Biometrics 70, 695–720
September 2014

DOI: 10.1111/biom.12191

Combining Biomarkers to Optimize Patient Treatment Recommendations

Chaeryon Kang,* Holly Janes, and Ying Huang

Vaccine and Infectious Disease Division and Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, U.S.A.
**email:* ckang2@fhcrc.org

Summary. Markers that predict treatment effect have the potential to improve patient outcomes. For example, the *Oncotype DX*® RecurrenceScore® has some ability to predict the benefit of adjuvant chemotherapy over and above hormone therapy for the treatment of estrogen-receptor-positive breast cancer, facilitating the provision of chemotherapy to women most likely to benefit from it. Given that the score was originally developed for predicting outcome given hormone therapy alone, it is of interest to develop alternative combinations of the genes comprising the score that are optimized for treatment selection. However, most methodology for combining markers is useful when predicting outcome under a single treatment. We propose a method for combining markers for treatment selection which requires modeling the treatment effect as a function of markers. Multiple models of treatment effect are fit iteratively by upweighting or "boosting" subjects potentially misclassified according to treatment benefit at the previous stage. The boosting approach is compared to existing methods in a simulation study based on the change in expected outcome under marker-based treatment. The approach improves upon methods in some settings and has comparable performance in others. Our simulation study also provides insights as to the relative merits of the existing methods. Application of the boosting approach to the breast cancer data, using scaled versions of the original markers, produces marker combinations that may have improved performance for treatment selection.

Key words: Biomarker; Boosting; Model mis-specification; Treatment selection.

1. Introduction

Discovering and describing heterogeneity in treatment effects across patient subgroups has emerged as a key objective in clinical trials and drug development. If the treatment effect can be predicted given marker values such as biological measurements and clinical characteristics, providing patients and clinicians with these marker values can help them make more informed treatment decisions. For example, the *Oncotype DX* Recurrence Score is a leading marker for predicting the benefit of adjuvant chemotherapy over and above tamoxifen among breast cancer patients with estrogen receptor-positive (ER-positive) tumors (Albain et al., 2010a). The Recurrence Score is a proprietary combination of expression levels of 21 genes (16 cancer-related and 5 reference) measured in breast cancer tumor tissue, and is used to identify a subgroup of patients for whom the likelihood of benefitting from adjuvant chemotherapy is small. These patients can therefore avoid unnecessary and potentially toxic treatment.

There is a large literature on statistical methods for combining markers, but the vast majority of them have focused on combining markers for predicting outcome under a single treatment (e.g., Etzioni et al., 2003; Pepe, Cai, and Longton, 2005; Zhao et al., 2011). However, combinations of markers for risk prediction or classification under a single treatment are not optimized for treatment selection. Being at high risk for the outcome does not necessarily imply a larger benefit from a particular treatment (Henry and Hayes, 2006; Janes et al., 2011; Janes, Pepe, and Huang, in press). In particular, the Recurrence Score was originally developed for predicting the risk of disease recurrence or death given treatment with tamoxifen alone (Paik et al., 2004), and was later shown to have value for predicting chemotherapy benefit (Paik et al., 2004; Albain et al. 2010a, 2010b). Therefore, it is of interest to explore alternative combinations of gene expression measures that are optimized for treatment selection.

Statistical methods for combining markers for treatment selection are being developed (see Gunter, Zhu, and Murphy, 2007; Brinkley, Tsiatis, and Anstrom, 2010; Cai et al., 2011; Claggett et al., 2011; Foster, Taylor, and Ruberg, 2011; Gunter, Zhu, and Murphy, 2011a; Zhang et al., 2012; Zhao et al., 2012; Lu, Zhang, and Zeng, 2013). A simple approach uses generalized linear regression to model the expected disease outcome as a function of treatment and markers, including an interaction between each marker and treatment (Gunter et al., 2007; Cai et al., 2011; Lu et al., 2013; Janes et al., in press). This model is difficult to specify, particulary with multiple markers as in the breast cancer example, and hence an approach that is robust to model mis-specification is warranted. This is a key motivation for our approach to combining markers for treatment selection. We call our approach "boosting" since it is a natural generalization of the Adaboost (Adaptive boosting) method used to predict disease outcome under a single treatment Freund and Schapire (1997) and Friedman, Hastie, and Tibshirani (2000).

Candidate approaches for combining markers should be compared with respect to a clinically relevant performance

measure, and yet a few of the existing studies have performed such comparisons. In a simulation study and in our analysis of the breast cancer data, we evaluate methods for combining markers using the cardinal measure of model performance: the improvement in expected outcome under marker-based treatment (Song and Pepe, 2004; Brinkley et al., 2010; Gunter, Zhu, and Murphy, 2011b; Zhang et al., 2012; Janes et al., 2014). To the best of our knowledge, only two other articles (Qian and Murphy, 2011; Zhang et al., 2012) have used this approach for evaluating new methodology.

The structure of the article is as follows. In Section 2, we introduce our approach to evaluating marker combinations for treatment selection and describe the boosting method. A simulation study used to evaluate the boosting approach in comparison to other candidate approaches is described in Section 3. Section 4 describes our application of the boosting approach to the breast cancer data. We conclude with a discussion of our findings and further research topics to pursue.

2. Methods

2.1. *Context and Notation*

Let D be a binary indicator of an adverse outcome following treatment which we refer to as "disease." In the breast cancer example, D indicates death or cancer recurrence within 5 years of study enrollment. We assume that D captures all the consequences of treatment, such as subsequent toxicity, morbidity, and mortality; more general settings are addressed in Section 5. Suppose that the task is to decide, for each individual patient, between two treatment options denoted by T, where we call $T = 1$ "treatment" and $T = 0$ "no treatment." We assume that the default treatment strategy is to treat all patients. The marker, $Y \in \mathbb{R}^p$, may be useful for identifying a subgroup of patients who can avoid treatment. This setup is motivated by the breast cancer context, wherein adjuvant chemotherapy in addition to hormone therapy ($T = 1$) is the standard of care and markers are used to identify women who can forego adjuvant chemotherapy ($T = 0$). The setting where $T = 0$ is the default and Y is used to identify a subgroup to treat can be handled by simply switching treatment labels ($T = 0$ for treatment and 1 for no treatment). We assume that the data $\{D_i, T_i, Y_i\}_{i=1}^n$ come from the ideal setting for evaluating treatment efficacy, a randomized clinical trial comparing $T = 0$ to $T = 1$ where Y is measured at baseline and D is a clinical outcome observed for all subjects.

2.2. *Measures for Evaluating Marker Performance*

Let $\Delta(Y) \equiv P(D = 1|T = 0, Y) - P(D = 1|T = 1, Y)$ denote the marker-specific treatment effect. Given marker values Y for all subjects, the treatment policy that minimizes the population disease rate is to recommend no treatment if $\phi(Y) = \mathbf{1}\{\Delta(Y) \le 0\} = 1$, where $\mathbf{1}(\cdot)$ is the indicator function (Vickers, Kattan, and Sargent, 2007; Brinkley et al., 2010; Zhang et al., 2012; Lu et al., 2013). In the breast cancer example, this policy would recommend hormone therapy alone to patients with negative treatment effects and adjuvant chemotherapy to patients with positive treatment effects. The function $\Delta(Y)$ is therefore the combination of markers that we seek, and $\phi(Y)$ is the associated treatment rule. Given data $\{D_i, T_i, Y_i\}$ for $i = 1, \ldots, n$ subjects, we estimate the marker-specific treatment effect by fitting a model for $P(D = 1|T, Y)$, termed the "risk model," and calculate $\widehat{\Delta}(Y) = \widehat{P}(D = 1|T = 0, Y) - \widehat{P}(D = 1|T = 1, Y)$ and $\widehat{\phi}(Y) = \mathbf{1}\{\widehat{\Delta}(Y) \le 0\}$.

We characterize the performance of an arbitrary estimated treatment rule $\widehat{\phi}(Y)$ by evaluating the benefit of marker-based treatment (Song and Pepe, 2004; Brinkley et al., 2010; Gunter et al., 2011b; Zhang et al., 2012; Janes et al., 2014). This is measured by the difference in the disease rate under marker-based treatment assignment versus the default strategy of providing treatment to all patients:

$$\begin{aligned}\theta\{\widehat{\phi}(Y)\} \equiv\ & P(D = 1|T = 1) - [P\{D = 1|T = 0, \widehat{\phi}(Y) = 1\} \\ & \times P\{\widehat{\phi}(Y) = 1\} + P\{D = 1|T = 1, \widehat{\phi}(Y) = 0\} P\{\widehat{\phi}(Y) = 0\}] \\ =\ & [P\{D = 1|T = 1, \widehat{\phi}(Y) = 1\} \\ & - P\{D = 1|T = 0, \widehat{\phi}(Y) = 1\}] \times P\{\widehat{\phi}(Y) = 1\}.\end{aligned}$$

In the breast cancer example, θ denotes the reduction in the 5-year death or recurrence rate under marker-based treatment; in general, a higher value of θ indicates greater marker value. Using the standard empirical measure $\mathbb{P}_n(\delta) \equiv \sum_{i=1}^n n^{-1}\delta_i$, θ is estimated empirically as follows:

$$\begin{aligned}\widehat{\theta}\{\widehat{\phi}(Y)\} = \Bigg[& \frac{\mathbb{P}_n \mathbf{1}\{D = 1, T = 1, \widehat{\phi}(Y) = 1\}}{\mathbb{P}_n \mathbf{1}\{T = 1, \widehat{\phi}(Y) = 1\}} \\ & - \frac{\mathbb{P}_n \mathbf{1}\{D = 1, T = 0, \widehat{\phi}(Y) = 1\}}{\mathbb{P}_n \mathbf{1}\{T = 0, \widehat{\phi}(Y) = 1\}} \Bigg] \times \mathbb{P}_n \mathbf{1}\{\widehat{\phi}(Y) = 1\}.\end{aligned}$$

Another important measure of the population performance of the marker is the rate of incorrect treatment recommendations, which we call the misclassification rate of treatment benefit, $\mathrm{MCR}_{\mathrm{TB}}\{\widehat{\phi}(Y)\} \equiv P\{\phi(Y) \ne \widehat{\phi}(Y)\}$, and estimate by $\widehat{\mathrm{MCR}}_{\mathrm{TB}}\{\widehat{\phi}(Y)\} = \mathbb{P}_n \mathbf{1}\{\phi(Y) \ne \widehat{\phi}(Y)\}$. Similar measures have been used by Foster et al. (2011) and Lu et al. (2013). Although this measure cannot be evaluated in practice since $\Delta(Y)$ is unknown, it can be evaluated in simulated data where $\Delta(Y)$ is known.

2.3. *The Boosting Method of Combining Markers for Treatment Selection*

A simple approach for estimating $\Delta(Y)$ is to use a generalized linear model for the outcome, D, as a function of markers, Y, and treatment, T, including interactions between each marker and treatment. That is, to stipulate that

$$h\{P(D = 1|T, Y)\} = \eta(T, Y), \tag{1}$$

where the linear predictor $\eta(T, Y) = \widetilde{Y}\beta_1 + T\widetilde{Y}\beta_2$, $\widetilde{Y} = (1^T, Y^T)^T$, β_1 and β_2 are $(p+1)$-dimensional vectors of regression coefficients for the markers' main effects and interactions with treatment, respectively, and h is a link function. The logit link is the most common choice for a binary outcome. This risk model, if correctly specified, produces the combination of markers, $\Delta(Y)$, with optimal performance, that is, $\theta\{\phi(Y)\}$. However if the risk model is mis-specified, it will produce a biased estimate of treatment effect, resulting in a suboptimal combination of markers and

rule for assigning treatment. With multiple markers, the likelihood of risk model mis-specification is increased. Our method seeks to improve upon logistic regression by providing an estimate of treatment effect, and a combination of markers, that is more robust to risk model mis-specification.

To achieve this goal, we adopt the idea of Adaboost, which iteratively fits classifiers, at each stage assigning higher weights to subjects whose outcomes are misclassified at the previous stage in order to minimize classification error. Analogously, we repeatedly fit a "working model" for $P(D = 1|T, Y)$, and at each stage give more weight to subjects who lie close to the decision boundary, $\Delta(Y) = 0$, who have greater potential to be recommended the incorrect treatment. In other words, we extend Adaboost from the classification setting, where the outcome to be predicted is D, to the treatment selection setting, where the outcome is $\mathbf{1}\{\Delta(Y) \le 0\}$. The added complexity is that $\mathbf{1}\{\Delta(Y) \le 0\}$ is not directly observable. Details of the boosting algorithm are given below.

Boosting algorithm

(1) With initial weight $w_i^{(0)} = w^{(0)}(Y_i)$ for subject i, $i = 1, \ldots, n$, fit the working risk model and calculate $\widetilde{P}^{(0)}(D = 1|T = t, Y)$, $t = 0, 1$, and $\widetilde{\Delta}^{(0)}(Y) = \widetilde{P}^{(0)}(D = 1|T = 0, Y) - \widetilde{P}^{(0)}(D = 1|T = 1, Y)$.

(2) Update weights according to $w_i^{(1)} = w_i^{*(1)} / \sum_{i=1}^n w_i^{*(1)}$, where $w_i^{*(1)} = \min[\widetilde{w}\{\widetilde{\Delta}^{(0)}(Y_i)\}, C_{\mathrm{M}}]$, for $\tilde{w}(u)$ decreasing in $|u|$ and a specified maximum weight C_{M}. In our simulations, we use $\widetilde{w}\{\widetilde{\Delta}^{(0)}(Y_i)\} = |\widetilde{\Delta}^{(0)}(Y_i)|^{-\frac{1}{3}}$ and $C_{\mathrm{M}} = 500$. This upweights subjects with small $|\widetilde{\Delta}^{(0)}(Y)|$ and limits the maximum size of the weights.

(3) Re-fit the working model with updated weights $w_i^{(1)}$ to obtain $\widetilde{P}^{(1)}(D = 1|T = t, Y)$, $t = 0, 1$ and $\widetilde{\Delta}^{(1)}(Y_i) = \widetilde{P}^{(1)}(D = 1|T = 0, Y_i) - \widetilde{P}^{(1)}(D = 1|T = 1, Y_i)$ for all subjects.

(4) Repeat steps (2)–(3) until either a pre-specified convergence criterion is satisfied or a specified maximum number of iterations ($M_{\max}$) is reached. In our simulations, we set $M_{\max} = 500$ as an upper limit on the number of iterations that would be necessary.

(5) After the last iteration, denoted by $M \le M_{\max}$, we have $\{\widetilde{P}^{(1)}(D = 1|T = t, Y_i), \ldots, \widetilde{P}^{(M)}(D = 1|T = t, Y_i)\}$, $t = 0, 1$, and $\{\widetilde{\Delta}^{(1)}(Y_i), \ldots, \widetilde{\Delta}^{(M)}(Y_i)\}$ for $i = 1, \ldots, n$. The estimated disease rate and treatment effect for subject i are $\widehat{P}(D_i = 1|T_i = t, Y_i) = M^{-1} \sum_{m=1}^M \widetilde{P}^{(m)}(D_i = 1|T_i = t, Y_i)$ for $t = 0, 1$, and $\widehat{\Delta}(Y_i) = M^{-1} \sum_{m=1}^M \widetilde{\Delta}^{(m)}(Y_i)$, and the estimated treatment rule is $\widehat{\phi}(Y_i) = \mathbf{1}\{\widehat{\Delta}(Y_i) \le 0\}$.

(6) Given a new subject with covariate Y^0, say in an independent test data set, we apply the set of working risk models in (5) and calculate $\widetilde{\Delta}^{(m)}(Y^0)$ for $m = 1, \ldots, M$. The estimated treatment effect is $\widehat{\Delta}(Y^0) = M^{-1} \sum_{m=1}^M \widetilde{\Delta}^{(m)}(Y^0)$ and $\widehat{\phi}(Y^0) = \mathbf{1}\{\widehat{\Delta}(Y^0) \le 0\}$.

We explore use of the linear logistic regression model (1) and a binary classification tree (Breiman et al., 1984) as working models. However, the boosting method applies to any arbitrary model for $P(D = 1|T, Y)$. Choice of the weight function, $\widetilde{w}(u)$, maximum weight, C_{M}, and algorithm stopping rule are discussed in Web appendix A. With the logistic working model, we stop the iterations when $\|\widetilde{\beta}^{(k)} - \widetilde{\beta}^{(k-1)}\| \le 10^{-7}$, where $\widetilde{\beta}^{(k)}$ is the vector of estimated regression coefficients at the kth iteration, or when $M = M_{\max}$; and with the classification tree working model, we stop the iterations when $M = M_{\max}$.

3. Simulation Study

A simulation study was performed to compare the boosting method to existing approaches for combining markers for treatment selection. The boosting method is compared to four comparator approaches: (1) using Adaboost (Friedman et al., 2000) to combine classification trees for predicting disease outcome under each treatment separately; (2) fitting a classification tree to both treatment groups including all marker-by-treatment interactions as predictors; (3) the classic logistic regression approach which fits model (1) using maximum likelihood; and (4) the approaches of Zhang et al. (2012) that maximize the Inverse Probability Weighted (IPW) or Augmented Inverse Probability (AIPW) estimators of θ.

3.1. *Comparator Methods for Combining Markers*

3.1.1. Applying Adaboost separately to each treatment group. A natural approach is to use a risk model to combine markers to predict outcome under each treatment separately. We consider the Adaboost algorithm (Freund and Schapire, 1997; Friedman et al., 2000) which combines predictions across multiple binary classification trees (Breiman et al., 1984) ("base trees" (Hastie, Tibshirani, and Friedman, 2001)). Hereafter, this is referred to as the "Adaboost trees" method. Each base tree is built by assigning higher weights to subjects that are misclassified at the previous stage. The associated risk model for each treatment group is a function of individual markers and, potentially, interactions between markers. We use Friedman et al.'s method (Friedman et al., 2000) for estimating $P(D = 1|T, Y)$. Adaboost trees is implemented by the R function **ada** (R package ada (Culp, Johnson, and Michailidis, 2012)) using the following default settings: exponential loss function, discrete boosting algorithm, and 500 base trees. Since Adaboost trees is a non-parametric approach, the obtained combination of markers is expected to be more robust than logistic regression. However, fitting a separate classifier to each treatment group may not yield the optimal marker combination for treatment selection.

3.1.2. A single classification tree with marker-by-treatment interactions. An alternative nonparametric approach is to fit a single classification tree to both treatment groups including $\{T, TY_1, \ldots, TY_p, (1 - T)Y_1, \ldots, (1 - T)Y_p\}$ as predictors. Using this classification tree, $P(D = 1|T, Y)$ can be estimated using the empirical proportion of $D = 1$ observations in each terminal node. We use the R function **rpart** (R package rpart (Therneau, Atkinson, and Ripley, 2012)) with default settings: the minimal number of observations required to split is 20, the minimum number of observations in any terminal node is 7, and the maximal number of nodes prior to terminal node is 30. We do not prune the tree to stabilize the probability estimates (Provost and Domingos, 2003; Chu

et al., 2011), but these estimates are improved by averaging across multiple tree classifiers (Chu et al., 2011).

3.1.3. Maximizing the IPW or AIPW estimators of θ. Recently, Zhang et al. (2012) proposed an approach that finds a combination of markers by directly maximizing the mean outcome (in our context, minimizing the disease rate) under marker-based treatment. This is equivalent to maximizing $\widehat{\theta}$, the estimated decrease in disease rate under marker-based treatment. Zhang et al. (2012) consider maximizing both IPW and AIPW estimators.

Briefly, let $D(t)$ denote the potential disease outcome under treatment t. For arbitrary treatment rule $g: Y \mapsto \{0, 1\}$ (in our context, assigning no treatment), the goal is to estimate the optimal treatment rule defined by $g^{\text{opt}}(Y) = \arg\min_{g \in \mathcal{G}} E\{D(g)\} = \mathbf{1}\{\Delta(Y) \le 0\}$, where $D(g) \equiv D(1)\{1 - g(Y)\} + D(0)g(Y)$. Given a parametric working risk model $P(D = 1|T, Y; \beta)$ parameterized by finite-dimensional parameter β, let $\eta = \eta(\beta)$ denote a scaled version of β satisfying $\|\eta\| = 1$ with $\|\cdot\|$ denoting the ℓ_2-norm. Treatment rules in this class of risk models are written $g(Y, \eta)$. The scaling is used to ensure that the solution $\eta^{\text{opt}} \equiv \arg\min_{\eta} Q(\eta)$, where $Q(\eta) \equiv E[D\{g(Y, \eta)\}]$, is unique. Specifically, η^{opt} is estimated by minimizing the IPW or AIPW estimators of $Q(\eta)$ as follows:

$$\text{IPWE}(\eta) = \mathbb{P}_n \left\{ \frac{C_\eta D}{\pi_c(Y; \eta, \widehat{\gamma})} \right\}, \tag{2}$$

$$\text{AIPWE}(\eta) = \mathbb{P}_n \left\{ \frac{C_\eta D}{\pi_c(Y; \eta, \widehat{\gamma})} - \frac{C_\eta - \pi_c(Y; \eta, \widehat{\gamma})}{\pi_c(Y; \eta, \widehat{\gamma})} m(Y; \eta, \widehat{\beta}) \right\}, \tag{3}$$

where $\widetilde{Y} = (1, Y)$, $\pi(Y; \gamma) = P(T = 1|Y; \gamma) = \frac{e^{\widetilde{Y}\gamma}}{1+e^{\widetilde{Y}\gamma}}$ is a known or estimated probability of treatment (the "propensity score"), $\pi_c(Y; \eta, \widehat{\gamma}) = \pi(Y; \widehat{\gamma})^T + \{1 - \pi(Y; \widehat{\gamma})\}^{1-T}$, $C_\eta = T\{1 - g(Y, \eta)\} + (1 - T)g(Y, \eta)$ is the treatment recommend by the rule $g(Y, \eta)$, and $m(Y; \eta, \widehat{\beta}) = P(D = 1|T = 1, Y; \widehat{\beta})\{1 - g(Y, \eta)\} + P(D = 1|T = 0, Y; \widehat{\beta})g(Y, \eta)$ is the model-estimated disease rate under $g(Y, \eta)$. In our randomized trial setting, the propensity score model is known by design. The IPW estimator (2) thus reduces to the empirical disease rate under marker-based treatment. The AIPW estimator (3) is more efficient in large samples. Maximizing (2) or (3) therefore yields the marker combination with the highest IPW or AIPW $\widehat{\theta}\{\widehat{\phi}(Y)\}$ in the training data within the class of the working risk model. However, when the working model is mis-specified, this combination may perform poorly, and it is in this setting where the boosting approach may generate marker combinations with closer-to-optimal performance.

To implement the approach, we find $\widehat{\eta}^{\text{opt}}$ that minimizes IPWE(η) or AIPWE(η) under the linear logistic working model (1) where $\eta = \beta/\|\beta\|$. Under this model, $\mathbf{1}\{\Delta(Y) \le 0\}$ is equivalent to $\mathbf{1}\{\widetilde{Y}\eta \le 0\}$, and so the class of treatment rules is $\mathcal{G}_\eta = \{g(\widetilde{Y}; \eta) = \mathbf{1}\{\widetilde{Y}\eta \le 0\}, \|\eta\| = 1, \widetilde{Y} = (1, Y_1, \ldots, Y_p)\}$. Following Zhang et al. (2012), the R function **genoud** (R package rgenoud (Mebane and Sekhon, 2011)) is utilized to minimize IPWE(η) (2) or AIPWE(η) (3) using the genetic algorithm (Sekhon and Mebane, 1998).

3.2. *Simulation Set-Up*

We generate simulated data sets with 500 or 5000 observations in the following fashion. Binary treatment indicators $T \sim$ Bernoulli (0.5). In most scenarios we generate three independent continuous markers Y_1, Y_2, and Y_3 ($Y = (Y_1, Y_2, Y_3)$) each following a standard normal distribution; exceptions are noted below. The binary outcome $D \sim$ Bernoulli $\{P(D = 1|T, Y)\}$. The risk model $P(D = 1|T, Y)$ varies among the seven scenarios as shown in Table 1 and described below. Figure 1 displays the distribution of $\Delta(Y)$ for each scenario. The linear logistic regression model (1) and the classification tree including $\{T, TY, (1 - T)Y\}$ as predictors are used as working models for the boosting method.

Simulation scenarios

Scenario 1. The true risk model is linear logistic where Y_1, Y_2, and Y_3 have strong, intermediate, and weak interactions with treatment: $\text{logit}P(D = 1|T, Y) = 0.3 + 0.2Y_1 - 0.2Y_2 - 0.2Y_3 + T(-0.1 - 2Y_1 - 0.7Y_2 - 0.1Y_3)$. The marker combination obtained by fitting the linear logistic working model with maximum likelihood estimation (MLE) is expected to achieve the best performance. However, it is of interest to determine the extent to which other methods produce comparable results.

Scenario 2. The true risk model is the same as in Scenario 1, but now Y_1 has high leverage points. Specifically, a random 2% of Y_1 values are replaced with draws from a Uniform (8, 9) distribution. This scenario is used to compare the performance of the approaches that use the correct linear logistic working model in the context of high leverage observations.

Scenario 3. The true risk model is $\log\{-\log P(D = 1|T, Y)\} = -0.7 - 0.2Y_1 - 0.2Y_2 + 0.1Y_3 + T(0.1 + 2Y_1 - Y_2 - 0.3Y_3)$, where Y_1, Y_2, and Y_3 have strong, intermediate, and weak interactions with treatment. The linear predictor of the linear logistic working model is correct but the link function is incorrect. This scenario is used to compare the robustness of the boosting approach to other approaches in the context of minor working model mis-specification.

Scenario 4. The true risk model is $\log\{-\log P(D = 1|T, Y)\} = 2 - 1.5Y_1^2 - 1.5Y_2^2 + 3Y_1Y_2 + T(-0.1 - Y_1 + Y_2)$. Y_1 and Y_2 follow a Uniform $(-1.5, 1.5)$ distribution. The link function and main effects of the linear logistic working model are incorrectly specified, the latter due to omission of quadratic and marker-by-marker interaction terms, but the interaction terms are correct. This scenario is chosen for its similarity to the first scenario in Zhang

Table 1

True risk models and marker distributions for the seven simulation scenarios. The linear logistic regression model $\text{logit}P(D=1|T,Y)=\widetilde{Y}\beta_1+T\widetilde{Y}\beta_2$, *with* $\widetilde{Y}=(1,Y)$, *and the classification tree including* $\{T, TY, (1-T)Y\}$ *as predictors are evaluated as working models in all scenarios.*

Scenario		True risk model	Marker distribution
The linear logistic working model is correctly specified	1	$\text{logit}P(D=1\|T,Y)=0.3+0.2Y_1-0.2Y_2-0.2Y_3$ $+T(-0.1-2Y_1-0.7Y_2-0.1Y_3)$	Y_1, Y_2, and Y_3 are independent $N(0,1)$
	2	$\text{logit}P(D=1\|T,Y)=0.3+0.2Y_1-0.2Y_2-0.2Y_3$ $+T(-0.1-2Y_1-0.7Y_2-0.1Y_3)$	Same as Scenario 1 except for 2% of high leverage observations where $Y_1\sim$ Uniform $(8,9)$
Link function in the linear logistic working model is incorrectly specified	3	$\log\{-\log P(D=1\|T,Y)\}=-0.7-0.2Y_1-0.2Y_2+0.1Y_3$ $+T(0.1+2Y_1-Y_2-0.3Y_3)$	Same as Scenario 1
Link function and main effects in the linear logistic working model are incorrectly specified	4	$\log\{-\log P(D=1\|T,Y)\}=2-1.5Y_1^2-1.5Y_2^2+3Y_1Y_2$ $+T(-0.1-Y_1+Y_2)$	Y_1, and Y_2 are independent Uniform $(-1.5,1.5)$
Main effects and interactions are incorrectly specified in the linear logistic working model	5	$\text{logit}P(D=1\|T,Y)=-0.1-0.2Y_1+0.2Y_2-0.1Y_3+Y_1^2$ $+T(-0.5-2Y_1-Y_2-0.1Y_3+2Y_1^2)$	Same as Scenario 1
	6	$\text{logit}P(D=1\|T,Y)=0.1-0.2Y_1+0.2Y_2-Y_1Y_2$ $+T(-0.5-Y_1+Y_2+3Y_1Y_2)$	
Linear logistic working model is mis-specified for outlying observations	7	$P(D=1\|T,Y)=\mathbf{1}\{Y_1<8\}\frac{1}{1+e^{-\eta}}+\mathbf{1}\{Y_1\geq 8\}\left(1-\frac{1}{1+e^{-\eta}}\right)$, where $\eta=0.3+0.2Y_1-0.2Y_2-0.2Y_3$ $+T(-0.1-2Y_1-0.7Y_2-0.1Y_2)$	Same as Scenario 2

et al. (2012) who found that, in a continuous outcome setting, maximizing the IPW or AIPW estimators of θ yielded substantial improvement over standard linear regression.

Scenario 5. The true risk model is $\text{logit}P(D=1|T,Y)=-0.1-0.2Y_1+0.2Y_2-0.1Y_3+Y_1^2+T(-0.5-2Y_1-Y_2-0.1Y_3+2Y_1^2)$ including a non-linear main effect and interaction of Y_1 with treatment. The linear logistic working model mis-specifies these Y_1 effects, but the classification tree working model should be able to detect them.

Scenario 6. The true risk model is a logistic regression model including an interaction between Y_1 and Y_2 where Y_1, Y_2 and Y_1Y_2 have intermediate, intermediate, and strong interactions with treatment: $\text{logit}P(D=1|T,Y)=0.1-0.2Y_1+0.2Y_2-Y_1Y_2+T(-0.5-Y_1+Y_2+3Y_1Y_2)$. The linear logistic working model does not include Y_1Y_2 and TY_1Y_2 interaction terms whereas a classification tree working model does allow for them.

Scenario 7. The true risk model is the same linear logistic model as in Scenario 1 except for the presence of 2% outlying observations. Specifically, for a random 2% sample, Y_1 is replaced with a draw from a Uniform $(8,9)$ distribution and D is replaced with $1-D$.

For each scenario, 1000 data sets are generated and used as training data to build a prediction model and treatment assignment rule, $\widehat{\phi}(Y)=\mathbf{1}\{\widehat{\Delta}(Y)\leq 0\}$. To avoid overoptimism associated with fitting and evaluating the risk model using the same data, a single large independent test data set with $n=10^5$ observations is generated and used to evaluate the performance of the fitted treatment rule, $\theta\{\widehat{\phi}(Y)\}$. Mean and Monte-carlo standard deviation (SD) of $\theta\{\widehat{\phi}(Y)\}$ and mean $\text{MCR}_{\text{TB}}\{\widehat{\phi}(Y)\}$ are reported. The performance of the true treatment rule, $\theta\{\phi(Y)\}$, is calculated as an average of $\widehat{\theta}\{\phi(Y)\}$ over 100 Monte-carlo simulations where each $\widehat{\theta}\{\phi(Y)\}$ is obtained using $n=3\times 10^7$ observations.

3.3. *Results of the Simulation Study*

Tables 2 and 3 summarize the simulation results for sample sizes $n=500$ and $n=5000$, respectively. The performances of marker combinations obtained using the following methods are compared: Logistic regression with maximum likelihood estimation (hereafter "linear logistic MLE"), the boosting method described in Section 2.3 with linear logistic working model ("linear logistic boosting"), maximizing the IPW or AIPW estimators of θ as proposed by Zhang et al. (2012) ("maximizing IPWE or AIPWE of θ"), a single classification

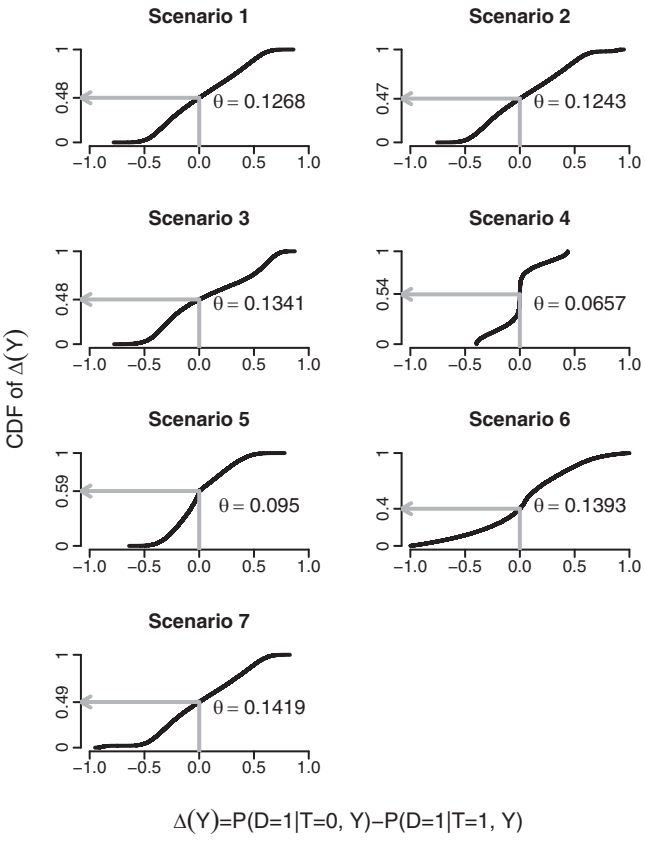

Figure 1. Distribution of the marker-specific treatment effect, $\Delta(Y) = P(D = 1|T = 0, Y) - P(D = 1|T = 1, Y)$, for each of the seven simulation scenarios. The proportion of individuals with negative treatment effects is indicated on the Y-axis, and $\theta = [P\{D = 1|T = 1, \phi(Y) = 1\} - P\{D = 1|T = 0, \phi(Y) = 1\}] \times P\{\phi(Y) = 1\}$, measuring the impact of marker-based treatment assignment, is shown.

tree with marker-by-treatment interactions ("single classification tree"), the boosting method with a classification tree working model including marker-by-treatment interactions ("classification tree boosting"), and applying Adaboost trees to each treatment group separately ("separate Adaboost"). For each scenario, the method with the highest mean θ is marked in **bold**.

When the linear logistic working model was correctly specified (Scenario 1), as expected the combination of markers obtained using linear logistic MLE had the highest mean θ, smallest SD of θ, and smallest MCR_{TB}. Linear logistic boosting produced almost identical results, whereas all other methods produced modestly lower mean θ and substantially higher SD of θ and MCR_{TB}.

In the presence of high leverage points (Scenario 2), linear logistic MLE continued to produce the highest mean θ and smallest MCR_{TB}. However, the SD of θ was slightly lower with linear logistic boosting and substantially lower when maximizing the AIPWE of θ, or employing classification tree boosting, even while the associated mean θ's were close to optimal. This suggests that, as expected, linear logistic MLE yields variable estimates in the presence of high leverage points; this effect disappears with large n (Table 3). Another observation is that only linear logistic boosting produced MCR_{TB} near that of linear logistic MLE; all other methods produced substantially higher classification error.

Mis-specifying the link function of the logistic working model (Scenario 3) had minimal impact on θ and both linear logistic MLE and linear logistic boosting produced nearly optimal mean θ and similarly low SD of θ and MCR_{TB}. All other methods yielded slightly lower mean θ and substantially higher SD of θ and MCR_{TB}. The superiority of the linear logistic regression methods persisted with larger n (Table 3). When both the link function and main effects were mis-specified (Scenario 4), methods with linear logistic working models produced similar mean θ (close to the optimal value) but linear logistic boosting had some advantage in terms of lower SD of θ and MCR_{TB}. Differences among methods were smaller again for larger n (Table 3).

Scenarios 5 and 6 explore substantial mis-specification of the linear logistic working model; the mean θ for linear logistic MLE is far from the optimal value. In these scenarios, boosting improved upon linear logistic MLE. Classification tree boosting yielded the best performance with the most dramatic improvement over logistic regression in the highly nonlinear setting of Scenario 6. These results persisted for large n (Table 3).

When the risk model mis-specification was due to outlying observations (Scenario 7), maximizing the AIPWE of θ and boosting provided marker combinations with improved performance over those generated by linear logistic MLE.

In summary, these simulation results demonstrate that the boosting method can improve upon existing methods for combining markers in certain settings. Under a substantially mis-specified working model, boosting can dramatically improve model performance. When the working model is mis-specified but not far from the true risk model, boosting may slightly improve performance. When high leverage points exist, boosting reduces variability without compromising mean performance. Boosting can perform better than direct maximization of the IPWE and AIPWE of θ, under mild or substantial working model mis-specification. As expected, linear logistic boosting performs best with minor mis-specification of the logistic risk function while classification tree boosting better captures nonlinear main effects and interactions with treatment.

4. Breast Cancer Data

The boosting method was then applied to the breast cancer data. The performance of the *Oncotype DX* Recurrence Score was most recently evaluated in the Southwest Oncology Group (SWOG)-SS8814 trial (Albain et al., 2010a), which randomized women with node-positive, ER-positive breast cancer to tamoxifen plus adjuvant chemotherapy (cyclophosphamide, doxorubicin, and fluorouracil before or concurrent with tamoxifen) or tamoxifen alone. For 367 women (219 on tamoxifen plus adjuvant chemotherapy sequentially ($T = 1$) and 148 on tamoxifen alone ($T = 0$)), expression levels of 16 breast cancer-related and 5 reference genes were measured on tumor samples obtained at surgery (before adjuvant chemotherapy), and the Recurrence Score was calculated.

Table 2

*Results of the simulation study with sample size $n = 500$. Marker combinations obtained using the following methods are compared: Logistic regression with maximum likelihood estimation (MLE), the boosting method of Section 2.3 with linear logistic working model, maximizing the IPW and AIPW estimators of θ as described by Zhang et al. (2012), a single classification tree with interactions between markers and treatment, the boosting method with classification tree working model, and applying Adaboost trees to each treatment group separately. Mean and Monte Carlo standard deviation (SD) of θ are shown, along with the mean misclassification rate for treatment benefit (MCR_{TB}). For each scenario, the method with the highest mean θ is marked in **bold**.*

Scenario		True θ		Working model	Linear logistic				Classification tree with interactions		
				Fitting algorithm	MLE	Boosting	Max $\hat{\theta}$ (IPWE)	Max $\hat{\theta}$ (AIPWE)	Single tree	Boosting	Separate Adaboost
The linear logistic working model is correctly specified	1	0.1268	θ	Mean	**0.1199**	0.1195	0.1076	0.1134	0.0971	0.1083	0.0854
				SD	0.00218	0.00258	0.01273	0.00728	0.01206	0.00649	0.00907
			MCR_{TB}	Mean	0.0517	0.0555	0.1261	0.0996	0.2351	0.1294	0.2111
	2	0.1243	θ	Mean	**0.1232**	0.1229	0.1099	0.1158	0.1004	0.1104	0.0876
				SD	0.01306	0.01156	0.01292	0.00760	0.01306	0.00740	0.00852
			MCR_{TB}	Mean	0.0512	0.0526	0.1240	0.0973	0.2372	0.1269	0.2053
Link function in the linear logistic working model is incorrectly specified	3	0.1341	θ	Mean	**0.1302**	0.1299	0.1176	0.1234	0.1056	0.1162	0.1038
				SD	0.00204	0.00224	0.01245	0.00712	0.01110	0.00654	0.00667
			MCR_{TB}	Mean	0.0418	0.0444	0.1045	0.0824	0.2072	0.1124	0.1539
Link function and main effects in the linear logistic working model are incorrectly specified	4	0.0657	θ	Mean	0.0574	**0.0607**	0.0561	0.0567	0.0221	0.0378	0.0352
				SD	0.01267	0.00986	0.01296	0.01482	0.01143	0.00867	0.00718
			MCR_{TB}	Mean	0.1511	0.1206	0.1719	0.1653	0.5397	0.3251	0.5304
Main effects and interactions are incorrectly specified in the linear logistic working model	5	0.0950	θ	Mean	0.0681	0.0702	0.0668	0.0694	0.0615	**0.0735**	0.0590
				SD	0.00703	0.00737	0.01303	0.01098	0.01359	0.00813	0.00854
			MCR_{TB}	Mean	0.2667	0.2540	0.2588	0.2503	0.2737	0.1838	0.2478
	6	0.1393	θ	Mean	0.0236	0.0438	0.0498	0.0544	0.0978	**0.1186**	0.1010
				SD	0.01875	0.01276	0.01238	0.00893	0.01996	0.01057	0.00807
			MCR_{TB}	Mean	0.3865	0.3542	0.3452	0.3330	0.2697	0.1762	0.2433
Linear logistic working model is mis-specified for outlying observations	7	0.1419	θ	Mean	0.0879	0.1140	0.1099	**0.1153**	0.1042	0.1151	0.0856
				SD	0.02370	0.01183	0.01294	0.00816	0.01657	0.00944	0.00944
			MCR_{TB}	Mean	0.2163	0.1207	0.1436	0.1198	0.2399	0.1394	0.2339

Table 3

*Results of the simulation study with sample size $n = 5000$. Marker combinations obtained using the following methods are compared: Logistic regression with maximum likelihood estimation (MLE), the boosting method of Section 2.3 with linear logistic working model, maximizing the IPW and AIPW estimators of θ as described by Zhang et al. (2012), a single classification tree with interactions between markers and treatment, the boosting method with classification tree working model, and applying Adaboost trees to each treatment group separately. Mean and Monte Carlo standard deviation (SD) of θ are shown, along with the mean misclassification rate for treatment benefit (MCR_{TB}). For each scenario, the method with the highest mean θ is marked in **bold**.*

				Working model	Linear logistic				Classification tree with interactions		
Scenario		True θ		Fitting algorithm	MLE	Boosting	Max $\hat{\theta}$ (IPWE)	Max $\hat{\theta}$ (AIPWE)	Single tree	Boosting	Separate Adaboost
The linear logistic working model is correctly specified	1	0.1268	θ	Mean	**0.1258**	0.1257	0.1198	0.1206	0.1114	0.1143	0.0990
				SD	0.00038	0.00043	0.00216	0.00149	0.00523	0.00434	0.00237
			MCR_{TB}	Mean	0.0165	0.0182	0.0527	0.0443	0.3222	0.2099	0.1664
	2	0.1243	θ	Mean	**0.1252**	0.1252	0.1227	0.1236	0.1140	0.1165	0.1003
				SD	0.00038	0.00043	0.00239	0.00170	0.00513	0.00473	0.00245
			MCR_{TB}	Mean	0.0160	0.0173	0.0530	0.0430	0.3182	0.2251	0.1635
Link function in the linear logistic working model is incorrectly specified	3	0.1341	θ	Mean	0.1316	**0.1319**	0.1299	0.1308	0.1130	0.1215	0.1120
				SD	0.00046	0.00046	0.00228	0.00151	0.00644	0.00328	0.00194
			MCR_{TB}	Mean	0.0222	0.0188	0.0436	0.0355	0.2092	0.1148	0.1290
Link function and main effects in the linear logistic working model are incorrectly specified	4	0.0657	θ	Mean	**0.0640**	**0.0640**	0.0625	0.0636	0.0259	0.0549	0.0436
				SD	0.00041	0.00036	0.00218	0.00077	0.01029	0.00412	0.00270
			MCR_{TB}	Mean	0.0633	0.0544	0.1252	0.1034	0.4949	0.2041	0.2694
Main effects and interactions are incorrectly specified in the linear logistic working model	5	0.0950	θ	Mean	0.0772	0.0804	0.0787	0.0797	0.0766	**0.0839**	0.0703
				SD	0.00210	0.00238	0.00280	0.00222	0.00458	0.00617	0.00231
			MCR_{TB}	Mean	0.2485	0.2326	0.2008	0.1941	0.3287	0.1357	0.1829
	6	0.1393	θ	Mean	0.0218	0.0496	0.0570	0.0582	0.1143	**0.1290**	0.1200
				SD	0.00834	0.00485	0.00403	0.00345	0.01494	0.00846	0.00247
			MCR_{TB}	Mean	0.3851	0.3490	0.3460	0.3472	0.2212	0.1231	0.1804
Linear logistic working model is mis-specified for outlying observations	7	0.1419	θ	Mean	0.1019	0.1215	0.1226	0.1236	0.1315	**0.1355**	0.1179
				SD	0.00824	0.00167	0.00247	0.00172	0.00537	0.00436	0.00246
			MCR_{TB}	Mean	0.1684	0.0668	0.0742	0.0635	0.3285	0.1459	0.1632

We use the SS8814 data to explore alternative combinations of the 16 breast cancer related genes that are optimized for treatment selection. In these data, there were 80 deaths or breast cancer recurrences by 5 years (35 given $T = 0$ and 45 given $T = 1$). There was little censoring; for nine subjects censored before 5 years, we assume $D = 0$. Because the data are not currently available for public use, we modified the gene values but preserved the basic underlying structure of the data. Specifically, we use scaled versions of the markers (mean centered with unit variance) and un-labeled genes. A modified version of the original Recurrence Score was used. Combinations of the following marker sets were considered for their potential to guide treatment decisions: (1) The modified risk score (MRS); (2) three genes, G_1, G_2, and G_3, that showed evidence of marker-by-treatment interactions in a multivariate linear logistic regression model; and (3) two genes, G_4 and G_5, that exhibited a significant three-way interaction TG_4G_5 in a linear logistic regression model.

We implement the following approaches: Linear logistic MLE, linear logistic boosting, maximization of the IPWE or AIPWE of θ described by Zhang et al. (2012), a single classification tree with marker-by-treatment interactions, and classification tree boosting. The tuning parameters $M_{\max}$ and $\widetilde{w}\{\Delta(Y)\}$ varied across marker sets and were determined using cross-validation (see Web Appendix A); $C_{\max}$ was set to 500 (see Web Appendix A). To assess model performance, we calculate the apparent performance $(\widehat{\theta}\{\widehat{\phi}(Y)\})$ using the original (training) data and use the percentile bootstrap to calculate a 95% confidence interval. A bootstrap-bias-corrected estimate of model performance $(\widehat{\theta}_c\{\widehat{\phi}_b(Y)\})$ (Efron and Tibshirani, 1993) is also calculated along with a 95% confidence interval obtained using the double-bootstrap (see Web Appendix B).

Performance measures of the various marker combinations are shown in Table 4. For every set of markers, maximizing the AIPWE of θ, linear logistic boosting, a single classification tree, or classification tree boosting yields a combination of markers with better performance (higher $\widehat{\theta}$) than that obtained using linear logistic MLE. For example, for the models including G_4, G_5, and G_4G_5, classification tree boosting yields a marker combination associated with a 9% decrease in 5-year recurrence or death (95% CI: 8–18%) and the marker combination maximizing the AIPWE of θ yields a 3% decrease (95% CI: 2–11%). In contrast, the combination derived using linear logistic MLE yields a 0.3% decrease (95% CI: −2 to 8%). These new combinations of markers may have improved ability to identify a subgroup of women who can avoid adjuvant chemotherapy, in terms of providing a lower population rate of 5-year death or recurrence. For example, the best function of the MRS is estimated to yield a 5% reduction in 5-year death or recurrence (95% CI: 4–13%), while allowing 64% women to avoid adjuvant chemotherapy.

Observe in Table 5 that the differences in performance between models are due to a large proportion of subjects being differently classified according to treatment benefit using linear logistic MLE versus the other approaches. The results also suggest that the linear logistic model may not hold for the modified risk score since maximizing the AIPWE of θ and classification tree boosting produce substantially higher $\widehat{\theta}$ than linear logistic MLE.

These results must be interpreted with caution, however, since even our bootstrap-bias-corrected estimates of model performance may be overoptimistic. With the small sample size, cross-validation did not produce satisfactory estimates of test data performance; results were highly dependent on the random seed used to split the data. Other bias correction approaches such as the 0.632 bootstrap method (Efron and Tibshirani, 1993) do not appear to apply to the measure θ. Obtaining sufficiently large data sets to validate marker combinations is a pervasive challenge for the treatment selection field.

5. Discussion

This article describes a novel application of boosting to combining markers for predicting treatment effect. The approach is intended to build in robustness to risk model misspecification, by averaging across risk models fit by iteratively upweighting subjects potentially misclassified according to treatment benefit at the previous stage. We evaluate the performance of the approach using clinically relevant measures and find several settings in which the boosting method results in combinations of markers that have closer-to-optimal performance than combinations derived using less-robust existing approaches. Specifically, boosting appears advantageous under substantial risk model mis-specification and in settings with high leverage points. Our analysis of the breast cancer data suggests that, in these data, boosting can yield new marker combinations that may have superior ability to identify women who do not benefit from adjuvant chemotherapy.

A simple approach to combining markers for treatment selection is to apply one of the plethora of methods available for combining markers for classification separately to each treatment group. As discussed by Claggett et al. (2011), however, the two best performing risk models for each treatment group do not necessarily produce the best model for treatment effect. This strategy risks missing markers that are strongly associated with treatment effect but which have modest main effects, and risks including markers which have strong main effects but modest interactions with treatment. For example, human epidermal growth factor receptor 2 (HER-2) is not considered a significant predictor of cancer recurrence in breast cancer patients while it is an important predictor of the effects of some adjuvant chemotherapies and hormone therapies (Clark, 1995; Henry and Hayes, 2006). In our simulations, fitting a risk model to each treatment group separately tended to produce marker combinations with inferior performance compared to those that simultaneously considered both treatment groups, such as the novel boosting method.

When evaluating candidate approaches for combining markers, it is important that methods be compared with respect to compelling and clinically relevant measures of model performance. Measures such as the frequency of correct variables selected (Gunter et al., 2007; Lu et al., 2013), the area under the receiver operating characteristic curve (AUC) for each treatment group (Claggett et al., 2011) and the mean squared error (MSE) of model coefficients (Lu et al., 2013) suffer from lack of clinical interpretation and do not characterize the benefit of the marker combination. The rate of incorrect treatment recommendation, $\mathrm{MCR}_{\mathrm{TB}}$, is appealing and useful

Table 4

Performance of marker combinations in the breast cancer data. Models including the modified risk score (MRS); genes G_1, G_2 and G_3; and genes G_4, G_5 and $G_4 \times G_5$ are shown. The following methods are used to obtain marker combinations: Logistic regression with maximum likelihood estimation (MLE), the boosting method with linear logistic working model, maximizing the IPW or AIPW estimator of θ as described by Zhang et al. (2012), a single classification tree with interactions between markers and treatment, and the boosting method with classification tree working model. For each method the apparent model performance ($\widehat{\theta}\{\widehat{\phi}(Y)\}$) and corresponding bootstrap-based 95% confidence interval; and bootstrap-bias-corrected estimate of model performance ($\widehat{\theta}_c\{\widehat{\phi}_b(Y)\}$) and corresponding bootstrap-based 95% confidence interval, are shown.

Marker set (Y)	Working model	Linear logistic				Classification tree with interactions	
	Fitting algorithm	MLE	Boosting	Max $\hat{\theta}$ (IPWE)	Max $\hat{\theta}$ (AIPWE)	Single tree	Boosting
MRS	$\widehat{\theta}\{\widehat{\phi}(Y)\}$	0.026	0.034	−0.012	0.044	0.052	0.074
	95% CI	(−0.017 , 0.068)	(−0.015 , 0.072)	(−0.062 , 0.073)	(0.005 , 0.092)	(−0.011 , 0.153)	(0.027 , 0.159)
	$\widehat{\theta}_c\{\widehat{\phi}_b(Y)\}$	0.002	0.008	−0.009	0.033	0.040	0.051
	95% CI	(−0.010 , 0.060)	(−0.014 , 0.062)	(−0.042 , 0.072)	(−0.001 , 0.087)	(0.025 , 0.127)	(0.037 , 0.133)
(G_1, G_2, G_3)	$\widehat{\theta}\{\widehat{\phi}(Y)\}$	0.052	0.060	0.063	0.080	0.075	0.095
	95% CI	(−0.012 , 0.110)	(0.013 , 0.121)	(−0.034 , 0.115)	(0.053 , 0.146)	(0.031 , 0.169)	(0.074, 0.206)
	$\widehat{\theta}_c\{\widehat{\phi}_b(Y)\}$	0.037	0.048	0.028	0.063	0.054	0.081
	95% CI	(−0.009 , 0.102)	(0.000 , 0.117)	(−0.010 , 0.086)	(0.000 , 0.137)	(0.000 , 0.132)	(0.000 , 0.159)
$(G_4, G_5, G_4 \times G_5)$	$\widehat{\theta}\{\widehat{\phi}(Y)\}$	0.011	0.049	0.050	0.057	0.053	0.125
	95% CI	(−0.027 , 0.095)	(−0.013 , 0.102)	(−0.013 , 0.094)	(0.028 , 0.126)	(0.025 , 0.159)	(0.087 , 0.240)
	$\widehat{\theta}_c\{\widehat{\phi}_b(Y)\}$	0.003	0.015	0.022	0.034	0.042	0.089
	95% CI	(−0.016 , 0.075)	(−0.013 , 0.087)	(−0.018 , 0.076)	(0.018 , 0.110)	(0.032 , 0.124)	(0.080 , 0.175)

Table 5

Estimated treatment rules in the breast cancer data obtained using the following methods: Logistic regression with maximum likelihood estimation (MLE), the boosting method with linear logistic working model, maximizing the IPW or AIPW estimator of θ as described by Zhang et al. (2012), a single classification tree with interactions between markers and treatment, and the boosting method with classification tree working model. Cells show the number (%) of subjects in the data set cross-classified by treatment rules from linear logistic MLE and other marker combination approaches. Subjects with $\widehat{\Delta}(Y) \leq 0$ are recommended tamoxifen alone $(T = 0)$ and those satisfying $\widehat{\Delta}(Y) > 0$ are recommended adjuvant chemotherapy $(T = 1)$.

		Linear logistic						Classification tree with interactions				
		Boosting		Max $\hat{\theta}$ (IPWE)		Max $\hat{\theta}$ (AIPWE)		Single tree		Boosting		
Marker set (Y)	Linear logistic MLE	$\widehat{\Delta}(Y) \leq 0$	$\widehat{\Delta}(Y) > 0$	$\widehat{\Delta}(Y) \leq 0$	$\widehat{\Delta}(Y) > 0$	$\widehat{\Delta}(Y) \leq 0$	$\widehat{\Delta}(Y) > 0$	$\widehat{\Delta}(Y) \leq 0$	$\widehat{\Delta}(Y) > 0$	$\widehat{\Delta}(Y) \leq 0$	$\widehat{\Delta}(Y) > 0$	
MRS	$\widehat{\Delta}(Y) \leq 0$	169 (46)	0 (0)	169 (46)	0 (0)	169 (46)	0 (0)	169 (46)	0 (0)	169 (46)	0 (0)	169 (46)
	$\widehat{\Delta}(Y) > 0$	9 (2)	189 (51)	155 (42)	43 (12)	17 (5)	181 (49)	107 (29)	91 (25)	67 (18)	131 (36)	198 (54)
		178 (49)	189 (51)	324 (88)	43 (12)	186 (51)	181 (49)	276 (75)	91 (25)	236 (64)	131 (36)	367 (100)
(G_1, G_2, G_3)	$\widehat{\Delta}(Y) \leq 0$	171 (47)	13 (4)	109 (30)	75 (20)	73 (36)	111 (14)	160 (44)	24 (7)	138 (38)	46 (13)	184 (50)
	$\widehat{\Delta}(Y) > 0$	12 (3)	171 (47)	19 (4)	164 (46)	13 (2)	170 (48)	103 (28)	80 (22)	79 (22)	104 (28)	183 (50)
		183 (50)	184 (50)	128 (35)	239 (65)	86 (23)	281 (77)	263 (72)	104 (28)	217 (59)	150 (41)	367 (100)
$(G_4, G_5, G_4 \times G_5)$	$\widehat{\Delta}(Y) \leq 0$	88 (24)	26 (7)	50 (14)	64 (17)	98 (27)	16 (4)	98 (27)	16 (4)	42 (11)	72 (20)	114 (31)
	$\widehat{\Delta}(Y) > 0$	69 (19)	184 (50)	68 (19)	185 (50)	75 (20)	178 (49)	173 (47)	80 (22)	77 (21)	176 (48)	253 (69)
		157 (43)	210 (57)	118 (32)	249 (68)	173 (47)	194 (53)	271 (74)	96 (26)	119 (32)	248 (68)	367 (100)

for simulation studies evaluating new methods. The decrease in the disease rate under marker-based treatment, measured by θ, has clear relevance. This measure, or a variation on it, has been advocated in several recent articles on evaluating treatment selection markers (Song and Pepe, 2004; Gunter et al., 2007, 2011b; Brinkley et al., 2010; Qian and Murphy, 2011; Janes et al., 2011, 2014; Zhang et al., 2012). θ is comprised of the proportion of subjects who are marker-negative and the treatment effect in the marker-negative subgroup. While these constituents inform about the nature of markers' effect, neither can serve as the sole basis for comparing combinations of markers.

The relative performance of the different approaches to combining markers for treatment selection depends on the scale of the outcome. While many of the methods to-date have focused on the continuous outcome setting, this article compares approaches given a binary outcome. In particular, we present results on the IPWE or AIPWE of θ maximization approach compared to logistic regression MLE, whereas the original article (Zhang et al., 2012) focused on a continuous outcome and linear regression. In our simulation study, improving upon logistic regression proved difficult. Even under risk model mis-specification, maximizing the IPWE or AIPWE of θ only resulted in moderately higher mean θ in most scenarios. In Scenario 4, constructed to be similar to the first simulation scenario of Zhang et al. (2012), maximizing the IPWE or AIPWE of θ did not yield a marker combination with superior performance to that associated with logistic regression MLE. Based on these results, it appears more difficult to improve upon logistic regression for binary outcomes than it is to improve upon linear regression for continuous outcomes. Pepe et al. (2005) also found logistic regression to be remarkably robust in the classification context.

The boosting method described here warrants further research along several avenues. The method can be generalized naturally to settings where the outcome does not capture all consequences of treatment and therefore the optimal treatment rule is $\Delta(Y) \leq \delta$ for some $\delta > 0$ (Vickers et al., 2007; Janes et al., 2014). Continuous outcomes and time-to-event outcomes could be also accommodated. Further investigation of the optimal weight function for the boosting method is of interest. The method could be extended to settings with marker values missing at random, multiple treatment options, or to the observational study setting. Another challenge is doing variable selection in the treatment selection context. Application of boosting with a penalized regression working model is one potential approach that would accommodate high dimensional makers.

6. Supplementary Materials

Web Appendix A, referenced in Sections 2.3 and 4, and Web Appendix B, referenced in Section 4, and R code to perform the estimation are available with this paper at the *Biometrics* website on Wiley Online Library

Acknowledgements

This work was funded by R01 CA152089, P30CA015704, and R01 GM106177-01.

References

Albain, K., Barlow, W., Ravdin, P., Farrar, W., Burton, G., Ketchel, S., Cobau, C., Levine, E., Ingle, J., Pritchard, K., Lichter, A., Schneider, D., Abeloff, M., Henderson, I., Muss, H., Green, S., Lew, D., Livingston, R., Martino, S., and Osborne, C. (2010a). Adjuvant chemotherapy and timing of tamoxifen in postmenopausal patients with endocrine-responsive, node-positive breast cancer: A phase 3, open-label, randomised controlled trial. *The Lancet* **374**, 2055–2063.

Albain, K., Barlow, W., Shak, S., Hortobagyi, G., Livingston, R., Yeh, I., Ravdin, P., Bugarini, R., Baehner, F., Davidson, N., Sledge, G., Winer, E., Hudis, C., Ingle, J., Perez, E., Pritchard, K., Shepherd, L., Gralow, J., Yoshizawa, C., Allred, D., Osborne, C., and Hayes, D. (2010b). Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: A retrospective analysis of a randomised trial. *The Lancet Oncology* **11**, 55–65.

Breiman, L., Friedman, J., Stone, C., and Olshen, R. (1984). *Classification and Regression Trees*. Wadsworth, Belmont, CA: Chapman & Hall/CRC.

Brinkley, J., Tsiatis, A., and Anstrom, K. (2010). A generalized estimator of the attributable benefit of an optimal treatment regime. *Biometrics* **66**, 512–522.

Cai, T., Tian, L., Wong, P., and Wei, L. (2011). Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics* **12**, 270–282.

Chu, N., Ma, L., Liu, P., Hu, Y., and Zhou, M. (2011). A comparative analysis of methods for probability estimation tree. *WSEAS Transactions on Computers* **10**, 71–80.

Claggett, B., Zhao, L., Tian, L., Castagno, D., and Wei, L. (2011). Estimating subject-specific treatment differences for risk-benefit assessment with competing risk event-time data. *Harvard University Biostatistics Working Paper Series*, 125.

Clark, G. (1995). Prognostic and predictive factors for breast cancer. *Breast Cancer* **2**, 79–89.

Culp, M., Johnson, K., and Michailidis, G. (2012). *ada: An R Package for Stochastic Boosting*. R package version 2.0-3.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*, Vol. 57. London: CRC Press.

Etzioni, R., Kooperberg, C., Pepe, M., Smith, R., and Gann, P. (2003). Combining biomarkers to detect disease with application to prostate cancer. *Biostatistics* **4**, 523–538.

Foster, J. C., Taylor, J. M. G., and Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Statistics in Medicine* **30**, 2867–2880.

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* **55**, 119–139.

Friedman, J., Hastie, T., and Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Annals of Statistics* **28**, 337–407.

Gunter, L., Zhu, J., and Murphy, S. (2007). Variable selection for optimal decision making. *Artificial Intelligence in Medicine* **4594**, 149–154.

Gunter, L., Zhu, J., and Murphy, S. (2011a). Variable selection for qualitative interactions. *Statistical Methodology* **8**, 42–55.

Gunter, L., Zhu, J., and Murphy, S. (2011b). Variable selection for qualitative interactions in personalized medicine while controlling the family-wise error rate. *Journal of Biopharmaceutical Statistics* **21**, 1063–1078.

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Vol. 1. New York: Springer Series in Statistics.

Henry, N. and Hayes, D. (2006). Uses and abuses of tumor markers in the diagnosis, monitoring, and treatment of primary and metastatic breast cancer. *The Oncologist* **11**, 541–552.

Janes, H., Brown, M., Pepe, M., and Huang, Y. (in press). An approach to evaluating and comparing biomarkers for patient treatment selection. *International Journal of Biostatistics*.

Janes, H., Pepe, M., and Bossuyt, P. (2014). A framework for evaluating markers used to select patient treatment. *Medical Decision Making*, **34**, 159–167.

Janes, H., Pepe, M., Bossuyt, P., and Barlow, W. (2011). Measuring the performance of markers for guiding treatment decisions. *Annals of Internal Medicine* **154**, 253.

Lu, W., Zhang, H., and Zeng, D. (2013). Variable selection for optimal treatment decision. *Statistical methods in medical research*, **22**, 493–504.

Mebane, Jr., W. R. and Sekhon, J. S. (2011). Genetic optimization using derivatives: The rgenoud package for R. *Journal of Statistical Software* **42**, 1–26.

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., Baehner, F., Walker, M., Watson, D., Park, T., Hiller, W., Fisher, E., Wickerham, D., Bryant, J., and Wolmark, N. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine* **351**, 2817–2826.

Paik, S., Tang, G., Shak, S., Kim, C., Baker, J., Kim, W., Cronin, M., Baehner, F., Watson, D., Bryant, J., Costantino, J., Geyer Jr, C., Wickerham, D., and Wolmark, N. (2006). Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor–positive breast cancer. *Journal of Clinical Oncology* **24**, 3726–3734.

Pepe, M., Cai, T., and Longton, G. (2005). Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* **62**, 221–229.

Provost, F. and Domingos, P. (2003). Tree induction for probability-based ranking. *Machine Learning* **52**, 199–215.

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

Sekhon, J. S. and Mebane, Jr., W. R. (1998). Genetic optimization using derivatives: Theory and application to nonlinear models. *Political Analysis* **7**, 189–213.

Song, X. and Pepe, M. (2004). Evaluating markers for selecting a patient's treatment. *Biometrics* **60**, 874–883.

Therneau, T., Atkinson, B., and Ripley, B. (2012). *rpart: Recursive Partitioning*. R package version 4.0-1.

Vickers, A., Kattan, M., and Sargent, D. (2007). Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials* **8**, 14.

Zhang, B., Tsiatis, A., Laber, E., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68**, 1010–1018.

Zhao, X., Dai, W., Li, Y., and Tian, L. (2011). AUC-based biomarker ensemble with an application on gene scores predicting low bone mineral density. *Bioinformatics* **27**, 3050–3055.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118.

Received March 2013. *Revised May* 2013.
Accepted December 2013.

Discussions

Eric B. Laber,* Anastasios A. Tsiatis, Marie Davidian, and Shannon T. Holloway

Department of Statistics, North Carolina State University, Raleigh,
North Carolina 27695-8203, U.S.A.
**email:* laber@stat.ncsu.edu

1. Introduction

We congratulate the Kang, Janes, and Huang (hereafter KJH) on an interesting and powerful new method for estimating an optimal treatment rule, also referred to as an optimal treatment regime. Their proposed method relies on having a high-quality estimator for the regression of outcome on biomarkers and treatment, which the authors obtain using a novel boosting algorithm. Methods for constructing treatment rules/regimes that rely on outcome models are sometimes called indirect or regression-based methods because the treatment rule is inferred from the outcome model (Barto and Dieterich, 1988).

Regression-based methods are appealing because they can be used to make prognostic predictions as well as treatment recommendations. While it is common practice to use parametric or semiparametric models in regression-based approaches (Robins, 2004; Chakraborty and Moodie, 2013; Laber, Linn, and Stefanski, in press; Schulte et al., in press), there is growing interest in using nonparametric methods to avoid model misspecification (Zhao et al., 2011; Moodie, Dean, and Sun, 2013). In contrast, direct estimation methods, also known as policy-search methods, try to weaken or eliminate dependence on correct outcome models and instead attempt to search for the best treatment rule within a

pre-specified class of rules (Orellana, Rotnitzky, and Robins, 2010; Zhang et al., 2012a,b, 2013; Zhao et al., 2012). Direct estimation methods make fewer assumptions about the outcome model, which may make them more robust to model misspecification but potentially more variable.

We derive a direct estimation analog to the method of KJH, which we term *value boosting*. The method is based on recasting the problem of estimating an optimal treatment rule as a weighted classification problem (Zhang et al., 2012a; Zhao et al., 2012). We show how the method of KJH can be used with existing policy-search methods to construct a treatment rule that is interpretable, logistically feasible, parsimonious, or otherwise appealing.

2. Setup and Notation

We assume that the available data are (D_i, T_i, Y_i), $i = 1, \ldots, n$, which comprise n independent and identically distributed copies of (D, T, Y), where $Y \in \mathbb{R}^p$ denotes biomarker information; $T \in \{0, 1\}$ denotes treatment received; and $D \in \mathbb{R}$ denotes the outcome of interest coded so that higher values are better, as is customary in the treatment regime literature. To match the development of KJH, we assume that the data are collected in a randomized clinical trial so that $\pi(t|y) = P(T = t|Y = y)$ is known by design, where $\epsilon < \pi(t|Y) < 1 - \epsilon$ for some $\epsilon < 0$ with probability one. This setup includes the binary outcome considered by KJH as a special case with the roles of $D = 1$ and $D = 0$ interchanged.

A treatment rule g is a map from the domain of Y into that of T, so that a patient presenting with $Y = y$ is recommended treatment $g(y)$. Define the value of a rule g, denoted $V(g)$, as the expected outcome if all patients are treated according to g. An optimal regime, g^{opt}, satisfies $V(g^{\text{opt}}) \geq V(g)$ for all other rules g under consideration. Under suitable conditions (e.g., Zhang et al., 2012b) $V(g) = \mathbb{E}[\mathbb{E}\{D|Y, T = g(Y)\}]$. Letting $Q(y, t) = \mathbb{E}(D|Y = y, T = t)$, $g^{\text{opt}}(y) = \arg\max_t Q(y, t)$. The foregoing expression motivates estimating g^{opt} by first estimating the function $Q(y, t)$ by regressing D on Y and T to obtain $\widehat{Q}(y, t)$ and then obtaining $\widehat{g}_{\text{Reg}}(y) = \arg\max_t \widehat{Q}(y, t)$. We refer to approaches of this form as regression-based methods; the method proposed by KJH is of this type.

An alternative approach is to construct first an estimator for $V(g)$, say $\widehat{V}(g)$, and subsequently obtain $\widehat{g}_{\text{Val}} = \arg\max_{g\in\mathcal{G}} \widehat{V}(g)$ for some suitable class of rules $\mathcal{G}$. We refer to approaches of this form as policy-search methods. Policy-search and regression-based methods are not clearly delineated. For example, one could estimate $V(g)$ by $n^{-1}\sum_{i=1}^n \widehat{Q}\{Y_i, g(Y_i)\}$, in which case $\widehat{g}_{\text{Val}} = \widehat{g}_{\text{Reg}}$ provided that $\widehat{g}_{\text{Reg}} \in \mathcal{G}$. Let $\widehat{B}_{KJH}(Y)$ be the estimator of $\widehat{Q}(Y, 1) - \widehat{Q}(Y, 0)$ proposed by KJH. Because $\widehat{Q}\{Y, g(Y)\} = \widehat{Q}(Y, 0) + g(Y)\widehat{B}(Y)$, the method of KJH can be viewed as a policy-search estimator of the form $\arg\max_{g\in\mathcal{G}} n^{-1}\sum_{i=1}^n g(Y_i)\widehat{B}_{KJH}(Y_i)$. However, the success of this approach relies on the correctness of the outcome model.

Generally, policy-search methods employ estimators for $V(g)$ that do not rely heavily on correctness of an outcome model (Zhang et al., 2012b, 2013; Zhao et al., 2012). In the clinical trial setting considered here, both the inverse probability weighted estimator (IPWE) and the augmented inverse probability weighted estimator (AIPWE) described by KJH are consistent estimators for $V(g)$, pointwise in g, that do not require a consistent outcome model. We assume subsequently that the AIPWE is used to estimate $V(g)$. Let $\mathbb{P}_n$ denote the empirical measure so that $\mathbb{P}_n f(Z) = n^{-1}\sum_{i=1}^n f(Z_i)$. The AIPWE is

$$\widehat{V}_{\text{AIPWE}}(g) = \mathbb{P}_n\left[\frac{1_{T=g(Y)}D}{\pi_c(Y;g)} - \frac{1_{T=g(Y)} - \pi_c(Y;g)}{\pi_c(Y;g)}\widehat{Q}\{Y, g(Y)\}\right], \tag{1}$$

where 1_ν is one if ν is true and zero otherwise, $\pi_c(Y;g) = \pi\{g(Y)|Y\}$, and $\widehat{Q}(y, t)$ is an estimator for $Q(y, t)$. The AIPWE can be viewed as adding a mean zero but negatively correlated term to the IPWE to gain efficiency (Scharfstein, Rotnitzky, and Robins, 1999). Define

$$\widehat{C}(D, Y, T) = \frac{T}{\pi(1|Y)}D - \frac{T - \pi(1|Y)}{\pi(1|Y)}\widehat{Q}(Y, 1) - \frac{1 - T}{1 - \pi(1|Y)}D - \frac{T - \pi(1|Y)}{1 - \pi(1|Y)}\widehat{Q}(Y, 0),$$

and $\widehat{C}_i = \widehat{C}(D_i, Y_i, T_i)$. Let $\widehat{W}_i = |\widehat{C}_i|$ and $\widehat{Z}_i = 1_{\widehat{C}_i > 0}$. Zhang et al. (2012a) showed that

$$\arg\max_{g\in\mathcal{G}} \widehat{V}_{\text{AIPWE}}(g) = \arg\min_{g\in\mathcal{G}} \mathbb{P}_n \widehat{W} 1_{\widehat{Z}\neq g(Y)},$$

and thus the policy-search objective function using (1) is equivalent to a weighted misclassification error with predictor-label pairs $\{(Y_i, \widehat{Z}_i)\}_{i=1}^n$ and weights $\{\widehat{W}_i\}_{i=1}^n$. Taking this approach, any classification algorithm that can incorporate weights can be used to estimate an optimal treatment regime. In the next section, we use this framework to develop a policy-search boosting algorithm for estimating g^{opt}.

3. Value Boosting

3.1. *Policy-Search Boosting Algorithm*

To apply boosting using this classification perspective, we require an algorithm that can accommodate weights. However, standard boosting algorithms assume unweighted data. To provide intuition and to match the development of KJH, we describe a simple boosting algorithm with a logistic regression model as the base classifier, although other choices are possible. To gain intuition, consider first a thought experiment. Suppose that each weight $\widehat{W}_i$ is a rational number expressed as q_i/r_i, where q_i and r_i are positive integers. Let N denote the smallest common multiple of $r_1, \ldots, r_n$, and create a new, augmented data set with $q_i N/r_i$ copies of $(Y_i, \widehat{Z}_i)$. Let $\mathcal{D}_{\text{Aug}} = \{(Y_{ik}, \widehat{Z}_{ik})\}$ $k = 1, \ldots, q_i N/r_i$, $i = 1, \ldots, n$ denote this augmented data set, where $Y_{ik} \equiv Y_i$ and $\widehat{Z}_{ik} \equiv \widehat{Z}_i$ for $i = 1, \ldots, n$. In this case

$$\widehat{V}_{\text{AIPWE}}(g) \propto \sum_{i=1}^n \sum_{k=1}^{q_i N/r_i} 1_{\widehat{Z}_{ik}\neq g(Y_{ik})},$$

which is the (unweighted) misclassification rate of the rule g evaluated on the augmented data set. Thus, we can apply a boosting algorithm without modification to the augmented data set. The algorithm we consider is as follows:

1. Initialize weights $\omega_{ik}^{(1)} = 1, k = 1, \ldots, Nq_i/r_i,\ i = 1, \ldots, n$.
2. Repeat for $b = 1, \ldots, B$
   (a) Fit a logistic regression model $\widehat{g}_n^{(b)}$ to $\mathcal{D}_{\text{Aug}}$ with weights $\omega_{ik}^{(b)}$, $k = 1, \ldots, q_i N/r_i$, $i = 1, \ldots, n$.
   (b) Compute $m_{ik}^{(b)} = 1_{\widehat{g}_n^{(b)}(Y_{ik}) \neq \widehat{Z}_{ik}}$ and weighted misclassification error $e^{(b)} = \sum_{i,k} \omega_{ik}^{(b)} m_{ik}^{(b)} / \sum_{i,k} \omega_{ik}^{(b)}$.
   (c) Set $\omega_{iK}^{(b+1)} = \omega_{ik}^{(b)} \left\{(1 - e^{(b)})/e^{(b)}\right\}^{m_{ik}}$.
3. The final boosted treatment rule is

$$\widehat{g}_n(Y) = \begin{cases} 1 & \text{if } \sum_{b=1}^{B} \log\{(1 - e^{(b)})/e^{(b)}\} \widehat{g}_n^{(b)}(Y) \\ & \quad > \frac{1}{2} \sum_{b=1}^{B} \log\{(1 - e^{(b)})/e^{(b)}\}, \\ 0 & \text{otherwise.} \end{cases}$$

This algorithm is not computationally feasible because $\mathcal{D}_{\text{Aug}}$ is potentially very large. However, it suggests a simple weighted procedure that can be applied to the original data. First, note that step 2(a) is equivalent to fitting logistic regression to obtaining $\widehat{g}_n^{(b)}$ using the original data $\{(Y_i, \widehat{Z}_i)\}_{i=1}^n$ with weights $\omega_i^{(b)} = \omega_{i1}^{(b)} \widehat{W}_i$. Second, defining $m_i^{(b)} = 1_{\widehat{g}_n^{(b)}(Y_i) \neq \widehat{Z}_i}$, the weighted misclassification at iteration b of the boosting algorithm is

$$e^{(b)} = \frac{\sum_{i,k} \omega_{ik}^{(b)} m_{ik}^{(b)}}{\sum_{i,k} \omega_{ik}^{(b)}} = \frac{\sum_{i=1}^{n} (Nq_i/r_i) \omega_{i1}^{(b)} m_{i1}^{(b)}}{\sum_{i=1}^{n} (Nq_i/r_i) \omega_{i1}^{(b)}}$$
$$= \frac{\sum_{i=1}^{n} \widehat{W}_i \omega_{i1}^{(b)} m_i^{(b)}}{\sum_{i=1}^{n} \widehat{W}_i \omega_{i1}^{(b)}}.$$

This suggests the following algorithm, which we term value boosting.

1. Initialize weights $\omega_i^{(1)} = n^{-1}$, $i = 1, \ldots, n$.
2. Repeat for $b = 1, \ldots, B$
   (a) Fit a logistic regression model $\widehat{g}_n^{(b)}$ to $\{(Y_i, \widehat{Z}_i)\}_{i=1}^n$ with weights $\omega_i^{(b)}$, $i = 1, \ldots, n$.
   (b) Compute $m_i^{(b)} = 1_{\widehat{g}_n^{(b)}(Y_i) \neq \widehat{Z}_i}$ and weighted misclassification error $e^{(b)} = \sum_i \omega_i^{(b)} \widehat{W}_i m_i^{(b)} / \sum_i \omega_i^{(b)} \widehat{W}_i$.
   (c) Set $\omega_i^{(b+1)} = \omega_i^{(b)} \left\{(1 - e^{(b)})/e^{(b)}\right\}^{m_i}$.
3. The final boosted treatment rule is

$$\widehat{g}_n(Y) = \begin{cases} 1 & \text{if } \sum_{b=1}^{B} \log\{(1 - e^{(b)})/e^{(b)}\} \widehat{g}_n^{(b)}(Y) \\ & \quad > \frac{1}{2} \sum_{b=1}^{B} \log\{(1 - e^{(b)})/e^{(b)}\}, \\ 0 & \text{otherwise.} \end{cases}$$

The form of the above algorithm can be viewed as a version of the adacost algorithm (Fan et al., 1999) and is thus one of many potential boosting algorithms that could be used to construct a direct search analog of the algorithm proposed by KJH.

Table 1

Performance of value boosting and the method proposed by KJH in terms of mean θ based on 1000 Monte Carlo data sets, $n = 500$. Monte Carlo standard deviations are given in parentheses. Scenario indicators correspond to those defined in Section 3.2 of KJH.

Scenario	True θ	Value boosting	KJH
1	0.127	0.113 (0.007)	0.118 (0.004)
2	0.124	0.125 (0.007)	0.128 (0.004)
3	0.134	0.129 (0.007)	0.134 (0.003)
4	0.066	0.036 (0.016)	0.059 (0.014)
5	0.095	0.072 (0.012)	0.077 (0.007)
6	0.139	0.126 (0.005)	0.045 (0.015)
7	0.142	0.124 (0.011)	0.111 (0.010)

3.2. *Boosting the Value or Boosting the Contrast*

Value-boosting focuses on iteratively improving performance of a treatment regime in terms of the estimated value. The method of KJH iteratively attempts to improve an estimator of the contrast function $B(y) = Q(y, 1) - Q(y, 0)$. Because the optimal treatment regime is $g^{\text{opt}}(y) = 1_{B(y)>0}$ and for any treatment regime, say g, $V(g) - \mathbb{E}Q(Y, 0) = \mathbb{E}g(Y)B(y)$, the problems of estimating the contrast and maximizing the value are closely related. However, value-boosting directly targets g^{opt} rather than B, which may be advantageous in settings where g^{opt} is markedly more parsimonious than $B(y)$; because g^{opt} is a function of B, it is necessarily more parsimonious in some sense. An exaggerated example that illustrates this point is $Q(y, t) = t\left(\sum_{j=1}^{p} y_j^2\right) y_1$; in this case $B(y) = \left(\sum_{j=1}^{p} y_j^2\right) y_1$ is a complex nonlinear function of all p variables whereas $g^{\text{opt}}(y) = 1_{y_1>0}$ is a univariate linear threshold. However, the operating characteristics of policy-search and regression-based estimators have not been directly compared on examples of this type; we believe such a comparison would be provide valuable information on the differences between these two classes of methods.

4. Simulation Experiments

To examine the finite sample performance of the value boosting algorithm, we use the seven simulation scenarios considered by KJH with $n = 500$. We use a logistic working model of the form $Q(y, t) = P(D = 1|T = t, Y = y) = \text{expit}(\widetilde{y}^{\intercal}\beta_0 + t\widetilde{y}^{\intercal}\beta_1)$, where $\widetilde{y} = (1, y^{\intercal})^{\intercal}$; β_0, β_1 are unknown parameters; and $\text{expit}(u) = e^u/(1 + e^u)$. We estimate $Q(y, t)$ using maximum likelihood. Our base classifier is a logistic model of the form $P(Z = 1|Y = y) = \text{expit}(\psi_0^{\intercal}\widetilde{y} + \psi_{12}y_1y_2 + \psi_{13}y_1y_3 + \psi_{23}y_2y_3 + \psi_{11}y_1^2 + \psi_{22}y_2^2 + \psi_{33}y_3^2)$. We estimate the performance of value boosting and the method proposed by KJH using the mean θ value as proposed by KJH computed using a test set of size 10^4 and 1000 Monte Carlo replications. Table 1 shows the estimated performance of each method. Value boosting is comparable with KJH in all scenarios except Scenario 6, where it shows a marked advantage. However, value boosting is more variable than KJH, which is antic-

ipated given that policy-search methods are typically more variable than regression-based methods. The simulations suggest that the boosting method proposed by KJH and value boosting may warrant further investigation.

5. Parsimony, Feasibility, and Interpretability

Boosting and other "black-box" estimation algorithms have a strong track record for predictive performance, especially on benchmark data sets. However, it is not apparent that such methods are suited to inform clinical practice, guidelines, or research. One of the strongest features of policy-search methods is that it is possible to control the class of potential treatment rules. Consequently, one can constrain the estimated treatment rule to be interpretable, low-cost, logistically feasible, and so on. However, boosting these methods to improve value may destroy some or all of the foregoing features.

Consider three approaches for estimating an optimal treatment strategy: (A1) optimize within a small but informative class of regimes; (A2) optimize within a large but difficult to interpret class; and (A3) optimize within a large but difficult class and then project the estimated regime onto a smaller, interpretable class. Approach (A1) is taken by Orellana et al. (2010), Zhang et al. (2012a,b), and Zhang et al. (2013). In addition to the benefits of interpretability and feasibility imposed on the estimated regime, in some contexts, (A1) is appealing because the best regime within a pre-specified class of regimes is of independent interest. Approach (A2), taken by Zhao et al. (2011, 2012) and Moodie et al. (2013), uses the predictive power of machine-learning methods and reduces the risk of model misspecification. Approach (A3) attempts to maintain both interpretability and performance (e.g., Breiman and Shang, 1996); furthermore, the "residual" of the estimated treatment regime relative to its projection onto the smaller class may be informative about where and how the smaller class is not sufficiently expressive.

The trade off between flexibility and interpretability is ubiquitous in statistical modeling. However, the importance of each component depends on the context in which the model is estimated. Estimation of an optimal treatment regime is often conducted as a secondary analysis aimed at generating scientific hypotheses and informing follow-up studies. In contrast, KJH seem to be primarily interested in constructing a high-quality treatment regime. Their boosting method seems ideally suited for this purpose.

Acknowledgement

This work was supported by NIH grant P01 CA142538.

References

Barto, A. and Dieterich, T. (2004). Reinforcement learning and its relation to supervised learning. In *Handbook of Learning and Approximate Dynamic Programming*, J. Si, A. G. Barto, W. B. Powell, and D. Wunsch (eds), 45–63. New York: Wiley.

Breiman, L. and Shang, N. (1996). Born again trees. Available at ftp://ftp.stat.berkeley. edu/pub/users/breiman/BAtrees.ps.

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

Fan, W., Stolfo, S. J., Zhang, J., and Chan, P. K. (1999). Adacost: misclassification cost-sensitive boosting. In *Proceedings of the Sixteenth International Conference on Machine Learning (ICML'99)*, 97–105, Bled, Slovenia, June 1999.

Laber, E., Linn, K., and Stefanski, L. (in press). Interactive model-building for Q-learning. *Biometrika*.

Moodie, E. E. M., Dean, N., and Sun, Y. R. (2013). Q-learning: Flexible learning about useful utilities. *Statistics in Biosciences*, 1–21.

Orellana, L., Rotnitzky, A., and Robins, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal treatment regimes, part I: Main content. *International Journal of Biostatistics* **6**, Article 8.

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

Scharfstein, D. O., Rotnitzky, A., and Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association* **94**, 1096–1120.

Schulte, P., Tsiatis, A., Laber, E., and Davidian, M. (in press). Q-and A-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science*.

Zhang, B., Tsiatis, A. A., Davidian, M., Zhang, M., and Laber, E. (2012a). Estimating optimal treatment regimes from a classification perspective. *Stat* **1**, 103–114.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012b). A robust method for estimating optimal treatment regimes. *Biometrics* **68**, 1010–1018.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika* **100**, 681–694.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118.

Zhao, Y., Zeng, D., Socinski, M. A., and Kosorok, M. R. (2011). Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics* **67**, 1422–1433.

Lu Tian

Department of Health Research and Policy, Stanford University,
Palo Alto, California 94305, U.S.A
email: lutian@stanford.edu

I would like to congratulate Kang, Janes, and Huang for their interesting article on a novel boosting algorithm for combining multiple biomarkers to optimize patient treatment recommendations. Their results show the potential of modern machine learning methods especially when one cannot correctly specify a regression model for the complex

underlying relationship of interest. Since the publication of the AdaBoosting procedure by Freund and Schapire (1997), there has been great interest in seeking an explanation for its excellent performance and generalizing the method in various settings. For example, AdaBoosting can be viewed as a numerical algorithm performing forward stagewise regression with a set of weak classifiers to minimize a regularized exponential loss function. Along this line, several versions of boosting procedures have been proposed by changing the loss function to be minimized (Friedman, Hastie, and Tibshirani, 2000). Boosting is arguably one of the best off-the-shelf methods and it is not a surprise that when appropriately adapted, it can perform competitively well for differentiating patients with positive treatment benefit from those without. On the other hand, after carefully examining the proposed boosting algorithm in the article, I see some crucial departures from the classical version of boosting. Therefore, I will first discuss the differences between the proposed method and the conventional boosting algorithm and their potential implications. Secondly, I will propose simple alternatives to resolve some of the difficulties. To simplify the discussion, I will assume $P(T=1)=P(T=0)=0.5$ throughout.

1. The Difference between the Proposal and Convention Boosting

Suppose for the time being that the directions of the treatment effect, $S_i = I\{\Delta(Y_i) \le 0\}, i = 1, \ldots, n$ are observed, and then the real AdaBoosting (a generalized version of AdaBoosting returns a class probability) proceeds as follows:

(1) Start with weights $w_i^{(0)} = 1/n$ for subjects $i = 1, \ldots, n$, and fit a working model to calculate $\widetilde{\Delta}^{(0)}(y)$, the estimated probability $P(S=1|Y=y)$ based on the working model.

(2) For $m = 1, \ldots, M$,

- update the weights according to $w_i^{(m)} = w_i^{*(m)} / \sum_{i=1}^n w_i^{*(m)}$, where

$$w_i^{*(m)} = w_i^{(m-1)} \exp\left(-\frac{1}{2}(2S_i - 1)\log\left[\frac{\widetilde{\Delta}^{(m-1)}(y_i)}{1-\widetilde{\Delta}^{(m-1)}(y_i)}\right]\right)$$

  for $i = 1, \ldots, n$.
- Refit the working model with updated weights $w_i^{(m)}$ to obtain $\widetilde{\Delta}^{(m)}(y)$, the updated estimated probability $P(S=1|Y=y)$.

(3) After the last iteration, the estimated treatment rule is

$$\hat{\phi}(Y_i) = I\left(\sum_{m=1}^{M} \log\left[\frac{\widetilde{\Delta}^{(m)}(y_i)}{1-\widetilde{\Delta}^{(m)}(y_i)}\right] \ge 0\right).$$

Although the assumption that S_i are all observed is artificial, it reveals an important message in comparing the proposed boosting algorithm with the ideal counterpart above. Firstly, while the updated weight in the ideal Real AdaBoosting depends on the fitted probability, the old weight and the true outcomes S_i, the weight in the proposed algorithm depends only on the fitted probability. Intuitively, the Real AdaBoosting always tries to upweight misclassified observations in the previous step to ensure that the new classification rule likely adds value to existing ones, especially on subgroup of observations for which the performance of the old classifiers is unsatisfactory. Since S_i is unobserved, the new algorithm upweights observations close to the estimated decision boundary, whose classification (optimal treatment recommendation) is relatively ambiguous. It is clear that there is a genuine difference in the reweighting scheme: following the same logic of the new algorithm, the Real AdaBoosting would upweights observations with the estimated class probability close to 0.5, which may or may not be the same observations misclassified in the previous step. One certainly can imagine that if the initial working model is a poor approximation to the truth, then some misclassified observations could mistakenly have an estimated class probability close to 0 or 1 and receive less weight. Furthermore, even when the working model yields reasonable class probability estimates, the misclassification that occurred near the decision boundary affects the value-function θ the least. Therefore, one may wonder if the new proposal sometimes fails to upweight the "right" observations. Secondly, while the Real AdaBoosting can be viewed as an algorithm to minimize a regularized loss function, the proposed algorithm can not be easily fit into the same framework. One consequence is that it is not clear if the algorithm provides a consistent estimate to the optimal treatment rule $I\{\Delta(Y) \le 0\}$. To gain some insight into the consistency issue, lets consider the convergence limit of $\tilde{P}^{(1)}(D=1|T=t,Y)$ in the proposed boosting procedure. The final estimator $\widehat{\Delta}(Y)$ should be close to the difference between the two limits after sufficient number of iterations. When the simple logistic working model is used, the limit of $\tilde{P}^{(1)}(D=1|T=t,Y)$ is $\{1+\exp(-\beta_t' Y)\}^{-1}$ and β_t solves the estimating equation

$$E\left[Y\tilde{w}\left(\left|\frac{\exp(\beta_1' Y)}{1+\exp(\beta_1' Y)} - \frac{\exp(\beta_0' Y)}{1+\exp(\beta_0' Y)}\right|\right) \times \left\{P(D=1|T=t,Y) - \frac{\exp(\beta_t' Y)}{1+\exp(\beta_t' Y)}\right\}\right] = 0, t = 0, 1,$$

as $n \to \infty$, which implies that

$$E\left[Y\tilde{w}\left(|\widetilde{\Delta}(Y)|\right)\left\{\Delta(Y) - \widetilde{\Delta}(Y)\right\}\right] = 0,$$

where $\widetilde{\Delta}(Y) = \{1+\exp(-\beta_0' Y)\}^{-1} - \{1+\exp(-\beta_1' Y)\}^{-1}$ can be viewed as a reasonable approximation to $\Delta(Y)$, especially in the region where $\beta_0' Y \approx \beta_1' Y$. However, it seems that in general $\widetilde{\Delta}(Y)$ is different from $\Delta(Y)$ unless the simple logistic working model is correctly specified. The extra weight function does not solve this problem. This is in contrast to the appealing "consistency" property of the original AdaBoosting procedure, which does not require the "weak" learners to be "strong" or "correct."

2. Alternative Boosting Procedure

From the discussion in the previous section, it seems that the new proposal may sacrifice some important components in the AdaBoosting procedure due to the simple fact that

$S_i, i = 1, \ldots, n$ are not observable in practice. This may cause concerns about its performance in settings beyond those investigated in the comprehensive simulation study. On the other hand, noticing the fact that

$$E\{(2D-1)(2T-1)|Y\} = -\Delta(Y),$$

one may treat the binary variable $\tilde{S}_i = (2D_i - 1)(2T_i - 1)$ as the surrogate to the unobserved $S_i = I\{-\Delta(Y_i) \geq 0\}$ (Signorovitch, 2007) and simply replace S_i by $\tilde{S}_i$ in the Real AdaBoosting procedure described above. To justify this approach, one notes that AdaBoosting can be viewed as a numerical algorithm to minimize the empirical version of the loss function

$$J(d) = E\left[\exp\{-\tilde{S}d(Y)\}\right]$$

in terms of $d(\cdot)$, a function of y. In this case, the minimizer of $J(d)$ is

$$d_1(y) = \frac{1}{2}\log\left\{\frac{P(\tilde{S}=1|Y=y)}{P(\tilde{S}=-1|Y=y)}\right\} = \frac{1}{2}\log\left\{\frac{1-\Delta(y)}{1+\Delta(y)}\right\}.$$

It suggests that $I\{d_1(Y) \geq 0\}$, the treatment assignment rule based on $d_1(\cdot)$, is equivalent to our target $I\{\Delta(Y) \leq 0\}$. Therefore, as a numerical algorithm to estimate $I\{d_1(Y) \geq 0\}$, the output of AdaBoosting with $\tilde{S}_i, 1 \leq i \leq n$, as responses can be used to assign patients according to their individual treatment effects.

The second alternative is to estimate the optimal treatment rule by maximizing the mean outcome. As the authors advocated, the mean outcome is a clinically relevant and interpretable quantity to measure the value of a given treatment assignment rule. In the current setting, maximizing the mean outcome is equivalent to minimizing

$$\sum_{i=1}^{n} D_i I\{T_i \neq g(Y_i)\} = \sum_{i:D_i=1} I\{T_i \neq g(Y_i)\},$$

the misclassification error for predicting T_i among the subgroup of patients having outcome $D_i = 1$, with respect to the binary classification rule $g : Y \to \{0, 1\}$. This again provides a very natural platform for applying the AdaBoosting algorithm. Specifically, one may use AdaBoosting procedure to construct a classification rule to classify a patient to his/her actual randomized treatment assignment based on Y for the subgroup of patients with $D = 1$. To justify the validity of this boosting procedure, one only needs to note that the minimizer of the corresponding loss function $J(d) = E[\exp\{-(2T-1)d(Y)\} \mid D = 1]$ is

$$d_2(y) = \frac{1}{2}\log\left\{\frac{P(T=1|Y=y, D=1)}{P(T=0|Y=y, D=1)}\right\} = \frac{1}{2}\log\left\{\frac{P(D=1|T=1, Y=y)}{P(D=1|T=0, Y=y)}\right\}.$$

Table 1
The results of simulation study with a sample size of $n = 500$ for scenarios 4 and 6 from Table 2 of the article by Kang, Janes, and Huang. The results of the two new methods are summarized in the last two columns under the names "$\tilde{S}$" and "value-function."

			Working model	Linear logistic				Classification tree		Separate AdaBoosting	**New boosting**	
Scenario	True θ		Algorithm	MLE	Boosting	IPWE	AIPWE	Single tree	Boosting		$\tilde{S}$	Value-function
4	0.0657	θ	Mean	0.0574	0.0607	0.0561	0.0567	0.0221	0.0378	0.0352	0.0396	0.0351
		MCR	Mean	0.1511	0.1206	0.1719	0.1653	0.5397	0.3251	0.5304	0.3270	0.3374
6	0.1393	θ	Mean	0.0236	0.0438	0.0498	0.0544	0.0978	0.1186	0.1010	0.1006	0.1011
		MCR	Mean	0.3865	0.3542	0.3452	0.3330	0.2697	0.1762	0.2433	0.2290	0.2227

Thus, the $d_2(\cdot)$-based treatment assignment rule $I\{d_2(Y) \geq 0\} = I\{\Delta(Y) \leq 0\}$ and the corresponding boosting procedure provides a flexible algorithm to approximate the decision boundary of the optimal treatment rule.

I have selected the two most challenging scenarios (4 and 6) in the simulation study performed by Kang, Janes, and Huang to examine the relative performance of the above two proposals. For both cases, the "ada" function with default parameters in R is used to perform the boosting algorithm. The sample size of the training set is 500 and final results are based on 1000 Monte-Carlo replications (Table 1). It seems that the empirical performances of these two simple alternatives are comparable to those proposed by Kang, Janes, and Huang with classification tree as the base learner.

3. Remarks

We statisticians can learn a great deal from the rapid development of machine learning techniques, which oftentimes offer robust performance for a broad range of problems. The authors convincingly demonstrated the power of a version of the generalized boosting method for estimating the optimal strategy for assigning treatment. On the other hand, the proposed boosting method is different from its conventional counterpart in several key aspects and new explanations are needed to account for its good performance. Furthermore, there are alternative boosting procedures, which in my opinion can circumvent the difficulties caused by unobservable class labels $S_i = I\{\Delta(Y_i) \leq 0\}, i = 1, \ldots, n$, in more natural ways. Further research in these directions is certainly warranted.

References

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, **55**, 119–139.

Friedman, J., Hastie, T., and Tibshirani, R. (2000). Additive logistic regression: A statistics view of boosting (with discussion and a rejoinder by the authors). *Annals of Statistics*, **28**, 337–407.

Signorovitch, J. E. (2007). *Identifying informative biological markers in high-dimensional genomic data and clinical trials*. Ph.D. thesis, Harvard University.

Ying-Qi Zhao*

Department of Biostatistics and Medical Informatics,
University of Wisconsin-Madison, Madison, Wisconsin 53792, U.S.A.
**email:* yqzhao@biostat.wisc.edu

and

Michael R. Kosorok

Department of Biostatistics, University of North Carolina at Chapel Hill,
Chapel Hill, North Carolina 27599, U.S.A.

Department of Statistics and Operations Research,
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, U.S.A.

SUMMARY. Kang, Janes and Huang propose an interesting boosting method to combine biomarkers for treatment selection. The method requires modeling the treatment effects using markers. We discuss an alternative method, outcome weighted learning. This method sidesteps the need for modeling the outcomes, and thus can be more robust to model misspecification.

KEY WORDS: Boosting; Outcome weighted learning; Personalized medicine; Support vector machine.

1. Introduction

The problem of combining markers to optimize treatment selection has recently received significant attention among statistical researchers. We congratulate Kang, Janes, and Huang (in press; hereafter K.J.H.) on an elegant contribution to this important area. K.J.H. provide a novel application of boosting to find the best marker combination. In particular, with a binary outcome, the optimal treatment rule is determined by the conditional probability of having disease, that is, the risk, given the covariates and the treatment. If the risk models are correctly specified, the optimal rule can be deduced accordingly. While a generalized linear model is a simple and popular option, it may suffer from model misspecification. The proposed method in K.J.H. achieves a measure of robustness to such model misspecification through the use of boosting, combined with iteratively reweighting each subject's potential misclassification based on treatment benefit in the previous iteration.

Although K.J.H. indicate that the purpose of the proposed method is to classify subjects according to the unobserved optimal treatment decision rule, the approach does not utilize a clear objective function for optimization. Risk modeling is required to estimate the optimal rules and to further update the weights at each step. As shown in the simulation results, the performances vary with different working models. An alternative approach, outcome weighted learning (OWL), is given

Table 1

Results of the simulation study with sample size $n = 500$. OWL-HL: outcome weighted learning with hinge loss, linear kernel; OWL-HG: outcome weighted learning with hinge loss, Gaussian kernel; OWL-E: outcome weighted learning with exponential loss; L-Boosting: linear logistic boosting; C-boosting: classification tree boosting: Logit: logistic regression.

SN	True θ			OWL-HL	OWL-HG	OWL-E	L-Boosting	C-Boosting	Logit
1	0.1268	θ	Mean	0.1225	0.1057	0.1020	0.1240	0.1149	**0.1256**
			SD	0.0025	0.0038	0.0034	0.0018	0.0040	0.0019
		MCR_{TB}	Mean	0.0640	0.1363	0.1620	0.0517	0.1128	0.0309
2	0.1243	θ	Mean	0.1224	0.1129	0.0946	**0.1248**	0.0921	0.1241
			SD	0.0024	0.0052	0.0024	0.0017	0.0037	0.0017
		MCR_{TB}	Mean	0.0664	0.1160	0.1883	0.0481	0.1811	0.0545
3	0.1341	θ	Mean	0.1256	0.1191	0.1010	**0.1315**	0.1154	**0.1315**
			SD	0.0026	0.0045	0.0031	0.0015	0.0051	0.0016
		MCR_{TB}	Mean	0.0865	0.1062	0.1461	0.0498	0.1307	0.0499
4	0.0657	θ	Mean	0.0654	0.0061	0.0436	**0.0657**	0.0356	**0.0657**
			SD	0.0058	0.0070	0.0031	0.0031	0.0031	0.0040
		MCR_{TB}	Mean	0.0855	0.4638	0.2721	0.0202	0.2853	0.0489
5	0.0950	θ	Mean	**0.0798**	0.0646	0.0783	0.0594	0.0517	0.0592
			SD	0.0046	0.0034	0.0035	0.0030	0.0037	0.0030
		MCR_{TB}	Mean	0.2049	0.2250	0.1810	0.3110	0.2708	0.3070
6	0.1393	θ	Mean	0.0001	**0.093**1	0.0903	0.0313	0.0926	−0.0023
			SD	0.0049	0.0035	0.0035	0.0032	0.0034	0.0055
		MCR_{TB}	Mean	0.4010	0.2264	0.2267	0.3469	0.2408	0.4296
7	0.1419	θ	Mean	0.1295	0.1217	0.1140	0.1312	0.1185	**0.1313**
			SD	0.0028	0.0056	0.0038	0.0015	0.0044	0.0015
		MCR_{TB}	Mean	0.0718	0.1123	0.1455	0.0468	0.1448	0.0468

in Zhao et al. (2012), which estimates the optimal treatment decision rule through a weighted classification procedure that incorporates outcome information. To be more specific, the optimization target, which directly leads to the optimal treatment decision rule, can be viewed as a weighted classification error where each subject is weighted proportional to his or her clinical outcome. In the next section, we will briefly introduce the idea of OWL and modify it to the binary outcome setup. In Section 3, we present simulation studies comparing OWL with the boosting method proposed by KJH. We conclude with a brief discussion in Section 4.

2. Outcome Weighted Learning (OWL)

Using the same notation as K.J.H., we let $D \in \{0, 1\}$ be the binary indicator of an adverse outcome, T indicate treatment ($T = 1$) or not ($T = 0$), and Y be the marker which can be used to identify a subgroup. We assume that the data $\{D_i, T_i, Y_i\}_{i=1}^n$ are from a randomized clinical trial. For arbitrary treatment rule $g : Y \mapsto \{0, 1\}$, the expected benefit under g, that is, if g were implemented in the whole population, can be written as (Qian and Murphy, 2011)

$$\mathbb{P}\left\{\frac{I(D=0)I(T=g(Y))}{\pi_c(Y)}\right\}, \tag{1}$$

where $\pi_c(Y)$ is the known probability of treatment, and the optimal treatment rule $g^{\text{opt}}(Y)$ can be obtained by maximizing the above quantity. Equivalently, by minimizing

$$\mathbb{P}\left\{\frac{I(D=0)I(T\neq g(Y))}{\pi_c(Y)}\right\},$$

we can obtain that $g^{\text{opt}}(Y) = \mathbf{1}\{\Delta(Y) \le 0\}$, where $\Delta(Y) = P(D = 1|T = 0, Y) - P(D = 1|T = 1, Y)$. Indeed, this can be viewed as a weighted classification error in the setting where we wish to classify T in the responders group using the covariate Y, with $1/\pi_c(Y)$ as the weights. In particular, when $\pi_c(Y) = 0.5$, the problem falls into a regular classification framework if we only consider responders. In this case, we classify subjects according to their assigned treatments only among those who received benefit from the assignment (i.e., the responders).

Provided with the data, we could potentially minimize the empirical analog

$$\min_g \mathbb{P}_n\left\{\frac{I(D=0)I(T\neq g(Y))}{\pi_c(Y)}\right\} \tag{2}$$

to estimate $g^{\text{opt}}(Y)$, where $\mathbb{P}_n$ denotes the empirical average. Note that this is similar to the quantity IPWE(η) presented in Zhang et al. (2012). Due to the nonconvexity and discontinuity of the 0-1 loss, it is computationally difficult to minimize (2). We address this problem by using a convex surrogate loss function to replace the 0-1 loss, a common practice in the field of machine learning literature (Zhang, 2004; Bartlett, Jordan, and McAuliffe, 2006). In other words, instead of minimizing

Table 2

Results of the simulation study with sample size $n = 5000$. OWL-HL: outcome weighted learning with hinge loss, linear kernel; OWL-HG: outcome weighted learning with hinge loss, Gaussian kernel; OWL-E: outcome weighted learning with exponential loss; L-Boosting: linear logistic boosting; C-boosting: classification tree boosting: Logit: logistic regression.

SN	True θ			OWL-HL	OWL-HG	OWL-E	L-Boosting	C-Boosting	Logit
1	0.1268	θ	Mean	0.1228	0.1256	0.1218	**0.1265**	0.1177	**0.1265**
			SD	0.0025	0.0038	0.0034	0.0018	0.0040	0.0019
		MCR_{TB}	Mean	0.0776	0.0338	0.0709	0.0124	0.0961	0.0111
2	0.1243	θ	Mean	0.1251	0.1246	0.1233	**0.1264**	0.1083	**0.1264**
			SD	0.0024	0.0052	0.0024	0.0017	0.0037	0.0017
		MCR_{TB}	Mean	0.0442	0.0459	0.0598	0.0162	0.1201	0.0181
3	0.1341	θ	Mean	0.1319	0.1316	0.1284	**0.1338**	0.1204	0.1337
			SD	0.0026	0.0045	0.0031	0.0015	0.0051	0.0016
		MCR_{TB}	Mean	0.0408	0.0433	0.0628	0.0110	0.1095	0.0161
4	0.0657	θ	Mean	**0.0653**	0.0634	0.0640	0.0652	0.0530	0.0652
			SD	0.0058	0.0070	0.0031	0.0031	0.0031	0.0040
		MCR_{TB}	Mean	0.0357	0.1515	0.1189	0.0090	0.2256	0.0035
5	0.0950	θ	Mean	0.0792	**0.0923**	0.0893	0.0761	0.0750	0.0732
			SD	0.0046	0.0034	0.0035	0.0030	0.0037	0.0030
		MCR_{TB}	Mean	0.2165	0.0693	0.0878	0.2364	0.2215	0.2522
6	0.1393	θ	Mean	0.0417	**0.1315**	0.1258	0.0457	0.1116	0.0123
			SD	0.0049	0.0035	0.0035	0.0032	0.0034	0.0055
		MCR_{TB}	Mean	0.3574	0.1081	0.1349	0.3490	0.2585	0.4039
7	0.1419	θ	Mean	0.1294	0.1319	0.1288	**0.1327**	0.1218	**0.1327**
			SD	0.0028	0.0056	0.0038	0.0015	0.0044	0.0015
		MCR_{TB}	Mean	0.0928	0.0440	0.0789	0.0329	0.1094	0.0307

(2), we minimize

$$\min_{f\in\mathcal{F}} \mathbb{P}_n \left\{ \frac{I(D=0)\phi\{(2T-1)f(Y)\}}{\pi_c(Y)} \right\} + \lambda_n \|f\|^2, \qquad (3)$$

where $\mathcal{F}$ is the functional space that f resides in, $2T-1$ is used to rescale T to reside in $\{-1, 1\}$, $\phi(t)$ is a convex surrogate loss function, and $\lambda_n\|f\|^2$ is a regularization term to avoid overfitting, with $\|\cdot\|$ denoting the norm of f in $\mathcal{F}$ and λ_n controlling the amount of penalization. The estimated treatment rule is $\hat{g}(Y) = I(\hat{f}(Y) > 0)$, where $\hat{f}$ is the solution to (3). We can specify $\mathcal{F}$ to be a linear functional space if we are only interested in linear decision rules. We can also consider nonlinear functional spaces where treatment effects can potentially be complex and nonlinear. In the simulation section, we will examine the performances using two popular choices for $\phi(t)$, including the hinge loss $\phi(t) = \max(1-t, 0)$ and the exponential loss $\phi(t) = \exp(-t)$.

Since being at high risk does not necessarily imply a larger benefit from treatment, we aim to find methods that are optimized for treatment selection (Kang, Janes, and Huang, in press). We point out that OWL directly targets the optimal decision rule, hence the covariate-treatment interaction effects are separated from the main effects. Indeed, it does not involve a modeling step for the risk and $\Delta(Y)$ as required by K.J.H. If the functional space $\mathcal{F}$ is correctly specified for the interaction effects, we can consistently estimate $g^{\text{opt}}(Y)$ (Zhao et al., 2012). Specifically, using the exponential loss, we generalize Adaboost (Freund and Schapire, 1997; Friedman, Hastie, and Tibshirani, 2000) to select the optimal treatment. Rather than using the weight function $\tilde{w}\{\Delta(Y)\}$ to be updated at each step, we use instead $I(D_i = 0)\exp(-T_i f(Y_i))$ as the weight for each observation, where the $f(Y_i)$ are repeatedly updated in each iteration.

As a side note, the OWL method can be naturally generalized to continuous outcomes, which commonly occurs in practice. For example, if we let R denote the continuous outcomes, with larger values being more preferable, we only need to change $I(D = 0)$ to R in (1). The subsequent derivation and computation follow accordingly.

3. Simulation Studies

We compare the OWL method with logistic regression and with the boosting methods (both linear logistic boosting and classification tree boosting) proposed by K.J.H. The OWL methods are implemented using the hinge loss and the exponential loss as the convex surrogates. We use the same simulation scenarios as presented in K.J.H. Since patients are equally randomized to $T = 0$ or 1, $\pi_c(Y)$ can be dropped from the optimization objective (3). Thus, we can apply the standard classification algorithm to the simulated data by only considering the responders with $D = 0$. The adaboost (Freund and Schapire, 1997) or support vector machine (SVM)(Cortes and Vapnik, 1995) can be carried out for this subset of patients, by treating their assignments $T \in \{0, 1\}$ as the class labels and the biomarkers Y as the predictors. The adaboost is implemented by the R function **ada** (R package ada (Culp, Johnson, and Michailides, 2006)) using the default settings with exponen-

tial loss function. The SVM is implemented by the R function **svm** (R package e1071 (Dimitriadou et al., 2008)). Both linear and Gaussian kernels are used for comparison, yielding linear and nonlinear decision rules, respectively.

For each scenario, 1000 data sets are generated as training data to build the treatment decision rule, $\hat{g}(Y)$. A large independent test data set with $n = 10^5$ observations is generated to evaluate the performance of the obtained $\hat{g}(Y)$ under different methods. Mean and Monte-Carlo standard deviation (SD) of $\theta\{\hat{g}(Y)\}$ and mean $\text{MCR}_{\text{TB}}\{\hat{g}(Y)\}$ are reported, where θ and MCR_{TB} are defined in K.J.H. The method is marked in **bold** if it outperforms other methods by producing the highest mean θ. Logistic regression performs the best when the model is correctly specified, which is anticipated and has been noted in K.J.H. The linear logistic boosting has a similar performance to logistic regression if the models are correct, and can improve on logistic regression to some extent when the models are misspecified. However, if the effects are nonlinear, classification tree boosting may be a better option. An appropriate working model is important in practice, when the underlying truth is masked. The OWL method has comparable performances with linear treatment effects. When the outcome models are complex with nonlinear treatment interactions (Scenarios 5 and 6), the OWL methods lead to better performances. In general, OWL using hinge loss with Gaussian kernel has a favorable performance which is always close-to-optimal, especially when the sample size is large.

4. Discussion

In summary, the intuition underlying the proposed boosting method by classifying subjects according to their (unobserved) optimal treatment holds much promise. However, in practice, attention must be paid to the selection of the working models and the weight function $\tilde{w}$ since they may impact the results to some extent. The OWL procedure discussed for comparison utilizes a machine learning approach to find the best treatment rule through optimizing a target function which directly reflects the overall benefit of the decision rule. When the outcome is binary, the method proposed in K.J.H. can have a better performance with small sample sizes, given that the OWL essentially only uses information from the responders. Also, if there is prior information on the relationship between outcomes and candidate biomarkers, the K.J.H. method can be preferable. On the other hand, due to its flexibility in handling covariate-treatment interaction effects, OWL can yield better results when the dimension of the covariate space is high or when the true model is fairly complex. For large samples sizes, OWL with hinge loss and Gaussian kernel performs nearly optimally in every scenario. Additionally, while OWL can also readily handle additional data types such as continuous outcomes, this is not yet the case for the K.J.H. method and so it would be worthwhile to investigate the possibility of such a generalization.

References

Bartlett, P. L., Jordan, M. I., and McAuliffe, J. D. (2006). Convexity, classification, and risk bounds. *JASA* **101**, 138–156.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning* **20**, 273–297.

Culp, M., Johnson, K., and Michailides, G. (2006). ada: An r package for stochastic boosting. *Journal of Statistical Software* **17**, 1–27.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., and Weingessel, A. (2008). Misc Functions of the Department of Statistics (e1071), TU Wien. *R Package*, 1–5.

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* **101**, 119–139.

Friedman, J., Hastie, T., and Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *Annals of Statistics* **101**, 337–374.

Kang, C., Janes, H., and Huang, Y. (in press). Combining biomarkers to optimize patient treatment recommendations. *Biometrics*.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of Statistics* **101**, 1180–1210.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **101**, 1010–1018.

Zhang, T. (2004). Statistical behavior and consistency of classification methods based on convex risk minimization. *Annals of Statistics* **101**, 56–134.

Zhao, Y. Q., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *JASA* **101**, 1106–1118.

Menggang Yu*

Department of Biostatistics & Medical Informatics,
University of Wisconsin, Madison, Wisconsin, U.S.A.
**email:* meyu@biostat.wisc.edu

and

Quefeng Li

Department of Operations Research and Financial Engineering,
Princeton University, Princeton, New Jersey, U.S.A.

1. Introduction

Kang, Janes, and Huang (KJH) proposed a new approach to study an important problem of comparative treatment selection based on selected markers. Although correct prediction of outcomes under different treatments facilitates treatment selection, such prediction can be hard due to model

Table 1
Performance comparison from various methods

	Non-boosting methods		**Boosting methods**		
	Logistic	Single tree	Logistic	Tree	M-boost
Model 1: scenario 1 in K.J.H.; true $\theta = 0.1270$					
MCR_{TB}	**0.0368**	0.2367	0.0415	0.1300	0.0632
$\hat{\theta}$	**0.1255**	0.1089	0.1252	0.1088	0.1155
$\text{SD}(\hat{\theta})$	0.0025	0.0090	0.0027	0.0157	0.0024
Model 2: scenario 5 in K.J.H.; true $\theta = 0.0956$					
MCR_{TB}	0.3264	0.2235	0.3255	0.2102	**0.1246**
$\hat{\theta}$	0.0535	0.0664	0.0537	0.0674	**0.0876**
$\text{SD}(\hat{\theta})$	0.0053	0.0160	0.0048	0.0119	0.0036
Model 3: $\log\{-\log(1 - P(D = 1\|T, Y))\} = 0.1 - 0.2Y_1 + 0.2Y_2 - Y_1Y_2 + T(-0.5 - Y_1 + Y_2 + 3Y_1Y_2)$ *true* $\theta = 0.1575$					
MCR_{TB}	0.3676	0.2057	0.3689	0.1891	**0.1757**
$\hat{\theta}$	0.0414	0.1280	0.0414	0.1259	**0.1431**
$\text{SD}(\hat{\theta})$	0.0029	0.0158	0.0028	0.0240	0.0032
Model 4: $Pr(D = 1\|T, Y)/Pr(D = 0\|T, Y) = \|0.1 - 0.2Y_1 + 0.2Y_2 - Y_1Y_2 + T(-0.5 - Y_1 + Y_2 + 3Y_1Y_2)\|$ *true* $\theta = 0.2345$					
MCR_{TB}	0.3020	0.2514	0.3083	0.2719	**0.2544**
$\hat{\theta}$	0.1755	**0.1917**	0.1743	0.1804	0.1882
$\text{SD}(\hat{\theta})$	0.0096	0.0153	0.0089	0.0218	0.0091

mis-specification, especially when there are multiple markers involved. On the other hand, treatment selection basically depends on the sign of the contrast function $\Delta(Y)$. Direct modeling this contrast function or its sign may lead to more parsimonious and robust results. For example, it can be shown that, if the binary response D follows a single index model as $\text{logitP}(D = 1|T, Y) = h\{l(Y) + g(TY'\beta)\}$ where $h(\cdot)$ and $g(\cdot)$ are increasing functions and $l(\cdot)$ is arbitrary, then for any given Y, $\Delta(Y) \leq 0$ only and if only $Y'\beta \geq 0$. Therefore the treatment rule is substantially simpler than the outcome model. This formulation of the problem into a classification framework nicely connects with a vast amount of available machine learning algorithms. For example, the procedure can be easily generalized to treatment selection among three or more treatments by utilizing multicategory boosting (Zou, Zhu, and Hastie, 2008; Schapire and Freund, 2012, Chapter 10).

2. The Boosting Algorithm

Our main comments are related to the boosting algorithm used by K.J.H. This article seems to bring a new dimension to the classification problem. Even though we observe D, the target function is actually based on $\Delta(Y)$, a quantity not directly observable. So it is not a completely supervised learning problem. This brings challenges to the following aspects.

(1) Sensitivity of the working model

A basic principle of the boosting algorithm is to "boost" weak learners to an adequate learner. The boosting iteratively "trains" the weak classifiers using a weighting distribution. The weights are usually related to the weak learners' accuracy: misclassified examples gain weight and correctly classified examples lose weight. Thus, future weak learners focus more on the examples that previous weak learners misclassified. This scheme seems to suit the treatment selection problem very well. Because of the possibility of model misspecification, $\tilde{\Delta}(Y)$ is a weak learner. This weak learner can then be boosted into a strong one with properly chosen weights based on some function of $\tilde{\Delta}(Y)$. However different from the usual boosting setting, these weights may be incorrect. To conform with the principle of the boosting algorithm, one's hope is that misclassified examples would still gain weight and correctly classified examples lose weight. Therefore the working model plays a dual role here. Mis-specification of the working model may be severe enough that this principle does not hold anymore.

(2) Thresholds of weights

K.J.H. controls the influence from those observations with tiny $\tilde{\Delta}(Y)$ using a thresholding C_m. The choice of C_m was shown to have minimum influence on the results. With the weight function $\tilde{w}\{\tilde{\Delta}(Y)\} = |\tilde{\Delta}(Y)|^{-1/3}$, the threshold is effective only for those with $|\tilde{\Delta}(Y)| < (1/C_m)^3$. When $C_m = 300, 500$, and 1000, the thresholds are $3.7e-08, 8e-09, 1e-10$. The threshold might be too small to threshold observations with small $\tilde{\Delta}(Y)$. To a certain degree, if $\tilde{\Delta}(Y)$ converges to $\Delta(Y)$, misclassifying the subjects with small $\Delta(Y)$ should have a negligible effect on estimating θ. We wonder if the algorithm may spend too much efforts on these "non-

important" but easily misclassified examples. In other words, maybe more liberal thresholds can be explored. Alternatively we wonder if boosting algorithms that actually decrease the weight of repeatedly misclassified examples, for example, boost by majority (Schapire and Freund, 2012, Chapter 13) and BrownBoost Freund (2001) may be adopted here. In the meantime, overfitting can be an issue with possibly misspecified working models. This may be reflected by "over-confident" votes or very large $\tilde{\Delta}(Y)$. In such case, increasing the weights from these subjects can be helpful.

(3) Target function

The target function used in the algorithm is $1\{\Delta(Y) \leq 0\}$. The 0-1 loss function is non-continuous and non-convex. Similar to the usual boosting setting, we wonder if convex surrogate target functions can be used similar to (Hastie, Tibshirani, and Friedman, 2005, Chapter 10). In the original boosting setting, the surrogate functions were argued from statistical principles and lead to some nice algorithms including the well known functional gradient descent algorithm (Friedman, Hastie, and Tibshirani, 2000; Friedman, 2001). Another possible target function is actually $\theta\{\phi(Y)\}$. This function, or equivalently, $E\{P(T|Y)^{-1}1\{T = \phi(Y)\}D\}$, was used in Zhao et al. (2012) for individualized treatment selection based on a support vector machine framework. Different from $1\{\Delta(Y) \leq 0\}$, $E\{P(T|Y)^{-1}1\{T = \phi(Y)\}D\}$ is empirically estimable given $\phi(Y)$ when $P(T|Y)$ is known, for example, from clinical trial settings. A corresponding boosting algorithm can then be designed to maximize $E\{P(T|Y)^{-1}1\{T = \phi(Y)\}D\}$.

3. Improvement Using the Component-Wise Boosting

In a small simulation study we explored the sensitivity of the results to the working model. To reduce the undue influence of a multivariate working model, which is likely to be misspecified, we adopted BinomialBoosting with a component-wise generalized additive model as the "weak learner" (Bühlmann and Hothorn, 2007). More explicitly, BinomialBoosting iteratively approximates $2^{-1}\text{logit}P(D = 1|\cdot)$ by $\hat{f}^{[m]}(\cdot) = \nu\sum_{k=1}^{m}\hat{g}^{[k]}(\cdot)$, where ν is a step-length factor. At the mth iteration, $\hat{g}^{[m]}(\cdot)$ was given by the following procedure. Let $\boldsymbol{X} = (Y_1, \ldots, Y_p, T, T \times Y_1, \ldots, T \times Y_p)$ be a $2p + 1$ vector of biomarkers, treatment, and treatment-biomarker interactions. Let $X^{(j)}$ be the jth component of $\boldsymbol{X}$ for $1 \leq j \leq 2p + 1$. Denote $U_1^{[m-1]}, \ldots, U_n^{[m-1]}$ as the current negative gradients of the binomial loss function (c.f. equation (3.1) of Bühlmann and Hothorn (2007)) corresponding to n subjects in the training set. We fit $U_1^{[m-1]}, \ldots, U_n^{[m-1]}$ against $X_1^{(j)}, \ldots, X_n^{(j)}$ by a one-dimensional smoothing spline function $\hat{h}^{(j)}(\cdot)$. Then, $\hat{g}^{[m]}(\cdot)$ was chosen to be the best component-wise fitting,

$$\hat{g}^{[m]}(\cdot) = \hat{h}^{(\hat{j})}(x^{(\hat{j})}) \text{ where } \hat{j} = \underset{1\leq j\leq 2p+1}{\arg\min}\sum_{i=1}^{n}(U_i^{[m-1]} - \hat{h}^{(j)}(X_i^{(j)}))^2.$$

Therefore simple smoothing splines based on each (univariate) biomarker, treatment, and treatment-biomarker interactions formed a pool of candidate weak learners. At each iteration, we chose the weak learner as the simple smoothing spline that best predicted the current negative gradient of the binomial loss function.

The intuition behind our choice of the component-wise boosting is 2-fold. First, the flexibility of the smoothing splines may overcome possible misspecification of variable functional forms. Note that due to the ensemble nature of the boosting algorithm, we just need to get the contribution of each variable right. Second, fitting the gradients of the binomial loss function may reduce the undue influence of wrong working models on the weighting distributions. The component-wise boosting procedure is quite flexible and can also incorporate variable selection (Bühlmann and Hothorn, 2007). This may be attractive when the number of biomarkers is large.

We compared the above boosting method with K.J.H. methods through four models. Models 1 and 2 were the same as Scenarios 1 and 5 in K.J.H. Model 3 used the $\log(-\log)$ link function and had the complexity between Scenarios 3 and 4 in K.J.H. Model 4 had severe deviation from the logistic regression setting. For each model, we ran 100 simulations with $n = 1000$ in each simulation. We used the R package `mboost` to implement the above boosting algorithm. We chose each $\hat{h}^{(j)}(\cdot)$ to be a P-spline of three degrees of freedom with a B-spline basis by using the option `bbs` in the `mboost` package. We used the default choices of ν and number of iterations in the package. The performances from various methods are summarized in Table 1. Both MCR_{TB} and $\hat{\theta}$ were calculated based on testing data sets of size 10,000. Best performers are highlighted. One thing we noticed from the K.J.H. simulation results was that the performances of the boosting methods can depend on the working and true models. The misclassification rates (MCR_{TB}) were quite high when the true model was logistic and the boosting was based on classification trees (see Scenarios 1 and 2 in J.K.H.), or when the true model was not logistic and the boosting was based on logistic working models (see Scenarios 5 and 6 in J.K.H.). Without knowing the truth, it is therefore hard to determine which boosting algorithm to use. In comparison, the component-wise boosting seems to have improved all-around performance under various true models. A more comprehensive investigation is warranted.

References

Bühlmann, P. and Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting (with discussion). *Statistical Science* **22**, 477–505.

Freund, Y. (2001). An adaptive version of the boost by majority algorithm. *Machine Learning* **43**, 293–318.

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics* **29**, 1189–1232.

Friedman, J., Hastie, T., and Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *Annals of Statistics* **28**, 337–374.

Hastie, T., Tibshirani, R., and Friedman, J. (2005). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd edition. New York, NY: Springer.

Schapire, R. and Freund, Y. (2012). *Boosting Foundations and Algorithms*, 1st edition. Cambridge: The MIT Press.

Zhao, Y., Zeng, D., Rush, A., and Kosorok, M. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118.

Zou, H., Zhu, J., and Hastie, T. (2008). New multicategory boosting algorithms based on multicategory Fisher-consistent losses. *The Annals of Applied Statistics* **2**, 1290–1306.

Rejoinder

Chaeryon Kang,
Holly Janes,
and Ying Huang

We thank co-editor Jeremy M. G. Taylor for organizing this discussion and the discussants for their insightful comments and suggestions. In this rejoinder, we will address the broad points made by individual discussants and draw connections between them.

We agree with Laber, Tsiatis, Davidian, and Holloway (hereafter LTDH) that a taxonomy of methodology for deriving treatment rules is useful for discussing the relative merits of the various approaches. LTDH classified statistical approaches to find marker-based treatment rules into two classes based on the estimation method: "regression-based methods" obtain a rule by first modeling the outcome using a regression model; and "policy search methods" directly maximize a criterion of interest, for example, the expected outcome under marker-based treatment, in order to derive a treatment rule. Our boosting approach was characterized as a regression-based approach, whereas outcome weighted learning (OWL, Zhao et al., 2012), direct maximization of the expected outcome under marker-based treatment using the inverse probability weighted estimator (IPWE) and the augmented inverse probability weighted estimator (AIPWE) (Zhang et al., 2012a,b), and modeling marker-by-treatment interactions through Q- and A-learning (e.g., Murphy, 2003; Zhao, Kosorok, and Zeng, 2009) were characterized as policy search methods.

We prefer to group methods using somewhat different labels. We call "policy search methods" those that yield a treatment rule. In contrast, "outcome prediction methods" yield a model for the expected outcome given marker and treatment, which can then be used to derive a treatment rule. Using this terminology, our boosting approach, OWL, direct maximization of the expected outcome under marker-based treatment, and Q- and A-learning approaches are all examples of policy search methods: they yield treatment rules only and do not produce a model for the outcome. The methods differ in whether they are "direct" in that they search for treatment rules by directly maximizing a criterion of interest such as the expected outcome under marker-based treatment; or "indirect" in that they search for treatment rules by maximizing a criterion which is different from, but presumably related to the criterion of interest. Our boosting method is an indirect approach. The first method proposed by Tian, which minimizes the rate at which subjects are misclassified according to treatment benefit (using a surrogate variable for this unobserved outcome), is also an indirect policy search method. This taxonomy is helpful, we believe, in that it makes plain the fact that the approaches mentioned in our article and by the discussants are all policy search methods, except for the method suggested by Yu and Li (hereafter YL), which is an outcome modeling approach that is designed to be robust to model misspecification. They are therefore limited in that they are suitable only for addressing the problem of identifying a treatment rule, and not for the more difficult task of predicting outcome given marker value and treatment assignment.

Several discussants proposed novel direct policy search approaches that also use boosting ideas. Several rely on the fact that maximizing the expected outcome under marker-based treatment can be refomulated as a classification problem with weights that are functions of the outcome (Zhao et al., 2012; Zhang et al., 2012a,b). Using this formulation, Zhao and Kosorok (hereafter ZK) and Tian proposed solving an approximation of the weighted classification problem and applied AdaBoost to improve weak classifiers, while LTDH proposed "value boosting" that allows more general weights such as those from AIPWE. We agree that these methods have broad appeal and deserve in-depth investigation.

YL and Tian both raised questions about our proposed strategy of upweighting subjects with small estimated treatment effects, near the decision boundary, who are more likely to be incorrectly classified with respect to treatment benefit. They raised an interesting and fundamental question: should subjects who lie close to the decision boundary have more influence on the classifier? Or should subjects who lie far from the decision boundary but whose incorrect treatment recommendations will have greater impact have more influence? Many traditional classification methods have focused on subjects who are difficult to classify, for example, support vector machines and AdaBoost. In contrast, other recently developed boosting methods such as BrownBoost (Freund, 2001) focus on subjects whose estimated class labels are consistently correct across iterations, and give up on "noisy subjects" whose estimated class labels are consistently incorrect. We agree that, in the treatment selection context, boosting subjects whose estimated treatment effects are large is worth further investigation. We suspect that the optimal weighting

Table 1

Results of the simulation study for the continuous outcome setting. Marker combinations obtained using linear regression with maximum likelihood estimation (Linear MLE) and the boosting method described in our article with linear regression working model (Linear Boosting) are compared. Scenarios similar to 5 and 6 in our article are examined.[a] *For each scenario, 1000 training datasets ($n = 500$) and one test dataset ($N = 10^5$) were generated. Mean and Monte-carlo standard deviation (SD) of θ are shown, along with the mean and SD of the misclassification rate for treatment benefit (MCR_{TB}), calculated using test data.*

Scenario		True		Linear MLE	Linear Boosting
Scenario 5	θ	3.385	Mean	1.821	2.121
			SD	0.365	0.427
	MCR_{TB}		Mean	0.352	0.303
			SD	0.062	0.061
Scenario 6	θ	3.049	Mean	2.396	2.477
			SD	0.379	0.157
	MCR_{TB}		Mean	0.198	0.154
			SD	0.054	0.025

[a]In Scenario 5, the true risk model is $D = 0.1 + 0.2Y_1 - 0.2Y_2 + 0.1Y_3 - Y_1^2 + T(0.5 + Y_1 + 0.5Y_2 + 0.1Y3 - Y_1^2) + \epsilon$, and in Scenario 6, the true risk model is $D = -0.1 - 0.2Y_1 + 0.2Y_2 + Y_1Y_2 + T(-0.5 - Y_1 - 0.5Y_2 + 2Y_1Y_2) + \epsilon$, given T and Y, where Y_1, Y_2, and Y_3 are independent $N(0, 1)$, and ϵ follows $N(0, 1)$ and is independent of Y_1, Y_2, and Y_3. The same weight function $\tilde{w}\{\widehat{\Delta}(Y)\} = |\widehat{\Delta}(Y)|^{-1/3}$, maximum number of iterations ($M = 500$), and the maximum weight ($C_M = 500$) were used as in our article.

strategy will depend on the particular setting, and will be affected by factors such as the distribution of the markers and their associations with the treatment effect.

We agree with the point raised by ZK and Tian that the performance of our boosting approach depends on the choice of working model. In practice, prior biological knowledge and cross-validation techniques are useful for guiding the choice of working model. There appears to be similar sensitivity of the OWL method of ZK to the choice of "kernel" parameterizing the treatment rule boundary. Comparing the finite sample performance of the boosting and OWL methods to one another in simulations will be challenging, particularly under model misspecification, given that each requires specification of a different set of inputs.

One simple question raised by ZK is how to extend our boosting approach from the binary outcome setting, which is the focus of our article, to other types of outcomes such as continuous and count outcomes. The method extends naturally as illustrated in Table 1. We let $D \in \mathbb{R}^1$ be a continuous outcome in which a smaller value is preferable, T be treatment assignment ($T = 0/1$, where $T = 1$ is the default), and $Y \in \mathbb{R}^p$ be a set of markers. Denote by $\Delta(Y) = E(D|T = 0, Y) - E(D|T = 1, Y)$ the marker-specific treatment effect, $\phi(Y) = \mathbf{1}\{\Delta(Y) \leq 0\}$ the optimal treatment rule, and

$$\theta\{\phi(Y)\} = E(D|T = 1) - [E\{D|T = 1, \phi(Y) = 0\}P\{\phi(Y) = 0\} + E\{D|T = 0, \phi(Y) = 1\}P\{\phi(Y) = 1\}]$$

the primary measure of its performance. Using the boosting approach with linear regression working model can yield marker combinations with slightly higher θ and smaller misclassification of treatment benefit than using classical linear regression with maximum likelihood estimation. We note, however, that further investigation is needed to specify reasonable ranges for the tuning parameters of the boosting method in the continuous outcome setting.

We conclude with two observations that emerge from this discussion. First, there is much to be gained from bringing researchers from different areas of statistics and biostatistics together around a single topic. This discussion highlights the connections between the fields of adaptive treatment regimes and risk prediction and biomarker evaluation. Undoubtedly, it has brought relevant work in one field to the attention of researchers in another. With such interactions our science will surely improve. Second, there is tremendous value in reproducible research. We applaud the journal for encouraging us to publish our simulation code along with our article. With this code, the discussants were able to efficiently compare alternative approaches to ours using the same simulation scenarios, thus expediting the scientific process.

Acknowledgements

This work was funded by R01 CA152089, P30CA015704, and R01 GM106177-01.

References

Freund, Y. (2001). An adaptive version of the boost by majority algorithm. *Machine Learning* **43**, 293–318.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B* **65**, 331–355.

Zhang, B., Tsiatis, A., Laber, E., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68**, 1010–1018.

Zhang, B., Tsiatis, A. A., Davidian, M., Zhang, M., and Laber, E. (2012). Estimating optimal treatment regimes from a classification perspective. *Stat* **1**, 103–114.

Zhao, Y., Kosorok, M. R., and Zeng, D. (2009). Reinforcement learning design for cancer clinical trials. *Statistics in Medicine* **28**, 3294–3315.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 1106–1118.