Check for updates

Received: 6 March 2021 | Revised: 15 October 2021 | Accepted: 15 October 2021

DOI: 10.1002/sim.9253

RESEARCH ARTICLE

Statistics in Medicine WILEY

Risk controlled decision trees and random forests for precision Medicine

Kevin Doubleday[1] | **Jin Zhou**[2] | **Hua Zhou**[2] | **Haoda Fu**[3]

[1]Department of Biostatistics, University of Arizona, Tucson, Arizona

[2]Department of Biostatistics, University of California, Los Angeles, California

[3]Eli Lilly and Company, Indianapolis, Indiana

Correspondence
Kevin Doubleday, Department of Biostatistics, University of Arizona, 1295 N Martin Ave, Tucson, AZ 85724, USA.
Email: kdoub5ha@gmail.com

Statistical methods generating individualized treatment rules (ITRs) often focus on maximizing expected benefit, but these rules may expose patients to excess risk. For instance, aggressive treatment of type 2 diabetes (T2D) with insulin therapies may result in an ITR which controls blood glucose levels but increases rates of hypoglycemia, diminishing the appeal of the ITR. This work proposes two methods to identify risk-controlled ITRs (rcITR), a class of ITR which maximizes a benefit while controlling risk at a prespecified threshold. A novel penalized recursive partitioning algorithm is developed which optimizes an unconstrained, penalized value function. The final rule is a risk-controlled decision tree (rcDT) that is easily interpretable. A natural extension of the rcDT model, risk controlled random forests (rcRF), is also proposed. Simulation studies demonstrate the robustness of rcRF modeling. Three variable importance measures are proposed to further guide clinical decision-making. Both rcDT and rcRF procedures can be applied to data from randomized controlled trials or observational studies. An extensive simulation study interrogates the performance of the proposed methods. A data analysis of the DURABLE diabetes trial in which two therapeutics were compared is additionally presented. An R package implements the proposed methods (https://github.com/kdoub5ha/rcITR).

KEYWORDS

decision trees, precision medicine, random forests, risk control, variable importance

1 | INTRODUCTION

In precision medicine, treatment selection is carried out based on an individual's particular set of characteristics. Optimizing clinical benefit, or treatment efficacy, is often the primary objective of a treatment recommendation, but secondary complications and risks should also be considered. Numerous statistical methods propose constructing individualized treatment rules (ITRs) to optimize efficacy,[1-9] including our previous work.[10] ITRs that optimize treatment efficacy with respect to several outcomes are a natural extension. For instance, Lizotte et al[11] proposed a reward function that scans across all linear combinations of trade-offs between the outcomes under consideration. Examining trade-offs between outcomes is commonly investigated in developing optimal dosing strategies (eg, Thall et al[12]) where striking a balance between response to treatment and toxicity is desired. Lipkovich et al[2] proposed to balance efficacy and risk via specifying a joint distribution of efficacy and risk scores, and used a splitting criteria based on an additive model that maximized the weighted group effects for efficacy and risk. Their method, however, assumed training data were from a randomized

 © 2021 John Wiley & Sons Ltd.

trial. Laber et al[13] avoided explicit trade-offs in ITR specification by returning a set of treatment recommendations demonstrating significant differential benefit in all outcomes.

Alternatively, one could construct treatment rules that maximize an expected clinical benefit while controlling expected risk at a given threshold. For instance, Wang et al[14] estimated the marginal mean efficacy and risk scores under a value-based framework (Qian and Murphy[5]). They provided two solutions to this constrained optimization problem, model-based benefit-risk learning (BR-M) and benefit-risk O-learning (BR-O). BR-M determines optimal ITR through estimating treatment contrasts for efficacy and risk scores. Treatment contrasts were modeled using a linear function of predictors, subjecting this approach to the potential for model misspecification leading to a suboptimal rule. BR-O is an extension of the outcome weighted learning method.[8] They utilized support vector machines (SVM) and kernel tricks to solve optimal risk-controlled ITR. However, nonlinear kernels (eg, Gaussian) in BR-O introduce substantial computational cost in problems of modest dimensionality. Variable importance measures were proposed for BR-O as a rank ordering of coefficient magnitude. For both BR-M and BR-O, interpreting the final rule is difficult. Note that the methods proposed both here and by Wang et al[14] conceptualize "riskε as an undesirable, measurable outcome that is potentially correlated with treatment assignment and patient profile.

In this work, two nonparametric, risk-controlled ITR (rcITR) estimation procedures are proposed. Both methods utilize decision trees[15,16] to identify an ITR with maximized expected treatment efficacy while maintaining expected risk within a clinically relevant threshold. This constrained problem was solved by formulating an objective function that maximizes expected efficacy while penalizing rules with expected risk greater than a clinically relevant threshold. The proposed approach avoids specifying trade-offs as was the case in Lizotte et al.[11] Instead, prespecification of a clinically relevant risk threshold is utilized in a constrained optimization-like framework to obtain a high efficacy rule that controls expected risk at the specified threshold. The proposed objective function is closely related to that proposed by Wang et al,[14] but has the advantage of preferring rules with lower risk, given similar efficacy. This corresponds well to patient preferences for treatment choices in the real world. Decision trees are employed to optimize the objective function. The resulting model is called a risk-controlled decision tree (rcDT). Traditional decision trees are comprised of a series of localized (ie, node-specific) models, and single tree models are highly variable. Accordingly, the rcDT model is extended to risk-controlled random forests (rcRF). rcRF aggregates rcDT learners with notable advantages. First, unbiased risk estimates can be obtained from the out-of-bag sample and leveraged in model selection. Second, rcRF is robust as it aggregates several weaker rcDT learners. Third, rcRFs can return a hard rule (ie, treat or not) and a probability measure to assess the strength of the treatment recommendation. Finally, rcRF naturally offers variable importance measures relevant to ITRs. Both rcDT and rcRF can be applied to randomized controlled trial data as well as observational studies. Performance of rcDT and rcRF procedures was assessed through extensive simulation studies. Data analysis was performed using randomized controlled trial data from the DURABLE (DURAbility of Basal vs Lispro mix 75/25 insulin Efficacy) trial.[17]

In summary, our contribution in this paper is: (1) development of two non-parametric rcITR estimation procedures using decision trees (rcDT) and random forests (rcRF); (2) implementation of three variable importance measures for an rcRF model; (3) extensive assessment of rcDT and rcRF methods via simulation studies; (4) demonstration of rcDT and rcRF procedures by analysis of DURABLE trial data; (5) implementation of rcDT and rcRF methods in a publicly available software package `rcITR` (https://github.com/kdoub5ha/rcITR) using the statistical computing language `R`.[18]

2 | STATISTICAL FRAMEWORK

Given a finite sample of n observations from a population of interest, the observed data are $(y_i, r_i, \boldsymbol{x}_i, a_i)$ for $i = 1, 2, \ldots, n$. Values of $y_i, r_i \in \mathbb{R}$ correspond to patient efficacy and risk scores, $\boldsymbol{x}_i \subset \mathbb{R}^p$ is the p-dimensional vector of covariates, and $a_i \in \{0, 1\}$ is the binary treatment indicator. Observed data is a realization from the underlying distribution $\mathcal{P} = (\boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{A})$. Without loss of generality, larger values of $\boldsymbol{Y}$ and smaller values of $\boldsymbol{R}$ are assumed to be desirable. Define the propensity score $p_i = \Pr(a_i|\boldsymbol{x}_i) \in (0, 1)$, that is, the probability of receiving treatment a_i given the covariates. For randomized controlled trials (RCTs) the value of p_i is viewed as fixed (eg, assuming two treatments and a 1:1 allocation ratio, then $p_i = 0.5$). In observational studies, p_i can be estimated using logistic regression.[19,20] A treatment rule d maps the predictor space to the treatment space, that is, $d : \boldsymbol{X} \rightarrow \boldsymbol{A}$ such that $d(\boldsymbol{x}_i) \in \{0, 1\}$. Let $\mathrm{E}^d(\boldsymbol{Y})$ and $\mathrm{E}^d(\boldsymbol{R})$ correspond to the expected efficacy and risk under rule d, respectively.

We aim to maximize the expected efficacy while controlling expected risk at a clinically relevant level.[14] Specifically, we seek to find d_o that solves

$$\begin{cases} \max_d & E^d(\boldsymbol{Y}) \\ \text{subject to} & E^d(\boldsymbol{R}) \leq \tau, \end{cases} \tag{1}$$

where τ is a predefined clinically relevant bound for the expected risk. When $\Pr(a_i|\boldsymbol{x}_i)$ is strictly positive, $\mathrm{E}^d(\boldsymbol{Y})$ and $\mathrm{E}^d(\boldsymbol{R})$[5] can be estimated as,

$$V_Y(d) = \mathrm{E}^d(\boldsymbol{Y}) = \left(\sum_{i=1}^n w_i\right)^{-1}\left(\sum_{i=1}^n w_i y_i\right), \tag{2}$$

$$V_R(d) = \mathrm{E}^d(\boldsymbol{R}) = \left(\sum_{i=1}^n w_i\right)^{-1}\left(\sum_{i=1}^n w_i r_i\right), \tag{3}$$

where $w_i = \mathrm{I}(a_i = d(\boldsymbol{x}_i))(\hat{\Pr}(a_i|\boldsymbol{x}_i))^{-1}$.

The current work aims to solve an optimization problem closely related to (1). Under a recursive partitioning framework, a splitting criterion needs to be defined that increases the "purity" of the daughter nodes relative to the common parent node. To that end, the constrained optimization problem (1) is translated to an unconstrained one,

$$L(d) = \mathrm{E}^d[\boldsymbol{Y}] - \lambda\left(\mathrm{E}^d[\boldsymbol{R}] - \tau\right), \tag{4}$$

where $\lambda > 0$ penalizes rules with expected risk greater than τ, that is, $E_R(d) > \tau$. $L(d)$ acts as a purity measure in the sense that an increase in $L(d)$ indicates either an increase in expected efficacy, a decrease in expected risk, or an increase in risk accompanied by an acceptable increase in efficacy (see Appendix S1 for more details). The optimal partition of the data is selected such that $L(d)$ is maximized and λ is treated as a tuning parameter. Note that optimizing $L(d)$ does not necessarily imply that $\mathrm{E}[\boldsymbol{Y}]$ is maximized when the risk constraint in system 1 is satisfied since rules with expected risk less than τ contribute positive values to $L(d)$. Optimizing $L(d)$ prefers rules with lower expected risk (for fixed expected efficacy), which is reflective of patient preferences. This differentiates the proposal from translating the optimization as a difference of convex functions as was done by Wang et al.[14] Given a set of values for λ, the rule both maximizing $L(d)$ and satisfying $\mathrm{E}[\boldsymbol{R}] < \tau$ can be selected, thereby enforcing the constraint from 1. In summary, the proposed splitting criteria balances expected efficacy and risk such that rules with smaller risk are preferred for fixed expected efficacy. See Section 3 for details about the optimization.

Apply the law of total expectation to Equation (4) and define $\delta_Y(\boldsymbol{X}) = \mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{A} = 1] - \mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{A} = 0]$ and $\delta_R(\boldsymbol{X}) = \mathrm{E}[\boldsymbol{R}|\boldsymbol{X}, \boldsymbol{A} = 1] - \mathrm{E}[\boldsymbol{R}|\boldsymbol{X}, \boldsymbol{A} = 0]$. Here δ_Y and δ_R correspond to the expected difference in efficacy and risk due to receiving active treatment versus control. Equation (4) is then equivalent to,

$$L(d) = \mathrm{E}\{\mathrm{I}(d(\boldsymbol{X}) = 1)\,[\delta_Y(\boldsymbol{X}) - \lambda\delta_R(\boldsymbol{X})]\} + \bar{\boldsymbol{Y}}_0 + \bar{\boldsymbol{R}}_0 + \lambda\tau, \tag{5}$$

where $\bar{\boldsymbol{Y}}_0$ and $\bar{\boldsymbol{R}}_0$ are the mean efficacy and mean risk in the control group, respectively. From Equation (5) the optimal rule satisfies $L(d) = \mathrm{E}\{\mathrm{I}(d(\boldsymbol{X}) = 1)\,[\delta_Y(\boldsymbol{X}) - \lambda\delta_R(\boldsymbol{X})]\} > 0$. This can be interpreted as $d(\boldsymbol{X})$ satisfying $\frac{\delta_Y(\boldsymbol{X})}{\delta_R(\boldsymbol{X})} > \lambda$ among participants assigned to treatment, that is, $d(\boldsymbol{X}) = 1$. Note that this assumes $\delta_R(\boldsymbol{X}) > 0$. Thus, the ratio of expected benefit to expected risk for those recommended to treatment must be "large enough," which is captured by λ.

Instead of modeling efficacy scores directly, residual efficacy scores are used as model inputs. The validity of this follows from the fact that,

$$\arg\max_d \mathrm{E}[\boldsymbol{w}\cdot\boldsymbol{Y}] = \arg\max_d \mathrm{E}[\boldsymbol{w}\cdot(\boldsymbol{Y} - m(\boldsymbol{X}))],$$

where $m(\boldsymbol{X})$ is some function of $\boldsymbol{X}$. Zhou et al[9] demonstrated that using residuals from a regression of $\boldsymbol{Y}$ on $\boldsymbol{X}$ stabilizes the variance of the expected efficacy estimator. Since efficacy is to be maximized and risk is to be constrained, analysis of residuals is restricted to efficacy scores.[14] The empirical estimate of the purity measure is,

$$\hat{L}(d) = \frac{\sum_{i=1}^n w_i(y_i - \hat{m}(\boldsymbol{x}_i))}{\sum_{i=1}^n w_i} - \hat{\lambda}\left(\frac{\sum_{i=1}^n w_i r_i}{\sum_{i=1}^n w_i} - \tau\right), \tag{6}$$

where $\hat{\lambda}$ is the tuning parameter. For simplicity $\hat{m}(\boldsymbol{x}_i)$ were estimated from a linear regression of $\boldsymbol{Y}$ on $\boldsymbol{X}$ throughout the rest of this work. Note that alternative models such as random forests can also be use to obtain $\hat{m}(\boldsymbol{x}_i)$.[9,10]

3 | ESTIMATION OF RISK CONTROLLED ITRs

3.1 | Risk-controlled decision trees

An rcDT model is constructed given values of λ and τ. Recall that λ is a tuning parameter and τ is a fixed quantity selected based on clinical relevance. First, an initial split is made using an exhaustive search of all possible cut points across all candidate predictors. A class of candidate rules $d(\boldsymbol{X})$ is constructed for continuous or ordinal covariates as $d(\boldsymbol{X}) = \mathrm{I}(\boldsymbol{X}_j \leq c)$ or $d(\boldsymbol{X}) = \mathrm{I}(\boldsymbol{X}_j > c)$ where c is a candidate cut point for covariate j ($\forall j \in \{1 \ \dots \ p\}$). This partition is denoted as $\Omega = \Omega^1 \bigcup \Omega^2$ where $\Omega = \{\boldsymbol{x}_i : i = 1 \ \dots \ N\}, \Omega^1 = \{\boldsymbol{x}_i : x_{ij} \leq c, i = 1 \ \dots \ N\}$, and $\Omega^2 = \{\boldsymbol{x}_i : x_{ij} > c, i = 1 \ \dots \ N\}$. Further, let $\Omega^{11} \cup \Omega^{12} = \Omega^1$ represent a similar partition of Ω^1. For instance, $\Omega^{11} = \{\boldsymbol{x}_i : x_i \in \Omega^1, x_{ij'} \leq c', i = 1 \ \dots \ N\}$ for splitting covariate j' and cut point c'. Through recursively splitting, the candidate rule d_o is selected which maximizes the purity measure (Equation (6)). For $j \in \{1, \ \dots \ , p\}$, all possible candidate cuts c, and treatment assignments $(d, d') \in \{(0,1), (1,0)\}$ for the two resulting subspaces, the initial and daughter node splits are estimated as

Initial Split of Ω

$$\max_{j,c,(d,d')\in\{(0,1),(1,0)\}} \left(\frac{\sum_{i\in\Omega^1} \hat{w}_i y_i}{\sum_{i\in\Omega^1} \hat{w}_i} + \frac{\sum_{i\in\Omega^2} \hat{w}'_i y_i}{\sum_{i\in\Omega^2} \hat{w}'_i} \right) - \lambda \left[\left(\frac{\sum_{i\in\Omega^1} \hat{w}_i r_i}{\sum_{i\in\Omega^1} \hat{w}_i} + \frac{\sum_{i\in\Omega^2} \hat{w}'_i r_i}{\sum_{i\in\Omega^2} \hat{w}'_i} \right) - \tau \right],$$

Daughter Split of Ω^1

$$\max_{j,c,(d,d')\in\{(0,1),(1,0)\}} \left(\frac{\sum_{i\in\Omega^{11}} \hat{w}_i y_i}{\sum_{i\in\Omega^{11}} \hat{w}_i} + \frac{\sum_{i\in\Omega^{12}} \hat{w}'_i y_i}{\sum_{i\in\Omega^{12}} \hat{w}'_i} \right) - \lambda \left[\left(\frac{\sum_{i\in\Omega^{11}} \hat{w}_i r_i}{\sum_{i\in\Omega^{11}} \hat{w}_i} + \frac{\sum_{i\in\Omega^{12}} \hat{w}'_i r_i}{\sum_{i\in\Omega^{12}} \hat{w}'_i} \right) - \tau \right] + \frac{\sum_{i\in\Omega^2} \hat{w}^o_i y_i}{\sum_{i\in\Omega^2} \hat{w}^o_i y_i} - \lambda \left[\left(\frac{\sum_{i\in\Omega^2} \hat{w}^o_i r_i}{\sum_{i\in\Omega^2} \hat{w}^o_i} \right) - \tau \right],$$

where $\hat{w}_i = \frac{\mathrm{I}(a_i = d(\boldsymbol{x}_i))}{\hat{\Pr}(a_i|\boldsymbol{x}_i)}$, $\hat{w}'_i = \frac{\mathrm{I}(a_i = d'(\boldsymbol{x}_i))}{\hat{\Pr}(a_i|\boldsymbol{x}_i)}$, and $\hat{w}^o_i = \frac{\mathrm{I}(a_i = d_o(\boldsymbol{x}_i))}{\hat{\Pr}(a_i|\boldsymbol{x}_i)}$. In the daughter node purity measure $\hat{d}_o$ is the estimated optimal rule corresponding to the initial split of Ω. When considering a partition of daughter node Ω^1, information from Ω^2, the other daughter node, is incorporated into the splitting criteria. This allows for risk to be properly controlled at the population level throughout the entire construction of the tree. The following stopping criteria are employed: (1) no candidate daughter split can be found to increase the purity of the tree, (2) parent nodes contain too few training observations, and (3) treatment group sizes in parent nodes are not large enough. Each daughter node is selected for splitting in a random order and splitting continues until a stopping criteria is reached. Note that stopping criterion (1) ensures that $L(d)$ functions as a proper purity measure (ie, the purity of the tree structure is guaranteed to increase after splitting a parent node).

3.1.1 | Pruning an rcDT for optimal tree selection

Optimal rcDT model selection consists of identifying risk penalty parameter λ and cost complexity (ie, tree size) parameter α. Out of a sequence of candidate values of $\boldsymbol{\Lambda} = \{\lambda_s : s = 1, 2, \ \dots \ , S\}$, select $\lambda = \lambda_s$ and grow large tree Γ^s_0. Tree Γ^s_0 maximizes $L(d)$ given in Equation (4) with $\lambda = \lambda_s$. Using a CART (Classification And Regression Tree) pruning procedure, a sequence of cost-complexity parameter values based on the structure of tree Γ^s_0 is identified using the "weakest link" criteria. Briefly, a sequence of subtrees $\Gamma^s_M \prec \Gamma^s_{M-1} \prec \cdots \prec \Gamma^s_1 \prec \Gamma^s_0$ is constructed where "$\prec$" means "is subtree of."

Note that Γ_M^s is the root node and Γ_0^s is the full tree. Each element Γ_m^s for fixed s and $m = 1, 2, \ldots, M$ is constructed by pruning off the branch from Γ_{m-1}^s that results in the smallest reduction in value, that is, minimizes $|L(\Gamma_m^s) - L(\Gamma_{m-1}^s)|$.

The cost-complexity function for tree Γ_m^s is,

$$L_\alpha(\Gamma_m^s) = L(\Gamma_m^s) - \alpha \cdot \left|\tilde{\Gamma}_m^s\right|, \tag{7}$$

where $L(\Gamma_m^s)$ evaluates the purity measure in tree Γ_m^s. $L(\Gamma_m^s)$ is obtained by evaluating Equation (6) for the decision rule generated by tree Γ_m^s. $\tilde{\Gamma}_m^s$ is the set of terminal nodes of tree Γ_m^s and $\alpha > 0$ penalizes more complex models. A tree with larger $L_\alpha(\Gamma_m^s)$ is desirable. A unique sequence $\alpha_0 < \alpha_1 \cdots < \alpha_M$ corresponds to each subtree in the optimally pruned sequence of subtrees. The optimal value of α is selected via a k-fold cross-validation as proposed previously.[10]

This procedure yields $\{\Gamma_*^s : s = 1, 2, \ldots, S\}$, a set of optimally pruned trees, one corresponding to each value in $\mathbf{\Lambda}$. The final tree Γ^* is selected such that the cross-validated risk is less than τ and the reward is maximized. Defining a parsimonious set of candidate values λ can considerably decrease the computational burden required. Values of λ are roughly bounded by

$$0 < \lambda < \frac{\mathrm{E}(\delta_Y(\boldsymbol{X})) + \bar{\boldsymbol{Y}}_0 + \bar{\boldsymbol{R}}_0}{\mathrm{E}(\delta_R(\boldsymbol{X})) - \tau}, \tag{8}$$

where the right-hand side of the inequality is derived from Equation (5) and can easily be estimated from the training data. In summary, rcDT model selection is completed in three main steps: (1) fit a large rcDT for each $\mathbf{\Lambda} = \{\lambda_s : s = 1, 2, \ldots, S\}$, (2) for each value of λ_s, select the optimally pruned subtree Γ_*^s via cross-validation, and (3) from the set of optimally pruned trees $\{\Gamma_*^s : s = 1, 2, \ldots, S\}$, obtain the final model Γ^* that has the largest expected efficacy and maintains expected risk at the τ level.

3.2 | Risk-controlled random forest

An rcRF is constructed by aggregating several rcDT predictors via bootstrap sampling of the training data, typically called "bagging".[21] Given n training samples indexed by $i = 1, 2, \ldots, n$, draw bootstrap sample $\mathcal{B}$ of size n with replacement from the training samples. An rcDT model, denoted Γ^b, is then constructed from bootstrap sample $\mathcal{B}$. At each split $\max(\lfloor p/3 \rfloor, 1)$ covariates are selected as candidates for splitting.[22] Repeating this procedure B times yields an rcRF comprised of a set of rcDT predictors, denoted as $\mathcal{F} = \{\Gamma^b : b = 1, 2, \ldots, B\}$.

Tuning of the risk penalty parameter λ for an rcRF is accomplished via risk estimation in the out-of-bag sample. Let $\mathcal{B}_b$ represent the bth bootstrap sample for $b = 1, 2, \ldots, B$ and $\mathcal{B}_b^*$ represent the bth out-of-bag sample corresponding to $\mathcal{B}_b$, that is, the observations not included in the bth bootstrap sample. Let $\hat{\theta}_i(\Gamma^b)$ denote the predicted treatment for the ith observation derived from Γ^b. The cumulative predicted probability of recommendation to active treatment for the ith sample up to the rth tree for $r \leq B$ is recorded as,

$$\hat{\mathrm{Pr}}_{i,r}^*(\hat{a} = 1) = \left(\textstyle\sum_{b=1}^r \mathrm{I}(i \in \mathcal{B}_b^*)\right)^{-1} \displaystyle\sum_{b=1}^r \hat{\theta}_i(\Gamma^b) \cdot \mathrm{I}(i \in \mathcal{B}_b^*). \tag{9}$$

The estimator $\hat{\mathrm{Pr}}_{i,r}^*(\hat{a} = 1)$ is an ensemble estimator that was cross-validated using the out-of-bag sample and did not rely on the training observations. Cumulative out-of-bag prediction allows for an estimator of risk in the forest up to the rth tree, namely $\hat{V}_R^{\hat{a}_{i,r}^*}(d)$. Tuning of λ is carried out by first setting $\lambda = \lambda_0$ using an initial value λ_0. An rcRF (denoted as $\mathcal{F}_0$) is constructed and the risk estimate, R_0, obtained using Equation (3). If $|R_0 - \tau| < \varepsilon$ for some $\varepsilon > 0$ then keep $\mathcal{F}_0$ as the final model. The value of ε is user defined. If $|R_0 - \tau| > \varepsilon$ and $R_0 < \tau$, then set $\lambda_1 = \delta \cdot \lambda_0$ for some $0 < \delta < 1$. Otherwise, set $\lambda_1 = \delta \cdot \lambda_0$ for some $\delta > 1$. Fit a new rcRF model using $\lambda = \lambda_1$ and iterate until the out-of-bag risk is controlled at the τ level within the specified ε tolerance. To avoid using a grid search of candidate values of δ, we obtain the updated value of λ_{s+1} based on λ_s and out-of-bag risk from the current iteration, R_s. For instance, let $\delta = \left(1 + \frac{R_s - \tau}{\tau}\right)$. This factor inflates λ_s, that is, $\lambda_{s+1} > \lambda_s$, if the current risk estimate is greater than τ, and decreases values of λ if the current risk estimate is less than τ.

The probability of recommendation to treatment given in (9) is noted as a "soft" decision. A "hard" decision could be defined as $\hat{a}^*_{i,r} = \mathrm{I}(\hat{\Pr}^*_{i,r}(\hat{a} = 1) > 0.5)$. However, more stringent thresholds could be used, for example, $\hat{a}^*_{i,r} = \mathrm{I}(\hat{\Pr}^*_{i,r}(\hat{a} = 1) > 0.8)$ to be assigned to treatment.

3.2.1 | rcRF variable importance

Variable importance measures are defined for an rcRF through a permutation scheme applied to the out-of-bag sample. Suppose we have an rcRF consisting of B rcDT learners. For the bth rule, run out-of-bag sample $\boldsymbol{B}^*_b$ down tree Γ^b to obtain $\hat{L}^*(\hat{\boldsymbol{\theta}}(\Gamma^b))$. Let j_m correspond to a covariate used in the construction of tree Γ^b. Permute the values of j_m in the out-of-bag sample and run the permuted data down tree Γ_b to obtain the objective value for the permuted covariate $\hat{L}^{*m}(\hat{\boldsymbol{\theta}}(\Gamma^b))$. Repeat this for all predictors included in tree Γ_b. Record the variable importance measure for predictor j_m derived from tree Γ_b as $\mathrm{VI}^b_m = \{\hat{L}^*\hat{\boldsymbol{\theta}}(\Gamma_b)) - \hat{L}^{*m}\{\hat{\boldsymbol{\theta}}(\Gamma_b))\}^+$, where $(x)^+ = x$ if $x > 0$; $(x)^+ = 0$ if $x \leq 0$. This considers a predictor as "important" if permuting its values results in the objective function decreasing in value and quantifies the importance as the magnitude of the decrease. The total importance of predictor j_m is then the sum of importances across the B trees. Importances are scaled to sum to one for ease of interpretation. This is called the total importance measure (ie, combines efficacy and risk information). Two additional importance measures can be similarly defined, one of each for efficacy and risk. Estimation of importance for efficacy and risk scores is accomplished by replacing $\hat{L}$ with $\hat{V}_Y(d)$ and $-\hat{V}_R(d)$, respectively. Given an rcRF model, the importance of a predictor is characterized in three ways, (1) a total importance which measures the relative contribution of the predictor to maximizing the purity measure, (2) an efficacy importance which measures the relative contribution of the predictor to achieving greater efficacy, and (3) a risk importance which measures the relative contribution of the predictor to maintaining risk at the τ level.

4 | SIMULATION STUDIES

Performance of rcDT and rcRF methods was assessed via simulation study (Table 1). Each scheme has 10 predictors, $X_1, X_2, \ldots, X_{10}$, generated from a uniform distribution. Treatment indicator $A \in \{0, 1\}$ was simulated to mimic an RCT design with a 1:1 allocation ratio to two treatment arms, that is, $\Pr(\boldsymbol{A}|\boldsymbol{X}) = 0.5$. Random noise for efficacy scores was generated from a standard normal distribution. For risk scores, random noise was generated as $N(\mu = 0, \sigma^2 = 0.5^2)$. In all schemes, the optimal rule for a plausible value of τ is defined by X_1 and X_2 detailed below. Efficacy and risk scores were based on formulas presented in Table 1. Schemes A to C have risk controlled ITR defined by a rectangle. Schemes D and E appeared in Wang et al[14] and have a risk controlled ITR favoring linear and quadratic boundaries, respectively.

Scheme A has optimal treatment regions defined by the rectangle $X_1 \leq 0.6$ and $X_2 \leq h$. The value of h depends on the desired risk constraint τ. If risk constraint is ignored the optimal rule recommends observations with $X_1 \leq 0.6$ to active treatment and all others to control, yielding efficacy score of 3.50 and a risk score of 3.20. For observations with $X_1 \leq 0.6$, the expected benefit from receiving active treatment vs control decreases as X_2 increases. For the entire population, larger values of X_2 increase expected risk from receiving treatment vs control. If no modeling was performed and all patients were assigned to control, that is, the least risky rule, the expected risk is 1.90. Hence, risk constraints were set to be $\tau = 2.20, 2.50, 2.80$. The optimal ITRs are defined by the rectangles $X_1 \leq 0.6$ and $X_2 \leq h$ for $h = 0.43, 0.65, 0.80$, respectively. This set of optimal ITRs recommends 27%, 39%, and 49% of observations to active treatment.

Scheme B has an optimal treatment region defined by the rectangle $X_1 \leq 0.7$ and $X_2 \leq h$. The value of h depends on the desired risk constraint τ. The unconstrained ITR recommends observations with $X_1 \leq 0.7 \cap X_2 \leq 0.7$ to treatment and control otherwise, with an efficacy score of 3.97 and risk score of 3.17. A constant efficacy reward of +2 is given to treated observations with $X_1 < 0.7$ and $X_2 < 0.7$ and also to control observations with $X_1 > 0.7$ or $X_2 > 0.7$. The risk function for Scheme B is identical to Scheme A. Assigning all observations to control has an expected risk of 2.40. Hence, risk constraints of $\tau = 2.50, 2.75$ and 3.00 are considered. The accompanying ITRs are defined by the rectangle $X_1 \leq 0.7$ and $X_2 \leq h$ for $h = 0.20, 0.45, 0.60$. These rules send 15%, 32%, and 42% of observations to active treatment. Simulation schemes A and B are similar in that the risk score generating functions both depend on $\boldsymbol{X}_2$ in the same way (ie, $(2\boldsymbol{X}_2 - 0.1)(2\boldsymbol{A} - 1)$). In scheme A, the benefit received from a treatment / subgroup match (ie, $[A = 1$ and in $S]$ OR $[A = 0$ and not in $S])$ decreases as X_2 increases. In scheme B, the same treatment / subgroup match always yields a constant benefit of 3 units. These two schemes were intended to assess performance under two "tree-favorable" scenarios where efficacy scales with one of the covariates vs a constant "on/off" efficacy benefit.

TABLE 1 Simulation schemes

Scheme	Model	Subgroup S optimally assigned to treatment if $\tau = \tau^*$
A	$Y = 3 - 2X_3 + X_4$	$\tau^* = +\infty \rightarrow S = \{X_1 \leq 0.6\}$
	$+(2 - 2X_2)(\mathrm{I}(x \in S) - \mathrm{I}(x \notin S))(2A - 1) + \varepsilon_Y$	$\tau^* = 2.80 \rightarrow S = \{X_1 \leq 0.6 \cap X_2 \leq 0.80\}$
	$R = 3 + X_3 - X_4 + (2X_2 + 0.1)(2A - 1) + \varepsilon_R$	$\tau^* = 2.20 \rightarrow S = \{X_1 \leq 0.6 \cap X_2 \leq 0.43\}$
B	$Y = 1 - X_3 + X_4 + 3\mathrm{I}(x \in S)\mathrm{I}(A = 1)$	$\tau^* = +\infty \rightarrow S = \{X_1 \leq 0.7 \cap X_2 \leq 0.7\}$
	$+3\mathrm{I}(x \notin S)\mathrm{I}(A = 0) + \varepsilon_Y$	$\tau^* = 3.00 \rightarrow S = \{X_1 \leq 0.7 \cap X_2 \leq 0.6\}$
	$R = 2 + 2X_3 + X_4 + (2X_2 + 0.1)(2A - 1) + \varepsilon_R$	$\tau^* = 2.50 \rightarrow S = \{X_1 \leq 0.7 \cap X_2 \leq 0.2\}$
C	$Y = 3 - 2X_3 + X_4 + 2(1 - \max(X_1, X_2))\mathrm{I}(x \in S)(2A - 1)$	$\tau^* = +\infty \rightarrow S = \{X_1 \leq 0.7 \cap X_2 \leq 0.7\}$
	$-\mathrm{I}(x \notin S)(2A - 1) + \varepsilon_Y$	$\tau^* = 2.20 \rightarrow S = \{X_1 \leq 0.65 \cap X_2 \leq 0.65\}$
	$R = 2 + X_3 + (\max(X_1, X_2))(2A - 1) + \varepsilon_R$	$\tau^* = 1.90 \rightarrow S = \{X_1 \leq 0.44 \cap X_2 \leq 0.44\}$
D	$Y = 1 - 2X_1 + X_2 - X_3 + 2(1 - X_1 - X_2)(2A - 1) + \varepsilon_Y$	$\tau^* = +\infty \rightarrow S = \{X_1 + X_2 \leq 1\}$
	$R = 2 + X_1 + (1 + X_1 - X_2)(2A - 1) + \varepsilon_R$	$\tau^* = 2.25 \rightarrow S = \{1.2X_1 + X_2 \leq 1\}$
		$\tau^* = 1.75 \rightarrow S = \{3X_1 + X_2 \leq 1\}$
E	$Y = 1 - 2X_1 + X_2 - X_3 + 8(1 - X_1^2 - X_2^2)(2A - 1) + \varepsilon_Y$	$\tau^* = +\infty \rightarrow S = \{X_1^2 + X_2^2 \leq 1\}$
	$R = 2 + X_1 + (X_1 + X_2 - 0.3)(2A - 1) + \varepsilon_R$	$\tau^* = 2.20 \rightarrow S \approx \{X_1^{1.5} + X_2^{1.5} \leq 0.86^{1.5}\}$
		$\tau^* = 2.00 \rightarrow S \approx \{X_1^{1.3} + X_2^{1.3} \leq 0.76^{1.3}\}$

Note: $\mathrm{I}(\cdot)$ is an indicator function that evaluates to 1 if "·" is true and 0 otherwise.

Scheme C has an optimal treatment region defined as the rectangle $X_1 \leq h$ and $X_2 \leq h$. The value of h depends on the risk constraint τ. The unconstrained optimal rule sends observations with $X_1 \leq 0.7 \cap X_2 \leq 0.7$ to treatment and all others to control, yielding an efficacy score of 3.52 and a risk score of 2.29. Observations with larger values of either X_1 or X_2 are at greater risk on treatment. Observations with larger values of either X_1 or X_2 receiving treatment are also expected to have smaller efficacy scores if $X_1 \leq 0.7$ and $X_2 \leq 0.7$. Observations with $X_1 > 0.7$ or $X_2 > 0.7$ receiving control receive an expected benefit of +1. Assigning all observations to control carries an expected risk of 1.75. Risk constraints of $\tau = 1.90, 2.05, 2.20$ are considered with optimal treatment regions defined by rectangles $X_1 \leq h$ and $X_2 \leq h$ for $h = 0.44, 0.57, 0.65$, respectively. These rules recommend 20%, 33%, and 43% of observations to treatment.

Scheme D has an optimal treatment region defined as a linear combination of X_1 and X_2. The slope of the line is dependent on the risk constraint level τ. The unconstrained rule recommends treatment if $X_1 + X_2 \leq 1$ and control otherwise, carrying an efficacy score of 0.670 with a risk score of 2.520. Assigning all observations to control lowers the expected risk to 1.50. Risk constraints of $\tau = 1.75, 2.00, 2.25$ were considered. **Scheme E** has an optimal treatment region defined on an ellipse for efficacy scores and a linear function in the risk scores. The unconstrained rule sends observations with $X_1^2 + X_2^2 \leq 1$ to treatment providing an efficacy score of 3.62 and a risk score of 2.67. Recommending control to all patients lowers risk to 1.80. Risk constraints of $\tau = 2, 2.1, 2.2$ were considered.

Figure 1 presents an X-ray plot of the optimal treatment regimes under the risk constraints τ. Regions colored black indicate optimal assignment to control. Regions colored gray indicate an optimal recommendation of treatment at a particular risk threshold. As the risk threshold is relaxed, the shade of gray darkens to black, and a greater proportion of patients are recommended to active treatment. For instance, in simulation scheme A, risk constraints of $\tau = 2.20, 2.50$, and 2.80 correspond to light, medium, and dark gray coloring, respectively.

Risk-controlled ITRs for all simulation studies were estimated using rcDT and rcRF methods along with comparator methods, BR-M and BR-O (Wang et al[14]). Briefly, BR-M fits linear models to efficacy and risk outcomes using first-order interactions between predictors and binary treatment indicator. Contrasts for efficacy and risk outcomes derived from the fitted models are then used to explicitly derive the optimal treatment rule. BR-O is a machine learning algorithm that translates the constrained optimization into an unconstrained O-learning problem with enforcement of risk control. Readers are directed to Wang et al[14] for further details on BR-M and BR-O model fitting.

Estimates of efficacy and risk along with treatment assignment accuracy were calculated and reported for each simulation from a 20 000 observation validation set. A total of 100 replicates for each simulation study were used. For each simulation setting training sample sizes of $n = 500$ and 1000 were considered. For rcDT model, tuning parameters λ and α

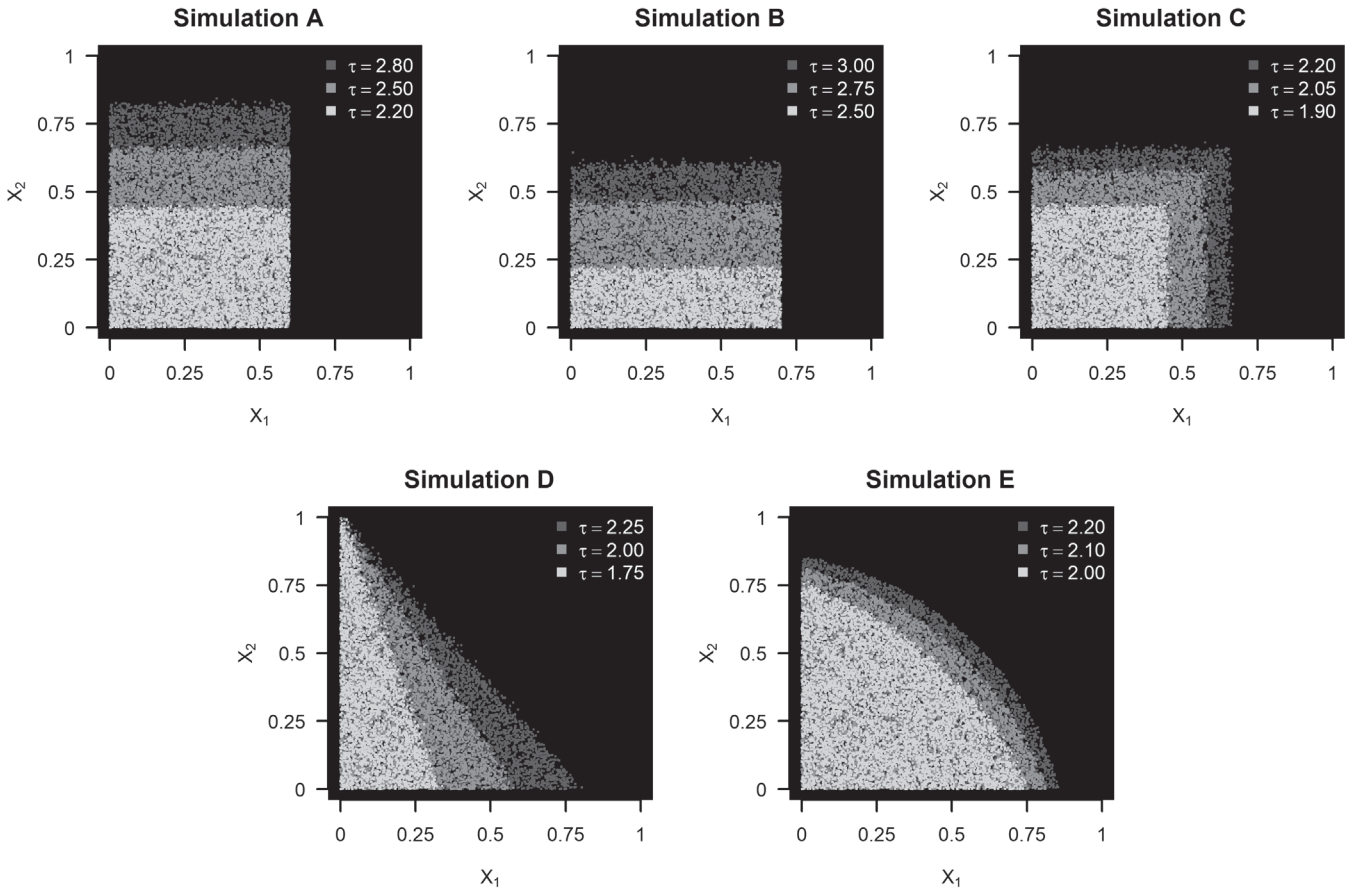

FIGURE 1 Optimal treatment assignment at different risk threshold levels for 20 000 validation observations. Shaded regions in each plot correspond to optimal treatment assignment under a given level of risk control (gray scale: active treatment; black: control)

were selected using 5-fold cross-validation. For rcRF model fitting, λ was selected using out-of-bag risk estimates. Each rcRF model consisted of 500 rcDT predictors, and variable importance measure were calculated for each rcRF model. Residuals for efficacy scores were estimated using fitted values from a linear model that regressed efficacy scores ($\boldsymbol{Y}$) on the 10 covariates ($\boldsymbol{X}$). Efficacy and risk summaries are presented as mean (SD) and median (median absolute deviation). Accuracy was calculated as proportion of the 20 000 validation observations recommended to the optimal treatment. Mean and median accuracy summaries were similar and median was selected for presentation. The final two columns of Tables 2 and 3 provide the percent of the 100 simulation replicates with mean validation set risk scores falling below the specified threshold (τ) or below a 5% increase of τ. Tables 2 and 3 summarize efficacy, risk, and accuracy results from simulation studies A and D. Results from simulation studies B, C, and E are provided in Appendix S1 (Tables S1, S2, and S3). BR-M models were fit using the procedure specified in Wang et al.[14]

In simulation scheme A, all four procedures control risk at the specified value of τ. Consider $\tau = 2.50$ and $n = 1000$ from simulation scheme A. Mean risk scores from rcDT, rcRF, BR-M, and BR-O were 2.491 (SD = 0.099), 2.465 (SD = 0.128), 2.499 (SD = 0.035), and 2.462 (SD = 0.109), respectively. This indicates that risk is controlled on average at the $\tau = 2.50$ level. Tree-based modeling identified rules with greater efficacy scores on average than BR-M or BR-O. Mean efficacy scores for rcDT, rcRF, BR-M, and BR-O methods were 3.350 (SD = 0.060), 3.329 (SD = 0.117), 3.203 (SD = 0.025), 2.831 (SD = 0.106), respectively. Note that for $\tau = 2.50$ the optimal rule has an efficacy score of 3.38. Median accuracies for rcDT, rcRF, BR-M, and BR-O were 95.2%, 95.8%, 83.4%, and 71.0%, respectively. In simulation scheme A, both BR-M and BR-O tend to propose rules that control risk, but fail to optimize efficacy.

Simulation scheme D has an optimal treatment assignment regime defined by a linear combination of X_1 and X_2. Consider $\tau = 1.75$ with $n = 1000$ training observations. All three methods control risk at the specified level. Validation set mean risk estimates for rcDT, rcRF, BR-M, and BR-O were 1.751 (SD = 0.053), 1.746 (SD = 0.038), 1.751 (SD = 0.026), and 1.711 (SD = 0.063), respectively. A less conservative risk constraint of $\tau = 2.25$ results in mean risk estimates for rcDT, rcRF, BR-M, and BR-O of 2.228 (SD = 0.083), 2.237 (SD = 0.053), 2.256 (SD = 0.037), and 2.135 (SD = 0.098), respectively.

TABLE 2 Simulation results for scheme A

Risk (τ)	Opt Efficacy	n	Method	Accuracy	Mean Risk	Mean Efficacy	Median Risk	Median Efficacy	% Risk $< \tau$	% Risk $< 1.05 \cdot \tau$
2.20	3.12	500	rcDT	0.863 (0.078)	2.173 (0.159)	2.900 (0.267)	2.149 (0.199)	2.853 (0.357)	57%	76%
			rcRF	0.904 (0.032)	2.174 (0.088)	3.002 (0.128)	2.165 (0.090)	3.002 (0.128)	64%	93%
			BRM	0.876 (0.013)	2.200 (0.053)	2.988 (0.050)	2.197 (0.054)	2.986 (0.053)	51%	97%
			BRO	0.740 (0.015)	2.190 (0.125)	2.590 (0.131)	2.180 (0.139)	2.577 (0.120)	58%	85%
		1000	rcDT	0.888 (0.076)	2.199 (0.160)	2.968 (0.267)	2.238 (0.166)	3.061 (0.243)	41%	67%
			rcRF	0.916 (0.021)	2.191 (0.077)	3.054 (0.107)	2.204 (0.074)	3.074 (0.101)	48%	96%
			BRM	0.880 (0.012)	2.196 (0.041)	2.997 (0.036)	2.199 (0.038)	2.997 (0.036)	55%	99%
			BRO	0.741 (0.012)	2.148 (0.090)	2.547 (0.095)	2.147 (0.091)	2.551 (0.091)	71%	97%
2.50	3.38	500	rcDT	0.930 (0.035)	2.491 (0.144)	3.325 (0.085)	2.485 (0.143)	3.332 (0.089)	51%	83%
			rcRF	0.932 (0.028)	2.465 (0.109)	3.310 (0.085)	2.471 (0.084)	3.330 (0.060)	61%	95%
			BRM	0.831 (0.013)	2.499 (0.054)	3.191 (0.034)	2.497 (0.061)	3.193 (0.036)	54%	100%
			BRO	0.719 (0.027)	2.462 (0.152)	2.862 (0.148)	2.449 (0.154)	2.866 (0.147)	59%	86%
		1000	rcDT	0.952 (0.036)	2.491 (0.099)	3.350 (0.060)	2.502 (0.089)	3.363 (0.053)	48%	92%
			rcRF	0.958 (0.021)	2.461 (0.128)	3.329 (0.117)	2.488 (0.080)	3.359 (0.042)	55%	96%
			BRM	0.834 (0.012)	2.496 (0.035)	3.203 (0.025)	2.494 (0.032)	3.207 (0.023)	60%	100%
			BRO	0.710 (0.018)	2.430 (0.109)	2.831 (0.106)	2.415 (0.105)	2.829 (0.091)	74%	97%
2.80	3.46	500	rcDT	0.925 (0.031)	2.770 (0.165)	3.407 (0.059)	2.768 (0.182)	3.426 (0.053)	53%	82%
			rcRF	0.929 (0.021)	2.746 (0.133)	3.395 (0.070)	2.754 (0.104)	3.401 (0.036)	71%	94%
			BRM	0.838 (0.011)	2.796 (0.057)	3.289 (0.029)	2.794 (0.066)	3.287 (0.028)	54%	100%
			BRO	0.780 (0.048)	2.746 (0.171)	3.115 (0.147)	2.756 (0.172)	3.135 (0.154)	60%	87%
		1000	rcDT	0.946 (0.029)	2.797 (0.107)	3.438 (0.033)	2.799 (0.113)	3.443 (0.030)	51%	91%
			rcRF	0.953 (0.019)	2.772 (0.091)	3.431 (0.026)	2.768 (0.092)	3.432 (0.030)	64%	96%
			BRM	0.842 (0.010)	2.795 (0.043)	3.305 (0.022)	2.793 (0.044)	3.303 (0.024)	55%	100%
			BRO	0.771 (0.030)	2.736 (0.107)	3.104 (0.103)	2.730 (0.114)	3.110 (0.105)	72%	99%

Abbreviations: "% Risk $< \tau$", Percent of simulation replicates with validation set risk estimate less than τ; "% Risk $< 1.05 \cdot \tau$", Percent of simulation replicates with validation set risk estimate less than $1.05 \cdot \tau$; "Accuracy", median (median absolute deviation) proportion receiving optimal treatment assignment; "Mean Efficacy" and "Mean Risk", mean (sd) summaries for predicted efficacy and risk scores; 'Median Efficacy' and "Median Risk", median (median absolute deviation) summaries for predicted efficacy and risk scores; "n", Training sample size; "Opt Efficacy", maximum achievable efficacy under the optimal treatment assignment; "Risk", Risk threshold.

Despite the true optimal treatment assignments aligning to a linear rule, both rcDT and rcRF are competitive with BR-M and BR-O in maximizing efficacy. When $\tau = 1.75$, mean validation set efficacy estimates from rcDT, rcRF, BR-M, and BR-O models were 0.327 (SD = 0.049), 0.336 (SD = 0.040), 0.356 (SD = 0.025), and 0.278 (SD = 0.068). When $\tau = 2.25$, mean validation set efficacy estimates for rcDT, rcRF, BR-M, and BR-O models were 0.558 (SD = 0.028), 0.597 (SD = 0.016), 0.618 (SD = 0.010), and 0.572 (SD = 0.042), respectively. When τ is highly conservative, $\tau = 1.75$, median accuracies for rcDT, rcRF, BR-M, and BR-O were 92.4%, 94.1%, 96.8%, and 90.6%, respectively. Note that rcDT can approximate the optimal treatment region for simulation scheme D with $\tau = 1.75$ reasonably well as a narrow rectangle. When τ is less conservative, for example, $\tau = 2.25$, median accuracy for rcDT dropped to 86.4% while rcRF was 92.1%, (BR-M and BR-O were 96.2% and 90.7%, respectively). This highlights the robustness of rcRF when the underlying optimal treatment assignments may not conform to a rectangular region of the predictor space. Readers can peruse results from simulation scheme E in Appendix S1 as another demonstration for the robustness of rcRF method.

Approximately 40% to 60% of simulation replicates across all methods for all simulation studies controlled risk at the τ level in the validation set. Risk is controlled by rcRF, BR-M, and BR-O in the validation set in greater than 90% of

TABLE 3 Simulation results for scheme D

Risk (τ)	Opt efficacy	n	Method	Accuracy	Mean Risk	Mean Efficacy	Median Risk	Median Efficacy	% Risk $< \tau$	% Risk $< 1.05 \cdot \tau$
1.75	0.36	500	rcDT	0.910 (0.024)	1.753 (0.071)	0.316 (0.066)	1.754 (0.062)	0.319 (0.068)	48%	90%
			rcRF	0.935 (0.015)	1.753 (0.053)	0.338 (0.050)	1.747 (0.051)	0.335 (0.050)	53%	92%
			BRM	0.956 (0.008)	1.755 (0.042)	0.354 (0.040)	1.754 (0.042)	0.354 (0.039)	46%	95%
			BRO	0.901 (0.011)	1.726 (0.095)	0.290 (0.095)	1.726 (0.088)	0.294 (0.091)	62%	86%
		1000	rcDT	0.924 (0.015)	1.751 (0.053)	0.327 (0.049)	1.750 (0.059)	0.332 (0.057)	51%	94%
			rcRF	0.941 (0.014)	1.746 (0.038)	0.336 (0.040)	1.748 (0.038)	0.341 (0.037)	53%	99%
			BRM	0.968 (0.006)	1.751 (0.026)	0.356 (0.025)	1.750 (0.025)	0.355 (0.029)	51%	100%
			BRO	0.906 (0.009)	1.711 (0.063)	0.278 (0.068)	1.715 (0.056)	0.285 (0.057)	75%	98%
2.00	0.54	500	rcDT	0.873 (0.030)	2.009 (0.103)	0.477 (0.057)	2.000 (0.094)	0.481 (0.050)	49%	82%
			rcRF	0.922 (0.018)	1.988 (0.060)	0.514 (0.033)	1.991 (0.070)	0.521 (0.035)	58%	98%
			BRM	0.953 (0.009)	2.009 (0.038)	0.542 (0.020)	2.009 (0.034)	0.542 (0.019)	39%	98%
			BRO	0.897 (0.021)	1.949 (0.117)	0.485 (0.081)	1.940 (0.122)	0.498 (0.078)	66%	91%
		1000	rcDT	0.883 (0.016)	1.996 (0.071)	0.485 (0.039)	1.996 (0.072)	0.490 (0.038)	57%	90%
			rcRF	0.928 (0.015)	1.989 (0.052)	0.517 (0.038)	1.989 (0.034)	0.523 (0.022)	62%	100%
			BRM	0.965 (0.007)	2.004 (0.033)	0.544 (0.017)	2.004 (0.034)	0.543 (0.016)	40%	99%
			BRO	0.899 (0.022)	1.924 (0.096)	0.473 (0.068)	1.923 (0.109)	0.481 (0.061)	78%	94%
2.25	0.62	500	rcDT	0.848 (0.023)	2.208 (0.118)	0.535 (0.040)	2.209 (0.111)	0.540 (0.040)	61%	89%
			rcRF	0.908 (0.021)	2.215 (0.082)	0.584 (0.027)	2.221 (0.069)	0.591 (0.021)	67%	96%
			BRM	0.948 (0.012)	2.258 (0.044)	0.615 (0.012)	2.256 (0.044)	0.616 (0.011)	42%	98%
			BRO	0.894 (0.033)	2.124 (0.149)	0.558 (0.070)	2.124 (0.130)	0.569 (0.052)	81%	95%
		1000	rcDT	0.864 (0.026)	2.228 (0.083)	0.558 (0.028)	2.243 (0.081)	0.560 (0.031)	58%	95%
			rcRF	0.921 (0.013)	2.237 (0.053)	0.597 (0.016)	2.240 (0.050)	0.599 (0.016)	57%	99%
			BRM	0.962 (0.009)	2.256 (0.037)	0.618 (0.010)	2.255 (0.040)	0.618 (0.010)	46%	100%
			BRO	0.907 (0.036)	2.135 (0.098)	0.572 (0.042)	2.130 (0.099)	0.575 (0.043)	87%	99%

Abbreviations: "% Risk $< \tau$", Percent of simulation replicates with validation set risk estimate less than τ; "% Risk $< 1.05 \cdot \tau$", Percent of simulation replicates with validation set risk estimate less than $1.05 \cdot \tau$"Accuracy", median (median absolute deviation) proportion receiving optimal treatment assignment; "Mean Efficacy" and "Mean Risk", mean (SD) summaries for predicted efficacy and risk scores; 'Median Efficacy' and "Median Risk", median (median absolute deviation) summaries for predicted efficacy and risk scores; "n", Training sample size; "Opt Efficacy", maximum achievable efficacy under the optimal treatment assignment; "Risk", Risk threshold.

simulation replicates at the $1.05 \cdot \tau$ level. For instance, in simulation scheme A with $\tau = 2.20$ and training sample of size $n = 500$, 64% of rcRF simulation replicates produced validation set risk estimates less than $\tau = 2.20$ and 93% of simulation replicates produced validation set risk estimates less than $1.05 \cdot \tau = 2.31$. Due to the variability inherent in tree models, rcDT modeling controlled risk at the τ and 5% excess of τ levels in 57% and 76% of simulation replicates, respectively, under the same simulation settings. Risk estimates under the stated settings for rcDT simulation replicates ranged from 1.92 to 2.51 and risk estimates from rcRF ranged from 2.00 to 2.44.

While asymptotic properties of the proposed methods cannot be derived, simulation studies reveal that as training sample size increases (1) accuracy of the rule increases, (2) efficacy increases, (3) risk is controlled closer to the τ level, and (4) the rule becomes more stable (ie, precision estimates decrease). A summary of the computational cost associated with model fitting for simulation schemes A and D is included in Appendix S1 (Table S8). Briefly, consider simulation scheme A. When the training sample size is $n = 500$, rcDT model fitting was typically completed in under 1 minute, rcRF in slightly more than 5 minutes, BR-M in under 1/10 of a second, and BR-O in 3 to 5 minutes. Times were fairly consistent across different values of τ. When the training sample size is $n = 1000$, model fitting was typically completed in under

four minutes for rcDT, 7 to 12 minutes for rcRF (faster times for small τ values), under 1/10 of a second for BR-M, and 31 to 36 minutes for BR-O (shorter times for larger τ values). Note that BR-M is expected to be much faster than the other methods since only simple linear model fits are required for estimating the treatment rule. As noted, the computational cost of BR-O grows rapidly as dimensionality of the training data increases even modestly.[14] The rcRF learner scales better than BR-O as dimensionality increases. In addition, the computational cost of rcRF can be controlled by specifying the number of decision trees and the depth of the individual rcDT learners. Variable importance measure summaries for each simulation scheme are presented in Appendix S1.

To demonstrate use of rcDT and rcRF modeling when original treatment assignments were not randomized, model fitting was performed using the efficacy and risk models outlined in simulation scheme D, but with original treatment assignments generated from an observational study design as opposed to an RCT design (see Table S4 in Appendix S1). This introduces additional variability to the modeling process as propensity scores are drawn from a generative model and then used to generate treatment assignments. Under the observational design, rcDT and rcRF models have mean risk scores above τ for smaller values of τ. For $\tau = 1.75$ and 2.00, risk is typically controlled. Under conservative risk thresholds, higher risk is accompanied by higher efficacy on average (eg, around +25%). The proportion receiving the optimal treatment assignment decreases slightly under the observational design, typically dropping by a couple of percentage points.

5 | ANALYSIS: DURABLE TRIAL

rcDT, rcRF, and BR-M methods were applied to the DURABLE trial.[17,23] Briefly, the trial investigated the use of twice-daily insulin lispro mix 75/25 (LMx75/25) vs once-daily basal insulin glargine (Glargine) in patients with type 2 diabetes. There were 18 covariates available at baseline including body mass index, height, weight, adiponectin, systolic and diastolic blood pressure, duration of diabetes, heart rate, HbA1c, fasting insulin, fasting glucose, and a 7-point self monitored blood sugar measure taken throughout the day. Of the 2187 patients enrolled, 1498 were retained for analysis. A more detailed patient flow diagram can be found in Appendix S1 (Figure S6). Patients were excluded from the analysis set if they had incomplete follow up data or extreme values in either the covariates or outcomes. The primary efficacy outcome was change in HbA1c from baseline to 24-week follow-up. The risk outcome was daily rate of hypoglycemia. Data from this study are not publicly available. Table 4 presents baseline covariate summaries for the DURABLE analysis dataset. All predictor values are well balanced at baseline between the two treatment groups. Mean decrease in HbA1c from baseline to end of follow-up was 1.79 (SD = 1.44) and 1.87 (SD = 1.45) for Glargine and LMx75/25 groups, respectively. Mean daily hypoglycemia event rates were 0.057 (SD = 0.069) and 0.074 (SD = 0.081) for Glargine and LMx75/25 groups, respectively. This provides some evidence that the more aggressive treatment, LMx75/25, yielded greater control of HbA1c over the 24-week follow-up, accompanied by higher rates of hypoglycemia. Three risk constraints were considered ($\tau = 0.063, 0.065, 0.067$) spanning the range of risk values obtained from recommending all patients receive more conservative (Glargine) to the more aggressive (LMx75/25) treatment. These risk control levels were selected to mirror analyses performed by Wang et al.[14]

Modeling performance was assessed using a 1:1 random split of the available observations. A model was trained using half of the data and validated using the other half. This was replicated 100 times for each modeling procedure. Each rcDT model was trained using 5-fold cross validation to select tuning parameters. Each rcRF model was trained using a forest of 500 trees with the tuning parameter selected using the out-of-bag estimate of risk. BR-M models were fit using the procedure outlined in Wang et al.[14] BR-O was not included here due to the high computational cost and the relatively good performance of BR-M in the simulation studies. Estimates of the efficacy and risk from training and validation sets was reported for each model fitting procedure. Results are summarized in Table 5. Risk is controlled close to the τ level in the validation sets for all three procedures across all risk constraints. BR-M and rcDT tended to produce a more conservative rule, especially as the risk constraint is relaxed, yielding more control over risk at the expense of loss in efficacy. rcRF models not only control risk well on average in the validation sets, but also pick up more efficacy compared to rcDT and BR-M methods. When $\tau = 0.063$, rcDT, rcRF, and BR-M methods produced validation risk estimates of 0.0638 (SD = 0.0061), 0.0643 (SD = 0.0035), and 0.0614 (SD = 0.0054). The accompanying efficacy estimates were 1.79 (SD = 0.11), 1.80 (SD = 0.07), and 1.76 (SD = 0.11), respectively. When the risk constraint is relaxed to $\tau = 0.067$, rcDT, rcRF, and BR-M produced respective validation risk estimates of 0.0645 (SD = 0.0058), 0.0666 (SD = 0.0035), and 0.0648 (SD = 0.0057) with validation efficacy estimates of 1.79 (SD = 0.12), 1.80 (SD = 0.06), and 1.79 (SD = 0.11), respectively.

TABLE 4 Cohort characteristics for DURABLE trial data

Characteristic	Glargine ($n = 754$)	LMx75/25 ($n = 744$)	P-value[a]
Reduction in HbA1c	1.79 (1.44)	1.87 (1.45)	.300
Daily hypoglycemic event rate	0.0568 (0.0686)	0.0741 (0.0808)	< .001
Baseline HbA1c	9.1 (1.2)	9.1 (1.3)	.521
Duration diabetes	9.5 (6.0)	10.0 (6.4)	.115
Heart rate	77.1 (10.0)	76.7 (10.1)	.448
Systolic BP	132.0 (16.4)	131.5 (16.4)	.526
Diastolic BP	78.4 (9.3)	78.3 (9.3)	.910
BMI	31.7 (6.0)	31.8 (6.0)	.719
Height	166.1 (10.9)	166.7 (10.5)	.325
Weight	87.9 (20.9)	88.8 (20.9)	.396
Glucose: Nighttime (3 AM)	198.8 (62.4)	197.6 (60.8)	.705
Glucose: Evening after meal	240.9 (67.7)	240.6 (64.5)	.934
Glucose: Evening before meal	205.4 (64.4)	202.2 (63.1)	.330
Glucose: Noon after meal	233.9 (67.7)	234.3 (65.6)	.923
Glucose: Noon before meal	205.1 (66.4)	206.1 (64.4)	.761
Glucose: Morning after meal	252.3 (62.7)	256.1 (63.3)	.247
Glucose: Morning before meal	198.2 (54.0)	194.9 (51.5)	.227
Fasting insulin	9.8 (7.3)	10.1 (7.4)	.402
Adiponectin	6.9 (5.2)	6.8 (5.5)	.955
Fasting glucose	11.0 (3.5)	11.3 (3.7)	.139

Abbreviations: BMI, body mass index; BP, blood pressure; HbA1c, Glycated hemoglobin; DURABLE trial, Assessing the DURAbility of Basal vs Lispro Mix 75/25 Insulin Efficacy.
[a] from two-sample t-test.

TABLE 5 DURABLE trial analysis results comparing rcDT, rcRF, and BR-M methods

Risk (τ)	Method	Efficacy (Training)	Efficacy (Validation)	Risk (Training)	Risk (Validation)
0.063	rcDT	1.835 (0.093)	1.789 (0.112)	0.0567 (0.0051)	0.0638 (0.0061)
	rcRF	1.749 (0.089)	1.799 (0.065)	0.0520 (0.0035)	0.0643 (0.0035)
	BR-M	1.847 (0.052)	1.756 (0.105)	0.0597 (0.0017)	0.0614 (0.0054)
0.065	rcDT	1.855 (0.098)	1.796 (0.122)	0.0579 (0.0054)	0.0641 (0.0059)
	rcRF	1.921 (0.083)	1.804 (0.061)	0.0615 (0.0050)	0.0657 (0.0040)
	BR-M	1.871 (0.057)	1.778 (0.111)	0.0620 (0.0020)	0.0630 (0.0057)
0.067	rcDT	1.855 (0.088)	1.792 (0.120)	0.0586 (0.0053)	0.0645 (0.0058)
	rcRF	2.004 (0.027)	1.804 (0.064)	0.0677 (0.0010)	0.0666 (0.0035)
	BR-M	1.889 (0.063)	1.792 (0.106)	0.0643 (0.0025)	0.0648 (0.0057)

Abbreviations: BR-M, model-based benefit-risk learning;[14] rcDT, risk controlled decision tree; rcRF, risk controlled random forest.

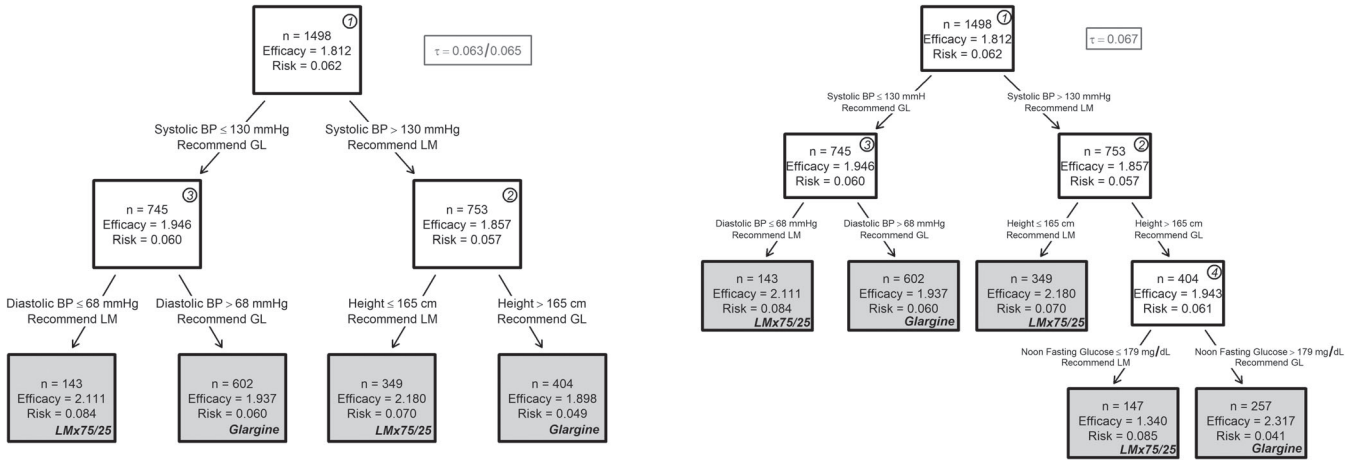

FIGURE 2 DURABLE risk-controlled decision tree structure for controlling daily hypoglycemic event rate at $\tau = 0.063$ or 0.065 (left) and $\tau = 0.067$ (right). Treatment recommendation ("GL": Glargine; "LM": LMx75/25) are given for terminal nodes in bold face. "Efficacy", Decrease in HbA1c from baseline to end of follow-up; "Risk", Daily hypoglycemic event rate (# events / days follow-up). Internal nodes contain the mean efficacy and risk scores from the training observations at the current node (annotated in upper right corner). Terminal nodes (gray highlighted) display the node level mean efficacy and risk estimates

Since rcDT and rcRF methods appear to control risk close to the τ level in the validation sets, rcDT and rcRF models were fit to the 1498 available observations from the DURABLE trial. Figure 2 presents rcDT structures fit using risk thresholds of $\tau = 0.063$, 0.065, and 0.067 hypoglycemic events per day. Identical tree structures were identified for daily hypoglycemia event rates of 0.063 and 0.065. The first split is made on systolic blood pressure with a cut point of 130 mmHg. Patients with systolic blood pressure less than 130 mmHg and diastolic blood pressure less than 68 mmHg are recommended to LMx75/25 while those with greater than 68 mmHg receive a treatment recommendation of Glargine. Among patients with systolic blood pressure above 130 mmHg, those under 165 cm in height are recommended to be on LMx75/25, and Glargine otherwise. This model for $\tau = 0.063$ and $\tau = 0.065$ controls risk in the training data at 0.060 hypoglycemic events per day. The training set efficacy estimate was a decrease in HbA1c of 1.946 mg/dL. When the risk constraint is relaxed to control daily rates of hypoglycemia at 0.067, an additional split among patients with systolic blood pressure greater than 130 mmHg and height above 165 cm is made on noon fasting glucose. Specifically, those with noon fasting glucose values less than 179 mg/dL to LMx75/25 and those above 179 mg/dL are recommended to Glargine. This model corresponding to $\tau = 0.067$ controls risk in the training data at 0.061 hypoglycemic events per day. The training set efficacy estimate was a decrease in HbA1c of 1.943 mg/dL. Both these models mirror the pattern observed in Table 5 where risk is over controlled in a training set and likely to be well controlled at the desired level in an external validation data set.

Internal nodes of each rcDT structure contain the global efficacy and risk estimates associated with the partition. In the upper right corner of each internal node is the place of that node in the order of splitting. For example, in the decision tree corresponding to $\tau = 0.063$ and 0.065 the initial partition sends patients with systolic blood pressure greater than 130 mmHg to LMx75/25 and Glargine otherwise. This results in global efficacy and risk estimates of 1.812 mg/dL decrease in HbA1c and 0.062 hypoglycemic events per day, respectively. The second split results is on height (cut point = 165 cm) within the group with higher systolic blood pressure. This second split yields efficacy and risk estimates of 1.857 mg/dL decrease in HbA1c and 0.057 hypoglycemic events per day, respectively. This can be repeated until no partitions remain. Terminal nodes display the efficacy and risk estimates for patients in that terminal node who received the recommended treatment. For instance, those with systolic blood pressure less than 130 mmHg, diastolic blood pressure less than 68 mmHg, and receiving LMx75/25 had mean decrease in HbA1c of 2.111 mg/dL and an average of 0.084 hypoglycemic events per day. In general, LMx75/25 recommendation carries with it a greater risk of hypoglycemia, but also a greater decrease in HbA1c.

For each of the three risk constraints considered in the DURABLE analysis, an rcRF model was trained. Risk estimates for rcRF models trained using $\tau = 0.063, 0.065$, and 0.067 hypoglycemic events per day were 0.057, 0.064, and 0.068, respectively. The associated decrease in HbA1c (efficacy) for each of the risk control values was 1.87, 1.97, and 2.01 for $\tau = 0.063, 0.065$, and 0.067, respectively. Table 6 summarizes rcDT and rcRF predictions for 10 DURABLE patients. The 10

TABLE 6 DURABLE trial analysis summaries for 10 patients

ID	Original Assignment	Efficacy	Risk	$\tau = 0.063$ Prob / Pred	$\tau = 0.065$ Prob / Pred	$\tau = 0.067$ Prob / Pred
1	Glargine	4.709	0.000	0.166 / 0	0.208 / 0	0.240 / 0
2	Glargine	3.361	0.012	0.186 / 0	0.240 / 0	0.312 / 0
3	LMx75/25	1.131	0.000	0.330 / 0	0.336 / 0	0.412 / 1
4	Glargine	0.973	0.048	0.408 / 1	0.510 / 1	0.558 / 1
5	Glargine	1.596	0.036	0.506 / 0	0.472 / 0	0.466 / 0
6	Glargine	2.136	0.006	0.516 / 0	0.540 / 0	0.586 / 1
7	LMx75/25	1.081	0.000	0.566 / 1	0.592 / 1	0.558 / 1
8	Glargine	−1.102	0.094	0.676 / 1	0.664 / 1	0.666 / 1
9	LMx75/25	4.730	0.012	0.776 / 1	0.824 / 1	0.858 / 1
10	Glargine	1.495	0.018	0.784 / 1	0.802 / 1	0.844 / 1

Abbreviations: "Efficacy", Observed decrease in HbA1c from baseline; "Original Assignment", Treatment originally received; "Pred", rcDT predicted treatment from rcDT model (0 = Glargine; 1 = LMx75/25); "Prob", rcRF probability of recommendation to LMx75/25; "Risk", Observed daily hypoglycemia event rate.

observations were selected across the range of rcRF predicted probabilities. For most observations, the probability of being assigned to LMx75/25 from the rcRF model increases as the risk threshold is relaxed from 0.063 to 0.067 hypoglycemic events per day. Consider Patient 1 who was originally assigned to Glargine, had an observed decrease in HbA1c of 4.7 mg/dL, and recorded no hypoglycemic events. This patient is recommended to remain on Glargine by the rcRF model (probabilities of recommending LMx75/25 all below 0.25). All three rcDT models recommend Glargine as well. Patient 6 was also originally assigned to Glargine, but had a smaller observed decrease in HbA1c (2.1 mg/dL) compared to Patient 1, and experienced 1 hypoglycemic event (0.006 per day over 24-week follow-up). For Patient 6, the rcRF model makes a weak recommendation to switch to LMx75/25 when $\tau = 0.063$ ($P = .516$). As τ increases the probability of recommending a change of treatment to LMx75/25 rises to 0.540 ($\tau = 0.065$) and 0.586 ($\tau = 0.067$). The rcDT models recommend remaining on Glargine when $\tau = 0.063$ and 0.065 and switching to LMx75/25 when $\tau = 0.067$. Since Patient 6 experienced a benefit from Glargine and had a low rate of hypoglycemia (0.006 vs Glargine group mean of 0.057), the recommendation to switch to active treatment is weak if desired risk control is highly conservative. Patient 10 was originally assigned to Glargine, but had a much weaker response (1.5 mg/dL decrease in HbA1c) and higher daily rate of hypoglycemic events (0.018 events / day over 24-week follow-up). Hence, there is a strong recommendation of switching to LMx75/25 regardless of risk control choice with rcRF probabilities all greater than 0.78 and all three rcDT models also recommend LMx75/25. Finally, consider Patients 3 and 7. Both were originally assigned to LMx75/25, experienced similar efficacy responses (decrease in HbA1c both roughly 1.1 mg/dL), and both with no hypoglycemic events. However, Patient 3 is recommended to Glargine (except for and rcDT model with $\tau = 0.067$) and Patient 7 is recommended to stay with the original assignment of LMx75/25. Digging into the data we find that Patient 3 is taller than Patient 7 (172 vs 160 cm) with a slightly greater systolic blood pressure measurement (140 vs 132 mmHg). Additionally, Patient 3 had blood glucose measures across all seven time points throughout the day that were about 33% lower than Patient 7. From the tree structures in Figure 2 it appears height and blood glucose measurement are driving the opposing treatment decisions for Patients 3 and 7.

Variable importance measures for the rcRF models are presented in Figure 3. Importance measures (total, efficacy, and risk) for $\tau = 0.065$ daily hypoglycemic events per day are summarized using box plots with individual points overlaid each corresponding to one of the 100 rcRF models. Height, systolic blood pressure, and weight were the top three predictors for all importance measures. This indicates that they are the strongest predictors of treatment recommendation among the available covariates. This corresponds to the rcDT models presented in Figure 2 as systolic blood pressure and height appear within the first two layers of every tree, regardless of the risk constraint imposed. Notably, LMx75/25 more effectively lowers postprandial blood glucose compared to basal insulin. The tree structures for the DURABLE data indicate that patients with greater systolic blood pressure are recommended to LMx75/25. One possible explanation is that patients with greater systolic blood pressure also consume more food at each meal and so the mix may yield greater benefit to these patients. Notably, heart rate and blood glucose readings taken after a meal (either morning or evening) were consistently among the top predictors in all three proposed importance measures.

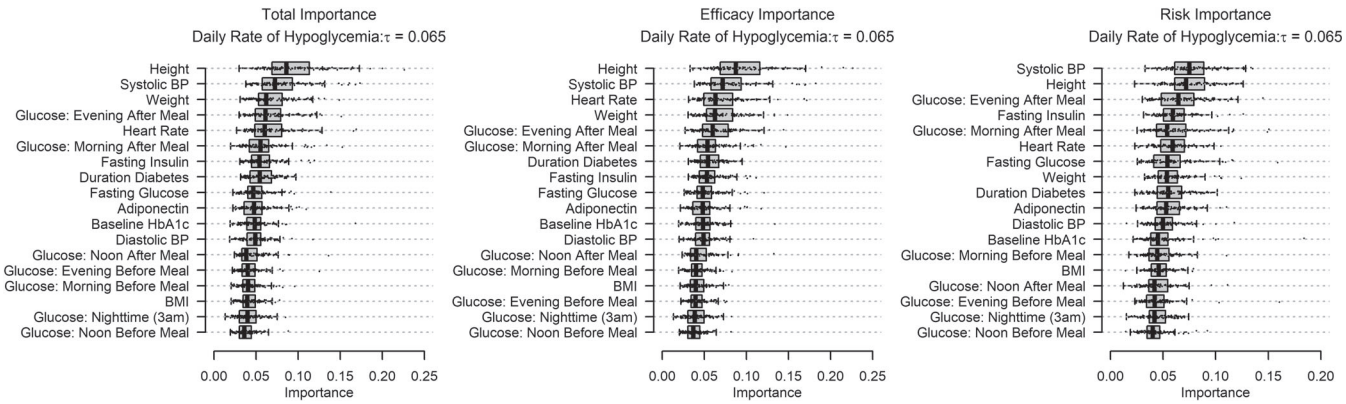

FIGURE 3 Variable importance measures for DURABLE trial with risk control level of $\tau = 0.065$ hypoglycemic events per day. Columns correspond to total, efficacy, and risk importances. Importance measure for each predictor are displayed as points with a box plot summary overlaid. Boxplots are ordered in each plot by mean variable importance

A variable importance measure was proposed by Wang et al[14] to accompany their BR-O proposal. Using the same set of predictors as this current work, their analysis identified baseline HbA1c, BMI, and fasting glucose as the three most important predictors of treatment assignment. Height and weight, which are highly correlated with BMI, were identified by rcRF modeling as highly important predictors of treatment. Participant height and weight were the fourth and sixth ranked predictors for BR-O. Postprandial glucose measures were identified by rcRF modeling, which may link to the phenomena of postprandial reactive hypoglycemia,[24,25] that is, elevated insulin levels trigger a hypoglycemic state. Given that both postprandial glucose measures had increased importance specifically related to risk prediction, this corresponds well with the current understanding of the relationship between food consumption, insulin production, and risk of hypoglycemia. Systolic blood pressure ranked 11th of 18 predictors in BR-O modeling, while it ranked in the top two predictors in rcRF modeling. This potential link between blood pressure and treatment assignment warrants further investigation.

6 | DISCUSSION

This work proposes two novel methods for discovery of risk-controlled ITRs, namely rcDT and rcRF. The goal of this class of ITR is to balance the expected efficacy and risk for a given treatment regime, aiming to control the expected risk at a predetermined, clinically meaningful threshold while achieving high efficacy. There is a pressing need for treatment rules that are easy to interpret as this facilitates communication with patients about why a particular treatment is being prescribed. To that end, the high interpretable rcDT model was developed. An rcDT model leverages a global estimator of efficacy and risk to construct a purity measure that is a composition of expected efficacy and risk scores computed under a "value"-based framework. The rcDT model, to our knowledge, is the only risk constrained optimization procedure available for ITR discovery that results in an easily interpretable rule and can be applied to both RCT and observational data. Additionally, the rcRF extension of rcDT is proposed. An rcRF model aggregates several bootstrap rcDT learners to construct an rcITR. Three variable importance measures calculated from an rcRF model were defined corresponding to the totality of the rule, efficacy, and risk. Importance measures can be directly compared for each predictor, representing a distinct advantage over other proposed importance measures in this class of ITR. Simulation studies demonstrated the robustness of the rcRF procedure (compared to rcDT, BR-M, and BR-O) to identify the optimal rcITR under a variety of efficacy and risk structures (see simulation schemes D and E).

We opt not to simply enforce the risk constraint at each split during tree construction, that is, define the splitting criterion as the constrained optimization 1. While this approach would surely control risk at the τ level, final rules may produce suboptimal efficacy estimates. As an illustration, suppose risk was to be constrained at each split and an initial partition of the data yields a set of candidate rules $\mathcal{D}$ such that $\forall d \in \mathcal{D}$, $\mathrm{E}^d(R) > \tau$. Under these conditions, no splitting would be performed as all candidate rules have expected risk above the threshold, and a null tree would be returned. However, it may be possible that given some initial partition of the data, call it d_o^1, with $\mathrm{E}^d(R) \leq \tau$, that there exists a second partition of daughter node 1, call it d_o^{11}, such that $\mathrm{E}^{d_o^{11}}(R) \leq \tau$ and $L(d_0^{11}) > L(d_o^1)$. Clearly, it would be preferred to consider splitting past the initial partition in order to capture potential complexity in the underlying association between efficacy,

risk, and the predictor space. The proposed purity measure would accommodate this complexity, whereas defining the splitting criterion based on the constrained optimization directly would not.

In the analysis of the DURABLE trial data rcRF controlled risk close to the τ level while returning the greatest estimates of efficacy compared to rcDT and BR-M procedures. The rcRF model also yielded a greater level of precision in risk and efficacy estimation than rcDT and BR-M modeling. rcDT returned efficacy estimates from validation sets that were close to those obtained using rcRF. In addition, rcDT models demonstrated the ability to control risk properly. A fitted rcDT model to the DURABLE trial data identified systolic and diastolic blood pressure and height as the defining covariates in the risk-controlled ITR. Analysis of the DURABLE trial data performed by Wang et al[14] using an outcome weighted learning approach with a linear kernel identified height, diastolic, and systolic blood pressure as the fourth, eighth, and eleventh most important predictors out of 18 available. Their procedure identified baseline HbA1c, BMI, and fasting glucose level as the most important predictors. Both rcRF and outcome weighted learning derived importance measures for the DURABLE analysis did not change meaningfully across the risk constraints considered. As elevated blood pressure is the most prevalent comorbidity condition among diabetes patients, systolic blood pressure is associated with poor patient blood glucose control. Similarly, height and BMI may be acting as surrogate measures of the same risk factor. Since rcDT and rcRF methods are able to control risk while achieving greater efficacy than linear ITR methods, this rule represents an intriguing view into diabetes management.

The methods proposed accommodate treatment efficacy and risk heterogeneity via construction of subgroups containing patients with similar baseline characteristics and subgroup specific treatment assignments. One can imagine that changes in τ may result in different treatment recommendations for some patients, indicating there may be a functional relationship between τ and the covariates. We defer further investigation of this relation to future research. The methods proposed here can easily be extended to include categorical predictors, multiple treatments, and multiple constraints. Alternative endpoints can also be accommodated such as survival outcomes (Zhao et al[26] investigated discovering ITRs for censored time-to-event data). A notable downside to the current risk controlled ITR procedures that utilize either random forests or SVMs is the computational cost associated with scaling for analyses with high dimensional data. Hence, analyzing predictors originating from the genome or proteome may not be feasible. Further work is needed to accommodate high-dimensional predictor sets into the modeling procedures. Since trial data is often collected over multiple follow-ups, the incorporation of time-varying covariates into risk-controlled modeling procedures warrants further exploration.

In conclusion, two novel methods were proposed for discovery of risk controlled ITR. The first, rcDT, is easy to interpret and performs well under a variety of circumstances. The second, rcRF, while not interpretable is robust against a wide variety of underlying rules as was demonstrated in the simulation studies. The rcRF model also allows for variable importance measures to be defined, further elucidating the contribution of predictors to the ITR. It is our recommendation that rcDT and rcRF models should be used in tandem, such that an interpretable tree structure is obtained (rcDT) along with importance measures from an rcRF model. This work contributes to a growing number of modern statistical techniques aimed at deciding proper treatment rules based on both efficacy and risk considerations.

DATA AVAILABILITY STATEMENT

Data from this study are not publicly available.

ORCID

Kevin Doubleday https://orcid.org/0000-0002-7263-2280
Jin Zhou https://orcid.org/0000-0001-7983-0274

REFERENCES

1. Su X, Tsai CL, Wang H, Nickerson DM, Li B. Subgroup analysis via recursive partitioning. *J Mach Learn Res*. 2009;10(Feb):141-158.
2. Lipkovich I, Dmitrienko A, Denne J, Enas G. Subgroup identification based on differential effect search—a recursive partitioning method for establishing response to treatment in patient subpopulations. *Stat Med*. 2011;30(21):2601-2621.
3. Cai T, Tian L, Wong PH, Wei LJ. Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics*. 2011;12(2):270-282.
4. Foster JC, Taylor JMG, Ruberg SJ. Subgroup identification from randomized clinical trial data. *Stat Med*. 2011;30(24):2867-2880.
5. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *Ann Stat*. 2011;39(2):1180.
6. Faries DE, Chen Y, Lipkovich I, Zagar A, Liu X, Obenchain RL. Local control for identifying subgroups of interest in observational research: persistence of treatment for major depressive disorder. *Int J Methods Psychiatr Res*. 2013;22(3):185-194.
7. Zhao L, Tian L, Cai T, Claggett B, Wei L-J. Effectively selecting a target population for a future comparative study. *J Am Stat Assoc*. 2013;108(502):527-539.

8. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *J Am Stat Assoc*. 2012;107(499):1106-1118.
9. Zhou X, Mayer-Hamblett N, Khan U, Kosorok MR. Residual weighted learning for estimating individualized treatment rules. *J Am Stat Assoc*. 2017;112(517):169-187.
10. Doubleday K, Zhou H, Fu H, Zhou J. An algorithm for generating individualized treatment decision trees and random forests. *J Comput Graph Stat*. 2018;27(4):849-860.
11. Lizotte DJ, Bowling M, Murphy SA. Linear fitted-q iteration with multiple reward functions. *J Mach Learn Res*. 2012;13(Nov):3253-3295.
12. Thall PF. Bayesian adaptive dose-finding based on efficacy and toxicity. *J Stat Res*. 2012;46(2):187-202.
13. Laber EB, Lizotte DJ, Ferguson B. Set-valued dynamic treatment regimes for competing outcomes. *Biometrics*. 2014;70(1):53-61.
14. Wang Y, Fu H, Zeng D. Learning optimal personalized treatment rules in consideration of benefit and risk: with an application to treating type 2 diabetes patients with insulin therapies. *J Am Stat Assoc*. 2018;113(521):1-13. doi:10.1080/01621459.2017.1303386
15. Morgan James N, Sonquist JA. Problems in the analysis of survey data, and a proposal. *J Am Stat Assoc*. 1963;58(302):415-434.
16. Breiman L, Friedman J, Olshen RA, Stone CJ. *Classification and Regression Trees*. Boca Raton, FL: CRC Press; 1984.
17. Buse JB, Wolffenbuttel BHR, Herman WH, et al. DURAbility of basal versus lispro mix 75/25 insulin efficacy (DURABLE) trial 24-week results. *Diabetes Care*. 2009;32(6):1007-1013.
18. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2017.
19. Rosenbaum Paul R, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
20. Borah BJ, Moriarty JP, Crown WH, Doshi JA. Applications of propensity score methods in observational comparative effectiveness and safety research: where have we come and where should we go? *J Comparat Effect Res*. 2014;3(1):63-78.
21. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
22. Friedman J, Hastie T, Tibshirani R. *The Elements of Statistical Learning*. Springer Series in Statistics. Berlin, Germany: Springer; 2001.
23. Fahrbach J, Jacober S, Jiang H, Martin S. The DURABLE trial study design: comparing the safety, efficacy, and durability of insulin glargine to insulin lispro mix 75/25 added to oral antihyperglycemic agents in patients with type 2 diabetes. *J Diabetes Sci Technol*. 2008;2(5):831-838.
24. Brun JF, Fédou C, Mercier J. Postprandial reactive hypoglycemia. *Diabetes Metab*. 2000;26(5):337-352.
25. Altuntaş Y. Postprandial reactive hypoglycemia. *Şişli Etfal Hastanesi tıp Bülteni*. 2019;53(3):215.
26. Zhao Y-Q, Zeng D, Laber EB, Song R, Yuan M, Kosorok MR. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika*. 2014;102(1):151-168.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Doubleday K, Zhou J, Zhou H, Fu H. Risk controlled decision trees and random forests for precision Medicine. *Statistics in Medicine*. 2022;41(4):719-735. doi: 10.1002/sim.9253