgow Outcome Scale														
unadjusted	−0.001	0.042	1.00	0.041		0.942		−0.001	0.042	1.00	0.041		0.943	
glm	−0.001	0.042	1.00	0.041	0.042	0.942	0.942	0.000	0.036	1.34	0.033	0.036	0.931	0.950
glmnet ($\alpha = 0$)	−0.001	0.042	1.00	0.041	0.042	0.942	0.944	0.000	0.036	1.35	0.033	0.036	0.937	0.950
glmnet ($\alpha = 1$)	−0.001	0.042	1.00	0.041	0.042	0.942	0.944	0.000	0.036	1.35	0.033	0.035	0.935	0.949
gam	−0.001	0.042	1.00	0.041	0.042	0.942	0.945	0.000	0.036	1.34	0.032	0.036	0.927	0.949
rpart	0.000	0.040	1.08	0.037	0.040	0.920	0.942	0.000	0.035	1.41	0.029	0.038	0.890	0.967
randomForest ($m = 2$)	0.000	0.039	1.14	0.037	0.039	0.938	0.948	0.000	0.035	1.45	0.023	0.034	0.811	0.949
randomForest ($m = 3$)	0.000	0.038	1.23	0.034	0.038	0.922	0.953	0.000	0.034	1.52	0.017	0.034	0.673	0.952
super learner	0.000	0.038	1.20	0.034	0.038	0.921	0.953	0.000	0.034	1.50	0.018	0.035	0.713	0.958
NIHSS														
unadjusted	0.000	0.053	1.00	0.052		0.948		−0.001	0.053	1.00	0.052		0.946	
glm	0.000	0.053	1.00	0.052	0.052	0.947	0.948	−0.001	0.046	1.33	0.042	0.046	0.920	0.950
glmnet ($\alpha = 0$)	0.000	0.053	1.00	0.052	0.052	0.948	0.948	−0.001	0.046	1.34	0.042	0.045	0.930	0.940
glmnet ($\alpha = 1$)	0.000	0.053	1.00	0.052	0.052	0.948	0.948	−0.001	0.046	1.35	0.042	0.045	0.929	0.941
gam	0.000	0.053	1.00	0.052	0.052	0.947	0.948	−0.001	0.046	1.32	0.041	0.046	0.916	0.950
rpart	0.000	0.050	1.12	0.046	0.050	0.935	0.954	0.000	0.045	1.42	0.036	0.048	0.885	0.962
randomForest ($m = 2$)	0.000	0.050	1.13	0.046	0.049	0.926	0.943	−0.001	0.044	1.45	0.029	0.043	0.803	0.943
randomForest ($m = 3$)	0.000	0.048	1.21	0.042	0.048	0.904	0.948	−0.001	0.044	1.48	0.021	0.042	0.671	0.936
super learner	0.000	0.049	1.20	0.043	0.048	0.916	0.950	−0.001	0.044	1.47	0.023	0.043	0.682	0.948

Abbreviations: NIHSS, National Institutes of Health Stroke Scale.

unbiased, but they differ substantially in efficiency. Under the indirect approach, $\widehat{\theta}(\widehat{a})$ is not necessarily more efficient than $\widetilde{\theta}$, and the super learner estimator $\widehat{\theta}(\widehat{a}_{\text{sl.ind}})$ does not necessarily outperform the estimators based on individual prediction algorithms. The direct approach clearly outperforms the indirect approach as well as the unadjusted method. Under the direct approach, $\widehat{\theta}(\widehat{a})$ is generally more efficient than $\widetilde{\theta}$, and the super learner estimator $\widehat{\theta}(\widehat{a}_{\text{sl.dir}})$ usually attains the best efficiency (approximately) among all estimators considered. The un–cross-validated standard errors sometimes exhibit a downward bias, leading to serious under-coverage of confidence intervals. The cross-validated standard errors generally perform better, and the associated confidence intervals have close-to-nominal coverage.

Table 2 shows the analogous simulation results for the dichotomized outcomes, which are qualitatively similar to those in Table 1.

TABLE 2 Simulation results for dichotomized outcomes: empirical bias, standard deviation (SD), relative efficiency (RE), median standard error (SE), and coverage probability (CP) for estimating a zero difference in proportion ($\theta = 0$). The RE of an estimator is the ratio of the variance of the unadjusted estimator to that of the estimator in question. SE and CP results are obtained with and without cross-validation (cv)

Method	Indirect Approach							Direct Approach						
(Algorithm)	Bias	SD	RE	SE	SE_{cv}	CP	CP_{cv}	Bias	SD	RE	SE	SE_{cv}	CP	CP_{cv}
Barthel Index														
unadjusted	0.000	0.069	1.00	0.070		0.958		0.000	0.069	1.00	0.070		0.956	
glm	0.000	0.069	1.00	0.070	0.071	0.960	0.963	0.001	0.063	1.22	0.058	0.063	0.921	0.952
glmnet ($\alpha = 0$)	0.000	0.069	1.00	0.070	0.071	0.958	0.960	0.000	0.062	1.25	0.059	0.062	0.939	0.952
glmnet ($\alpha = 1$)	0.000	0.069	1.00	0.070	0.071	0.958	0.960	0.000	0.061	1.26	0.059	0.063	0.936	0.955
gam	0.000	0.069	1.00	0.070	0.071	0.960	0.963	0.001	0.062	1.25	0.057	0.063	0.928	0.957
rpart	0.000	0.067	1.06	0.062	0.068	0.925	0.953	0.001	0.060	1.33	0.050	0.067	0.903	0.972
randomForest ($m = 2$)	0.000	0.065	1.14	0.062	0.067	0.941	0.960	0.000	0.059	1.34	0.041	0.060	0.829	0.953
randomForest ($m = 3$)	0.001	0.062	1.22	0.057	0.065	0.929	0.959	0.000	0.059	1.37	0.031	0.059	0.674	0.954
super learner	0.000	0.063	1.20	0.058	0.065	0.933	0.961	0.001	0.059	1.37	0.036	0.061	0.750	0.953
Modified Rankin Scale														
unadjusted	0.002	0.069	1.00	0.067		0.943		0.002	0.068	1.00	0.067		0.943	
glm	0.002	0.069	1.00	0.067	0.068	0.942	0.944	0.002	0.062	1.22	0.057	0.062	0.929	0.945
glmnet ($\alpha = 0$)	0.002	0.069	1.00	0.067	0.068	0.943	0.946	0.002	0.062	1.24	0.057	0.061	0.932	0.946
glmnet ($\alpha = 1$)	0.002	0.069	1.00	0.067	0.068	0.943	0.946	0.002	0.061	1.25	0.058	0.061	0.934	0.947
gam	0.002	0.069	1.00	0.067	0.068	0.942	0.944	0.001	0.061	1.27	0.055	0.061	0.921	0.948
rpart	0.003	0.064	1.14	0.059	0.065	0.926	0.953	0.003	0.059	1.36	0.048	0.064	0.894	0.963
randomForest ($m = 2$)	0.002	0.064	1.14	0.059	0.064	0.931	0.949	0.002	0.058	1.38	0.040	0.058	0.828	0.948
randomForest ($m = 3$)	0.002	0.062	1.23	0.055	0.062	0.910	0.950	0.002	0.058	1.39	0.029	0.057	0.684	0.941
super learner	0.002	0.062	1.22	0.055	0.062	0.912	0.957	0.002	0.058	1.40	0.032	0.058	0.726	0.948
Glasgow Outcome Scale														
unadjusted	0.002	0.070	1.00	0.069		0.940		0.002	0.070	1.00	0.069		0.938	
glm	0.002	0.070	0.99	0.069	0.069	0.939	0.941	0.001	0.061	1.31	0.057	0.062	0.937	0.951
glmnet ($\alpha = 0$)	0.002	0.070	1.00	0.069	0.069	0.940	0.941	0.001	0.061	1.31	0.058	0.061	0.939	0.951
glmnet ($\alpha = 1$)	0.002	0.070	1.00	0.069	0.069	0.940	0.941	0.001	0.061	1.31	0.058	0.061	0.940	0.951
gam	0.002	0.070	0.99	0.069	0.069	0.939	0.940	0.001	0.061	1.33	0.056	0.061	0.927	0.955
rpart	0.002	0.067	1.09	0.061	0.066	0.918	0.933	0.001	0.059	1.41	0.048	0.065	0.907	0.970
randomForest ($m = 2$)	0.001	0.065	1.16	0.060	0.065	0.928	0.949	0.001	0.059	1.41	0.040	0.058	0.808	0.950
randomForest ($m = 3$)	0.001	0.062	1.26	0.056	0.063	0.910	0.958	0.001	0.059	1.42	0.030	0.058	0.675	0.947
super learner	0.001	0.063	1.25	0.056	0.063	0.912	0.951	0.001	0.059	1.42	0.033	0.059	0.721	0.946
NIHSS														
unadjusted	−0.002	0.065	1.00	0.063		0.944		−0.001	0.064	1.00	0.063		0.947	
glm	−0.002	0.066	1.00	0.063	0.064	0.943	0.946	−0.001	0.059	1.22	0.055	0.060	0.928	0.948
glmnet ($\alpha = 0$)	−0.002	0.065	1.00	0.063	0.064	0.944	0.946	0.000	0.059	1.22	0.055	0.059	0.926	0.946
glmnet ($\alpha = 1$)	−0.002	0.065	1.00	0.063	0.064	0.944	0.946	−0.001	0.059	1.22	0.056	0.059	0.929	0.939
gam	−0.002	0.066	1.00	0.063	0.064	0.944	0.945	−0.001	0.058	1.25	0.053	0.059	0.927	0.948
rpart	−0.001	0.062	1.12	0.057	0.062	0.930	0.945	−0.002	0.057	1.31	0.046	0.063	0.877	0.972
randomForest ($m = 2$)	−0.002	0.062	1.12	0.057	0.061	0.928	0.947	−0.001	0.056	1.36	0.039	0.056	0.829	0.950
randomForest ($m = 3$)	−0.002	0.060	1.18	0.052	0.059	0.923	0.952	−0.001	0.056	1.37	0.028	0.055	0.674	0.948
super learner	−0.002	0.060	1.17	0.053	0.060	0.925	0.952	−0.001	0.056	1.36	0.031	0.057	0.698	0.946

Abbreviations: NIHSS, National Institutes of Health Stroke Scale.

3.3 | Analysis

Table 3 shows the results of analyzing the NINDS rt-PA trial data using the same methods compared in Tables 1 and 2. Note that, for all instruments but the Barthel Index, higher scores represent worse outcomes, and a negative treatment difference (rt-PA minus placebo) in mean standardized score represents a beneficial effect of rt-PA relative to placebo. The results in Table 3 demonstrate a substantial and statistically significant benefit of rt-PA for most outcomes and effect measures, with the possible exception of mean difference in the NIHSS score. In some cases, the cross-validated standard error is substantially larger than the un–cross-validated standard error, showing the effect of cross-validation. In other cases, the two standard errors are close to each other, suggesting that overfitting may be a lesser issue in these cases. The cross-validated standard errors of the augmented estimates are generally smaller for the direct approach than the indirect approach, and those for the direct approach are usually smaller than the standard error of the unadjusted estimator. The

TABLE 3 Analysis of the National Institute of Neurological Disorders and Stroke Recombinant Tissue Plasminogen Activator trial data: point estimates (PE) and standard errors (SE) with and without cross-validation (cv) for the treatment differences (rt-PA minus placebo) in the mean standardized score and the probability of success (defined using four different instruments)

Method (Algorithm)	Difference in Mean Standardized Score						Difference in Probability of Success					
	Indirect Approach			Direct Approach			Indirect Approach			Direct Approach		
	PE	SE	SE_{cv}	PE	SE	SE_{cv}	PE	SE	SE_{cv}	PE	SE	SE_{cv}
Barthel Index												
unadjusted	0.084	0.033		0.084	0.033		0.142	0.040		0.142	0.040	
glm	0.084	0.033	0.033	0.075	0.027	0.028	0.142	0.040	0.040	0.138	0.034	0.035
glmnet ($\alpha = 0$)	0.084	0.033	0.033	0.075	0.027	0.028	0.142	0.040	0.040	0.138	0.034	0.034
glmnet ($\alpha = 1$)	0.084	0.033	0.033	0.075	0.027	0.028	0.142	0.040	0.040	0.137	0.034	0.034
gam	0.084	0.033	0.033	0.074	0.027	0.028	0.142	0.040	0.040	0.127	0.033	0.034
rpart	0.087	0.027	0.029	0.089	0.027	0.029	0.099	0.034	0.038	0.129	0.034	0.036
randomForest ($m = 2$)	0.084	0.029	0.031	0.087	0.021	0.028	0.136	0.035	0.037	0.139	0.027	0.035
randomForest ($m = 3$)	0.084	0.032	0.033	0.085	0.028	0.030	0.140	0.038	0.039	0.139	0.035	0.036
super learner	0.085	0.028	0.030	0.085	0.027	0.028	0.130	0.034	0.037	0.125	0.035	0.034
Modified Rankin Scale												
unadjusted	−0.107	0.033		−0.107	0.033		0.164	0.038		0.164	0.038	
glm	−0.107	0.033	0.033	−0.095	0.027	0.028	0.164	0.038	0.038	0.151	0.033	0.034
glmnet ($\alpha = 0$)	−0.107	0.033	0.033	−0.096	0.027	0.028	0.164	0.038	0.038	0.152	0.033	0.034
glmnet ($\alpha = 1$)	−0.107	0.033	0.033	−0.097	0.027	0.027	0.164	0.038	0.038	0.153	0.033	0.034
gam	−0.107	0.033	0.033	−0.090	0.026	0.028	0.164	0.038	0.038	0.141	0.032	0.033
rpart	−0.085	0.030	0.029	−0.096	0.027	0.028	0.142	0.034	0.035	0.144	0.033	0.035
randomForest ($m = 2$)	−0.104	0.029	0.031	−0.106	0.021	0.028	0.156	0.033	0.036	0.150	0.026	0.034
randomForest ($m = 3$)	−0.106	0.032	0.032	−0.106	0.028	0.030	0.162	0.037	0.037	0.157	0.033	0.035
super learner	−0.098	0.029	0.029	−0.095	0.028	0.028	0.156	0.033	0.035	0.138	0.033	0.033
Glasgow Outcome Scale												
unadjusted	−0.060	0.024		−0.060	0.024		0.145	0.039		0.145	0.039	
glm	−0.060	0.024	0.024	−0.052	0.019	0.020	0.145	0.039	0.039	0.130	0.033	0.034
glmnet ($\alpha = 0$)	−0.060	0.024	0.024	−0.053	0.019	0.020	0.145	0.039	0.039	0.133	0.033	0.034
glmnet ($\alpha = 1$)	−0.060	0.024	0.024	−0.054	0.019	0.020	0.145	0.039	0.039	0.133	0.033	0.034
gam	−0.060	0.024	0.024	−0.052	0.019	0.020	0.145	0.039	0.039	0.122	0.033	0.034
rpart	−0.059	0.020	0.022	−0.055	0.019	0.021	0.122	0.033	0.034	0.132	0.032	0.036
randomForest ($m = 2$)	−0.060	0.021	0.022	−0.063	0.015	0.020	0.139	0.034	0.036	0.141	0.026	0.034
randomForest ($m = 3$)	−0.060	0.023	0.023	−0.060	0.020	0.021	0.143	0.038	0.038	0.142	0.033	0.036
super learner	−0.060	0.020	0.021	−0.054	0.020	0.020	0.135	0.034	0.035	0.126	0.029	0.034
NIHSS												
unadjusted	−0.052	0.029		−0.052	0.029		0.138	0.035		0.138	0.035	
glm	−0.052	0.029	0.030	−0.041	0.025	0.025	0.138	0.035	0.036	0.132	0.032	0.033
glmnet ($\alpha = 0$)	−0.052	0.029	0.030	−0.042	0.025	0.025	0.138	0.035	0.036	0.131	0.032	0.033
glmnet ($\alpha = 1$)	−0.052	0.029	0.030	−0.044	0.025	0.025	0.138	0.035	0.036	0.130	0.032	0.032
gam	−0.052	0.029	0.030	−0.043	0.024	0.026	0.138	0.035	0.036	0.123	0.031	0.032
rpart	−0.051	0.025	0.026	−0.049	0.025	0.027	0.124	0.031	0.033	0.132	0.030	0.036
randomForest ($m = 2$)	−0.053	0.026	0.028	−0.059	0.019	0.026	0.128	0.032	0.034	0.129	0.025	0.033
randomForest ($m = 3$)	−0.053	0.028	0.029	−0.053	0.025	0.027	0.135	0.034	0.035	0.130	0.032	0.033
super learner	−0.053	0.025	0.026	−0.056	0.020	0.025	0.127	0.031	0.033	0.124	0.030	0.033

cross-validated standard error of the direct super learner estimate is usually equal and sometimes close to the smallest cross-validated standard error among all estimates.

4 | DISCUSSION

Clinical trials can be very expensive in terms of time and money. Improving the efficiency of clinical trials via effective use of information, including baseline covariate information, is key to accelerating the development of new and effective medical products. This article shows that machine learning methods can bring about substantial improvements in the efficiency of clinical trials, which translate into big savings for medical product developers and the society as a whole.

There are different ways of using machine learning methods for this purpose and, in this article, we have considered and compared a direct approach and an indirect approach. The indirect approach is obvious and straightforward but not

well aligned with the analytical objective. As a result, it does not perform well and sometimes fails to improve upon the unadjusted estimator. The direct approach aims to minimize a risk function that is directly interpretable as the asymptotic variance of the treatment effect estimator. It is easy to implement via weighted least squares, and generally performs better than the indirect approach and the unadjusted estimator. Therefore, we recommend the direct approach over the indirect one.

There is a great variety of machine learning methods available with different properties and operating characteristics. The super learning principle allows us to combine many candidate algorithms into a super learner with a desirable oracle property. Our simulation results confirm that the super learner generally performs better (or no worse) than the candidate algorithms under the direct approach. In addition to its superior statistical performance, the direct super learner also has an important practical advantage, ie, it eliminates the need for trial planners to choose among many candidate algorithms, as we can simply include them all in the library for building a super learner. We do need to specify the library of candidate algorithms a priori to avoid data dredging.

ACKNOWLEDGEMENT

We thank Dr. Robert Dachs for providing the NINDS rt-PA trial data, and two anonymous reviewers for constructive comments on an earlier version of the paper.

ORCID

Zhiwei Zhang https://orcid.org/0000-0002-3187-9180

REFERENCES

1. Greenland S, Pearl J, Robins JM. Confounding and collapsibility in causal inference. *Stat Sci*. 1999;14(1):29-46.
2. Tsiatis AA, Davidian M, Zhang M, Lu X. Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. *Statist Med*. 2008;27(23):4658-4677.
3. Moore KL, van der Laan MJ. Covariate adjustment in randomized trials with binary outcomes: targeted maximum likelihood estimation. *Statist Med*. 2009;28(1):39-64.
4. Rosenblum M, van der Laan MJ. Simple, efficient estimators of treatment effects in randomized trials using generalized linear models to leverage baseline variables. *Int J Biostat*. 2010;6(1). Article No 13.
5. Zhang M, Tsiatis AA, Davidian M. Improving efficiency of inferences in randomized clinical trials using auxiliary covariates. *Biometrics*. 2008;64(3):707-715.
6. Rubin DB, van der Laan MJ. Empirical efficiency maximization: improved locally efficient covariate adjustment in randomized experiments and survival analysis. *Int J Biostat*. 2008;4(1). Article No 5.
7. Tian L, Cai T, Zhao L, Wei L-J. On the covariate-adjusted estimation for an overall treatment difference with data from a randomized comparative clinical trial. *Biostatistics*. 2012;13(2):256-273.
8. Shen C, Li X, Li L. Inverse probability weighting for covariate adjustment in randomized studies. *Statist Med*. 2014;33(4):555-568.
9. Balzer LB, van der Laan MJ, Petersen ML, the SEARCH Collaboration. Adaptive pre-specification in randomized trials with and without pair-matching. *Statist Med*. 2016;35(25):4528-4545.
10. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. New York, NY: Springer-Verlag; 2009.
11. van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol*. 2007;6(1).
12. van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York, NY: Springer Science+Business Media; 2011.
13. Bickel PJ, Klaassen CAJ, Ritov Y, Wellner JA. *Efficient and Adaptive Estimation for Semiparametric Models*. Baltimore, MD: Johns Hopkins University Press; 1993.
14. Tsiatis AA. *Semiparametric Theory and Missing Data*. New York, NY: Springer Science+Business Media; 2006.
15. van der Laan MJ, Dudoit S. *Unified Cross-Validation Methodology for Selection Among Estimators and a General Cross-Validated Adaptive Epsilon-Net Estimator: Finite Sample Oracle Inequalities and Examples*. Working Paper. Berkeley, CA: UC Berkeley Division of Biostatistics Working Paper Series. 2003. http://biostats.bepress.com/ucbbiostat/paper130
16. The National Institute of Neurological Disorders, Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. 1995;333:1581-1587.
17. Ingall TJ, O'Fallon WM, Asplund K, et al. Findings from the reanalysis of the NINDS tissue plasminogen activator for acute ischemic stroke treatment trial. *Stroke*. 2004;35(10):2418-2424.
18. Zhang Z, Tang L, Liu C, Berger VW. Conditional estimation and inference to address observed covariate imbalance in randomized clinical trials. Clinical Trials. 2018. https://doi.org/10.1177/1740774518813120

19. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B*. 2005;67(2):301-320.
20. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. Boca Raton, FL: Chapman & Hall/CRC; 1990.
21. Breiman L, Friedman J, Olshen RA, Stone CJ. *Classification and Regression Trees*. Boca Raton, FL: Chapman & Hall/CRC; 1984. *Wadsworth Statistics/Probability Series*.
22. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
23. van der Vaart AW, Wellner JA. *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York, NY: Springer-Verlag New York; 1996.
24. van der Vaart AW. *Asymptotic Statistics*. Cambridge, UK: Cambridge University Press; 1998.

How to cite this article: Zhang Z, Ma S. Machine learning methods for leveraging baseline covariate information to improve the efficiency of clinical trials. *Statistics in Medicine*. 2019;38:1703–1714. https://doi.org/10.1002/sim.8054

APPENDIX

JUSTIFICATION FOR EQUATION (7)

We write $\mathbb{P}_0$ for the true distribution of $(\boldsymbol{X}, T, Y)$ and $\mathbb{P}_n$ for the empirical distribution of $(\boldsymbol{X}_i, T_i, Y_i)$, $i = 1, \ldots, n$. Let $\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - \mathbb{P}_0)$ be the empirical process. We shall use operator notation for integrals, writing

$$\widehat{\theta}(\widehat{a}) = \widetilde{\theta} - \mathbb{P}_n\left\{(T-\pi)\widehat{a}(\boldsymbol{X})\right\}.$$

It follows that

$$\begin{aligned}\sqrt{n}\left\{\widehat{\theta}(\widehat{a}) - \theta\right\} &= \sqrt{n}(\widetilde{\theta} - \theta) - \sqrt{n}\mathbb{P}_n\left\{(T-\pi)\widehat{a}(\boldsymbol{X})\right\} \\ &= \mathbb{G}_n\left\{\psi(T,Y) - (T-\pi)\widehat{a}(\boldsymbol{X})\right\} + o_p(1),\end{aligned} \tag{A1}$$

where the second step follows from (3) and the fact that $\mathbb{P}_0\{(T-\pi)a(\boldsymbol{X})\} = 0$ for any function a because of randomization.

We assume that, with probability tending to 1, both $\widehat{a}$ and a^* belong to some Donsker class of functions with a square-integrable envelope (with respect to $\mathbb{P}_0$). Techniques for verifying the Donsker condition can be found in the work of van der Vaart and Wellner.[23] We also assume that $\widehat{a}$ converges in probability to a^* in $L_2(\mathbb{P}_0)$, ie,

$$\mathbb{P}_0\left[\left\{\widehat{a}(\boldsymbol{X}) - a^*(\boldsymbol{X})\right\}^2\right] = o_p(1).$$

Now, it follows from theorem 19.24 in the work of van der Vaart[24] that

$$\mathbb{G}_n\left\{\psi(T,Y) - (T-\pi)\widehat{a}(\boldsymbol{X})\right\} = \mathbb{G}_n\{\psi(T,Y) - (T-\pi)a^*(\boldsymbol{X})\} + o_p(1).$$

Substituting this into (A1) completes the proof of (7).