Published in final edited form as:

Biometrics. 2014 September ; 70(3): 719–720. doi:10.1111/biom.12192.

Rejoinder: 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, WA 98109, U.S.A

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 (for example, Murphy (2003); Zhao et al. (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

*ckang2@fhcrc.org

**hjanes@fhcrc.org

***yhuang@fhcrc.org

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 reformulated as a classification problem with weights that are functions of the outcome (Zhao et al. (2012) and 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 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 $\phi(Y) = E(D|T = 0, Y) - E(D|T = 1, Y)$ the marker-specific treatment effect, $\phi(Y) = \mathbf{1}\{\phi(Y) > 0\}$ 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.

Acknowledgments

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

References

- Freund Y. An adaptive version of the boost by majority algorithm. *Machine learning*. 2001; 43:293–318.
- Murphy SA. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2003; 65:331–355.
- Zhang B, Tsiatis A, Laber E, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012; 68:1010–1018. [PubMed: 22550953]
- Zhang B, Tsiatis AA, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. *Stat*. 2012; 1:103–114. [PubMed: 23645940]
- Zhao Y, Kosorok MR, Zeng D. Reinforcement learning design for cancer clinical trials. 2009
- Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*. 2012; 107:1106–1118. [PubMed: 23630406]

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[†]. 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.

Table 1

Scenario	True	Linear MLE		Linear Boosting	
Scenario 5	θ	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	Mean	2.396	2.477	
		SD	0.379	0.157	
	MCR_{TB}	Mean	0.198	0.154	
		SD	0.054	0.025	

[†]In Scenario 5, the true risk model is $D = 0.1 + 0.2Y_1 - 0.2Y_2 + 0.1Y_3 - Y_1^2 + T(0.5 + Y_1 + 0.5Y_2 + 0.1Y_3 - Y_1^2) + \varepsilon$, and in Scenario 6, the true risk model is $D = -0.1 - 0.2Y_1 + 0.2Y_2 + Y_1Y_2 + T(-0.5 - Y_1 - 0.5Y_2 + 2Y_1Y_2) + \varepsilon$, 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 $w(\hat{Y}) = |\hat{Y}|^{-1/3}$, maximum number of iterations ($M = 500$), and the maximum weight ($CM = 500$) were used as in our article.