Optimal Treatment Recommendation via Subgroup Identification in Randomized Control Trials

Yang Zhao[1], Haoda Fu[2]
[1]Gilead Sciences, 333 Lakeside Dr., Foster City, CA 94404
[2]Eli Lilly and Company Corporate Center, Indianapolis, IN 46285

Abstract
In an era of rapid medical treatment development and with various options available to patients, personalized medicine has become an important topic to both researchers and practitioners. A new subgroup identification algorithm developed by Fu et al. (2016) provides individualized treatment recommendation under the outcome weighted learning framework. We here focus on its applications in randomized clinical trials to generate easy-to-interpret results. We applied this method to a dataset from a real clinical trial, and identified the optimal treatment recommendations for patient subgroups.

Key Words: multiple treatments; personalized medicine; randomized control trials; subgroup identification; value function; machine learning

1 Introduction

Personalized medicine, also termed theranostics or precision medicine, is a medical procedure that separates patients into different groups -- with medical decisions, practices, interventions and/or products being tailored to the individual patient based on their predicted response or risk of disease. In a clinical trial context, sometimes we may not be able to establish statistically the treatment benefit over control in the overall tested population, yet obvious treatment effect may be observed in a subgroup of patients. In this case, if we could identify the subgroup by evaluating patients' heterogeneous responses to the treatment, we may be able to learn more about the investigational drug, the underlying disease, and ultimately be able to develop recommendations of most efficacious or safer drugs for subgroups of patients.

There have been quite a few statistical methods developed in recent years for this purpose. For example, interaction trees detect subgroups by recursively partitioning the patient population based on treatment-by-covariate interactions (e.g., Su *et al.* (2009), Loh *et al.* (2015)); the *Virtual Twin* (VT) method first estimate differential treatment effect of each individual patient measured by a score function and then use these scores as responses to find subgroups (e.g., Foster *et al.* (2011)); and Lipkovich *et al.* (2011) proposed a subgroup identification algorithm based on differential effect search (SIDES). To identify an optimal treatment for a given patient rather than finding the "best patient" for a given treatment, the Individualized Treatment Recommendation (ITR) method aims to find an optimal treatment rule that maximizes a value function based on patient benefit evaluation (e.g. Qian *et al.* (2011), Zhao *et al.* (2012), Zhang *et al.* (2012)). Most of these ITR algorithms generate treatment rules as a linear combination of covariates, which may not easy to interpret in a clinical trial context. A new ITR algorithm developed by Fu *et*

al. (2016) under the outcome weighted learning framework provides simple rules where subgroups of an open rectangle shape can be identified. In this paper, we focus on its applications in randomized control clinical trials to generate easy-to-interpret results. We applied this method to a dataset from a real clinical trial, and identified the optimal treatment recommendations for patient subgroups.

This paper is organized as follows. In Section 2, we describe the ITR framework and introduce the searching algorithm proposed by Fu *et al.* (2016). Application of this method in a real clinical trial is discussed in Section 3. We conclude in Section 4.

2 The Individualized Treatment Recommendation (ITR) Method

2.1 The Framework

Assume we have a random sample of N subjects from a large population. T_i is the treatment assignment for patient i, where $i = 1, \cdots, N$. Y_i is the response variable, X_i is a vector of covariates, and $(Y, \ T, \ X)$ is the generic random variable of $\{(Y_i, \ T_i, \ X_i)\}$. Without loss of generality, we assume that larger Y corresponds to a better clinical outcome. Let $\mathcal{P}$ be the distribution of $(Y, \ T, \ X)$, E the expectation w.r.t. $\mathcal{P}$, $\mathcal{P}^r$ the distribution of $(Y, \ T, \ X)$ given $T = r(X)$, where $r(\cdot)$ is a rule defining a treatment recommendation for each individual in a population. The value function, which is the expected value of treatment benefit with respect to r, is defined as: $V(r) = E^r(Y)$. Simple derivation gives:

$$V(r) = E^r(Y) = \int Y d\mathcal{P}^r = \int Y \frac{d\mathcal{P}^r}{d\mathcal{P}} d\mathcal{P} = E\left\{\frac{I_{T=r(X)}}{p(T|X)} Y\right\} \quad (1)$$

The goal is to estimate r_0 such that

$$r_0 \in \arg\max_{r \in R} V(r), \quad (2)$$

where R is a collection of ways to assign treatments. Here r_0 is the optimal treatment regime that, if followed by the entire population of patients, would lead to the best outcomes on average.

Several advantages of using this framework can be observed from equation (1). First of all, there is no restriction on the variable Y, X, or T. That is, Y can be binary, continuous, or time-to-event data types; X can incorporate a variety of covariates, e.g., if X includes study ID, the framework can be used for meta-analysis; and T can handle multiple treatments. Notably, $p(T|X)$ explicitly allows the treatment assignments to depend on covariates, therefore, the framework can handle both randomized control trials (RCTs) and observational studies. Also, there is an objective function to evaluate different treatment assignments, which is a feature of the ITR method.

2.2 An ITR Algorithm for Simple Optimal Rule

Focusing on meeting the needs for drug development, Fu *et al.* (2016) proposed an ITR algorithm that directly searches for subgroups with an open rectangle shape (e.g., age <= 75 & body mass index > 18). These simple rules are often more desirable in clinical trial settings. Here, subgroups are defined by a certain number of covariates, called *depth*. In practice, only *depth* <= 3 is allowed for interpretability. The optimal simple rule can be obtained using a comprehensive searching algorithm, as follows:

- With observed data (Y, T, X), use a logistic regression to estimate the propensity scores $e = P(T = 1|X)$.
- Fit a linear model $Y \sim X$ to obtain the residuals $\tilde{Y}$. The data input to the algorithm are $(\tilde{Y}, T, e, X)$.

- Select any *depth* number of covariates (e.g., when *depth* = 3, X_{k1}, X_{k2}, X_{k3}) and for each selected covariate, choose a split value (c_{k1}, c_{k2}, c_{k3}).
- For each split value, we select a direction ($\leq$ or $>$) to define the subgroup. E.g., when *depth* = 3, one possible subgroup can be $A_{k,j} = (X_{k1} \leq c_{k1} \cap X_{k2} \leq c_{k2} \cap X_{k3} > c_{k3})$, where $j = 1, \cdots, 8$.
- Assign treatment to subgroup $A_{k,j}$ and control otherwise. Evaluate the value function $V_{k,j} = \frac{1}{N}\sum \frac{\tilde{Y}_i I_{T_i = I_{X_i \in A_{k,j}}}}{T_i e_i + (1-T_i)(1-e_i)}$.
- By selecting different covariates, split values, and directions, we evaluate all the value functions $V_{k,j}$ and provide a subgroup associated maximal value of $V_{k,j}$.

The operating properties of this algorithm have been evaluated via simulations on various aspects including speed test, numerical stability, convergence, variable importance, etc. (see Fu *et al.* (2016)).

3 Application To Clinical Trials

In this section, we apply the ITR method for simple rule to a real clinical trial dataset.

This is a phase 2, randomized, placebo-controlled clinical study. Most of the trial information is masked here to maintain confidentiality. In total about 200 subjects were randomized into the study with around 120 subjects in the treatment group and 80 on placebo control. The efficacy endpoint is the occurrence of a clinical event for disease worsening, with 0 indicating an event and 1 otherwise. The baseline covariates included demographics, other characteristics, lab markers, imaging markers and derived disease risk scores. There were in total 37 covariates used, of which 30 were continuous, 1 ordinal and 6 categorical. The data collected were of very high quality. With this many variables, we only had < 8% incomplete cases. Therefore, we used the approximately 180 complete cases (with both non-missing response and covariates) for this analysis.

For the searching algorithm, we used split values as the deciles based on the observed data, i.e., 10%, 20%, …, 90% quantiles, for each covariate. Using *depth* = 3, the variable importance ranking is plotted as in Figure 1.

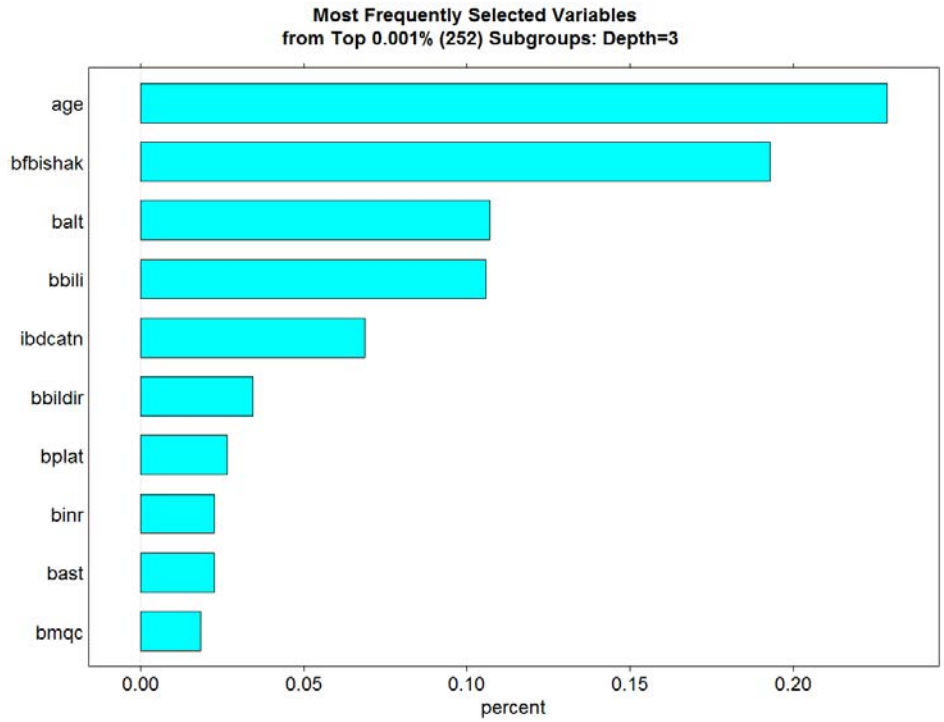

Figure 1: Variable Importance Ranking.

From the graph, we can see that from the second top variable to the 3rd, there is an obvious big drop in the variable selection frequency, which may indicate that a lower depth could be a more reasonable choice providing easier interpretation.

Using depth = 2, the ITR method selected a subgroup of patients with age <= 48 (60% quantile) and baseline Ishak fibrosis stage <= 4 (80% quantile). That is, patients in this subgroup are recommended to take the treatment, otherwise the control. This subgroup includes around 56% of the patients and non-subgroup 44%. In this example the top 2 variables from using depth = 3 happened to associate with the largest value function when using depth = 2. This may not always be the case, and if not, we may need to consider to compromise the best choice but to rely on clinical judgement to select the subgroup.

Figure 2 shows the treatment effect in the selected subgroup and non-subgroup in the original data. We can see that the patients in the non-subgroup might be better off if they were assigned to the control group.

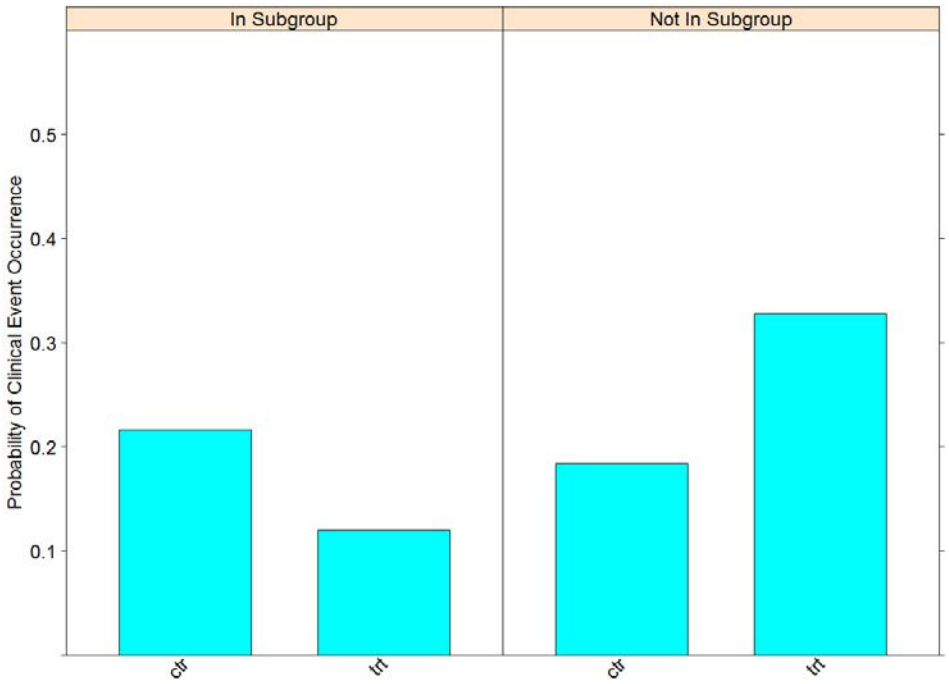

Figure 2: Probability of clinical event occurrence in the selected subgroup and non-subgroup.

To demonstrate the benefit of the ITR treatment allocation rule, in Table 1, we report the probability of clinical event occurrence before and after following the ITR rule.

Table 1: Probability of clinical event occurrence before and after following ITR

Original data		Following ITR		
0.204		0.185		
Control	Treatment		Control	Treatment
0.200	0.206	Subgroup	0.216	0.135
		Non-subgroup	0.310	0.328

If we were to follow the ITR optimal rule, that is, to assign treatment to the selected subgroup and control to the rest, then our estimate is that we would have approximately 18.5% event occurrence overall, as compared to the 20.4% in the original data. Furthermore, patients in both the subgroup and the non-subgroup would have a chance to improve their outcomes. For patients in the subgroup who received placebo control, they have a 21.6% event rate in the original data; if we follow the ITR allocation method, that is, to assign them to treatment, then they would have an estimated event rate of 13.5%. Similarly, for patients in the non-subgroup who received treatment, they have a 32.8% event rate in the original data, while under ITR assignment, their estimated event rate becomes approximately 31%.

The ability to come up with such an optimal rule is a unique feature of the ITR method. It provides us a useful tool to practice personalized medicine in subgroups of patients. And in doing this, we would be able to maximize the overall welfare of the patient population.

4 Conclusion

In this paper, we briefly reviewed the ITR framework and an algorithm that generates simple optimal rules in clinical trial settings. The algorithm is applied to a real clinical trial dataset. A subgroup of patients was identified and the benefit of following the ITR treatment rule was demonstrated. For future research, meta-analysis of multiple clinical trials may be considered, and angle-based classifiers may be utilized to accommodate high dimensional covariate space.

5 References

1. Su X, Tsai CL, Wang H, Nickerson DM, Li B. Subgroup analysis via recursive partitioning. *Journal of Machine Learning Research* 2009; 10: 141-158.
2. Loh WY, He X, Man M. A regression tree approach to identifying subgroups with differential treatment effects. *Statistics in Medicine* 2015; 34: 1818-1833.
3. Foster JC, Taylor JMC, Ruberg SJ. Subgroup identification from randomized clinical trial data. *Statistics in Medicine* 2011; 30: 2867-2880.
4. Lipkovich I, Dmitrienko A, Denne J, Enas G. Subgroup identification based on differential effect search (SIDES): a recursive partitioning method for establishing response to treatment in patient subpopulations. *Statistics in Medicine* 2011; 30: 2601-2621.
5. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *The Annals of Statistics* 2011; 39: 1180-1210.
6. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* 2012; 107(499): 1106-1118.
7. Zhang B, Tsiatis AA, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. *Statistic* 2012; 1(1): 103-114.
8. Fu H, Zhou J, Faries DE. Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies. *Statistics in Medicine* 2016; 35(19): 3285-3302.