Strategy-aware evaluation of treatment personalization

Félix Balazard, Gérard Biau, Philippe Ravaud, Raphaël Porcher

Strategy-aware evaluation of treatment personalization

Félix Balazard[1,2,3], Gérard Biau[1], Philippe Ravaud[4,5,6,7], and Raphaël Porcher*[4,5,6]

[1]Sorbonne Université, CNRS, LPSM, Paris, France
[2]Inserm, U1169, Kremlin-Bicêtre, Paris, France
[3]Owkin, Inc., New York, USA
[4]Université de Paris, CRESS, Paris, France
[5]Inserm, U1153, Paris, France
[6]Assistance Publique-Hôpitaux de Paris, Hôtel-Dieu, CEC, Paris, France
[7]Columbia University, Mailman School of Public Health, New York, USA

July 5, 2019

Abstract

Personalizing treatment according to patient's characteristics is at the core of stratified or precision medicine. There has been a recent surge of statistical methods aiming at identifying so-called optimal treatment strategies, i.e., strategies that assign a given treatment to a patient according to his/her characteristics. However, when data from a randomized controlled trial are used to estimate the optimal treatment strategy, it is not straightforward to estimate and test the benefit of the estimated strategy as compared to not personalizing treatment. In this context, we propose a principled approach for the estimation of the benefit of an estimated treatment strategy, accounting for its uncertainty. This leads to formalizing a strategy that we term the max lower bound strategy. Numerical simulations are used to show it allows proper type I error rate control and coverage probabilities. The approach is extended to multiple covariates using machine learning techniques. It is then applied to the data of a randomized trial in antineutrophil cytoplasmic antibody (ANCA)-associated vasculitis.
Keywords: Personalized medicine; Stratified medicine; Optimal treatment strategy; Machine learning.

1 Introduction

Personalized—or stratified—medicine consists in differentially treating patients based on their individual characteristics. If adapting the treatment to the patient is not a new idea in itself, it has attracted wide attention since the 2010s. Personalized medicine has been considered to have the potential to radically advance patient care by improving prevention and treatment efficacy while avoiding side effects (Hamburg and Collins, 2010). If this concept has successfully transformed the treatment of certain diseases, it has also been suggested that the hopes of personalized medicine

*E-mail: raphael.porcher@aphp.fr

were not fully matched by evidence (Khoury, 2010; Vivot et al., 2015). Of particular concerns are the emphasis on prognostic rather than predictive markers (Mandrekar and Sargent, 2009), and the widespread use of trials with an enrichment design (Freidlin and Korn, 2014; Vivot et al., 2016). Nonetheless, several methods have been recently developed to derive combinations of markers predicting improved response to treatment using data from randomized clinical trials (RCT, Cai et al., 2011; Zhao et al., 2012, 2013; Kang et al., 2014; Zhou et al., 2017), as well as observational studies (Qian and Murphy, 2011; Zhang et al., 2012; Shen and Cai, 2016; Athey and Wager, 2017; Künzel et al., 2019). One threat is however the risk of overfitting and associated treatment mistargeting (van Klaveren et al., 2019).

Let us consider a potential outcome framework, where it is assumed that each patient is associated with a vector (Y^0, Y^1) representing the outcome that would be observed under each treatment option so that the outcome is $Y = Y^0\mathbb{1}_{T=0} + Y^1\mathbb{1}_{T=1}$. Assuming that higher values of Y are beneficial, it would be natural to give treatment 1 to patients with $Y^1 \geq Y^0$ and treatment 0 to those with $Y^1 < Y^0$. Since both are never observed together, and cannot be known before administering the treatment, the approaches cited above mostly attempt to relate Y^0 and Y^1 to a set of covariates X representing the patient's characteristics. If we let $\Delta(X) = \mathbb{E}(Y^1|X) - \mathbb{E}(Y^0|X)$ and if this quantity is known, then giving treatment 1 to individuals with $\Delta(X) \geq 0$ and treatment 0 to individuals with $\Delta(X) < 0$ yields an optimal treatment strategy in that it maximizes the expectation of the outcome over the population (Zhang et al., 2012).

A treatment strategy—also termed treatment regime (e.g., Zhang et al., 2012), individualized treatment rule (e.g., Shen and Cai, 2016), or policy (e.g., Kang et al., 2014)—consists in formalizing a rule determining which treatment a patient should receive according to his/her covariates. To define the benefit of using a treatment strategy, one needs to compare the expectation of the outcome in the population under this treatment strategy to what would be obtained under the usual or reference treatment strategy (Janes et al., 2014). What should be the reference strategy is a complex issue, but we will simply consider here that there exists a treatment that is at some point viewed as the best treatment option for a given disease. Let us assume that one RCT compares a new treatment to this "old" one. If the new treatment is significantly superior to the old one in terms of average outcome, then the reference treatment strategy would be to now recommend treating all patients with the new treatment. In contrast, if the new treatment is not significantly better, then the reference treatment strategy is to treat all patients with the old treatment. In order to deal with both situations, we will simply refer throughout the manuscript to the reference treatment ($T = 1$), and the other treatment will be called the alternative treatment ($T = 0$). Later on, we will call the subset of patients for whom a treatment strategy recommends the alternative rather than the reference treatment the "personalized set".

In their comprehensive work on how to evaluate the performance of personalized treatment strategies, Janes et al. (2014) have proposed a plug-in estimator of the average gain under the optimal treatment strategy, and used bootstrap to obtain the corresponding percentile confidence intervals. They show that their estimators have good properties when an improvement is present (that is, there are patients who have better outcome under the alternative treatment than under the reference one) but warn the reader not to use their estimators to test for the presence of an improvement. While testing for presence of an improvement is not the focus of their work, they

suggest to use a composite test on the linear regression coefficients. In a previous work, Shuster and van Eys (1983) proposed to divide the range of X in regions of superiority of one treatment over the other, and a region of uncertainty where there is no significant difference between the treatment effects. This naturally allows to test for the benefit of personalizing treatment based on the covariates X.

The starting point of our paper is to note that, in practice, the optimal treatment strategy is not known, and therefore any actual personalization will be dependent on an estimated treatment strategy. Thus, our main objective is to propose a principled way to estimate the benefit of personalization of an estimated strategy. This allows us to study the choice of strategy and, accordingly, to design a new strategy with maximal guarantee on its gain. We also provide a statistical test for presence of improvement under the estimated strategy and show its good properties.

In Section 2, we adopt a Bayesian framework to estimate the quantities of interest and their credible quantiles. We start Section 3 by illustrating the problems posed by the naïve plug-in estimator of the optimal treatment strategy, and consider more general forms of strategies. We then propose in Section 4 a new treatment strategy, which we call the max lower bound strategy, and show that it defines a personalized set whenever presence of improvement under the optimal treatment strategy is detectable. Our approach is extended in Section 5 to other types of outcomes (binary and censored) and to the multivariate case, using machine learning techniques. Finally, we illustrate the procedure on a randomized clinical trial in Section 6.

2 Strategy-aware estimation

It is assumed throughout that each triplet patient/response/treatment is modeled by a random vector (X, Y, T), where X is a vector of covariates taking values in $\mathscr{X} \subset \mathbb{R}^d$ (the patient's characteristics), T is the treatment (alternative = 0, reference = 1), and Y is the patient's outcome. The latter depends on the treatment, and under the consistency assumption, we consider $Y = Y^0 \mathbb{1}_{T=0} + Y^1 \mathbb{1}_{T=1}$.

According to the definition given in the introduction, a treatment strategy can be represented by a function $\texttt{pol} : \mathscr{X} \to \{0, 1\}$, which assigns a possible treatment 0 or 1 to each possible realization x of X. For instance, the strategy consisting in giving the reference treatment to everyone could be termed $\texttt{ref} : x \mapsto 1$. In the context of personalized medicine, we are looking for a more complex strategy, which takes into account each patient's characteristics. In particular, the optimal treatment strategy introduced earlier can therefore be written $\texttt{opt}(x) = \mathbb{1}_{\Delta(x) \geq 0}$, where $\Delta(x) = \mathbb{E}(Y^1 | X = x) - \mathbb{E}(Y^0 | X = x)$. A general strategy (or policy), say $\texttt{pol}$, can be characterized by its average outcome, $\mathbb{E}Y^{\texttt{pol}}$, as well as the gain in average outcome as compared to using the reference strategy $\texttt{ref}$, that is:

$$\Theta(\texttt{pol}) \stackrel{\text{def}}{=} \mathbb{E}Y^{\texttt{pol}} - \mathbb{E}Y^1 = \mathbb{E}_X[-\Delta(X)\mathbb{1}_{\texttt{pol}(X)=0}],$$

where $\mathbb{E}_X$ is the expectation under $\mathbb{P}_X$, the probability distribution of X.

Taking the optimal strategy $\texttt{opt}$ gives

$$\Theta(\texttt{opt}) = \mathbb{E}_X[-\Delta(X)\mathbb{1}_{\Delta(X)<0}], \tag{1}$$

the quantity used in Janes et al. (2014), who have further proposed additional measures for the benefit of personalization, such as the proportion of marker-negative (or positive) patients $P_{\text{neg}} = \mathbb{P}(\Delta(X) < 0)$ and the average benefits of no treatment among marker-negatives $B_{\text{neg}} = \mathbb{E}[-\Delta(X)|\Delta(X) < 0]$, so that $\Theta(\texttt{opt}) = P_{\text{neg}}B_{\text{neg}}$. We note ithat $\Theta(\texttt{opt}) \geq 0$, and that $P_{\text{neg}} > 0$ is a necessary condition to have $\Theta(\texttt{opt}) > 0$. Of course, for a general strategy `pol`, $\Theta(\texttt{pol})$ can have an arbitrary sign.

In practice, the distribution of (X, Y, T) is unknown, and so is the optimal strategy `opt`. We consider an i.i.d. sample corresponding to RCT data $\mathscr{D}_n = (X_i, Y_i, T_i)$, $1 \leq i \leq n$, where each triplet (X_i, Y_i, T_i) is distributed as the generic (X, Y, T). So, for each i, X_i represents the patient's characteristics that we want to use to personalize treatment, Y_i is the observed continuous outcome, and $T_i \in \{0, 1\}$ is the treatment allocated in the trial. We consider the potential outcome framework introduced earlier, i.e., $Y_i = Y_i^0$ if $T_i = 0$ and $Y_i = Y_i^1$ if $T_i = 1$, and assume throughout that X_i is independent of T_i (randomization).

In this section, and until Subsection 5.2, we make the assumption that X is a bounded scalar (that is, $d = 1$, with $\mathscr{X} = \text{support}(\mathbb{P}_X) = [x_0, x_1]$, $-\infty < x_0 < x_1 < \infty$), and that data arise from a linear model. The linear model assumes that Y^t, $t = 0, 1$, can be expressed as a sum of an intercept β_0, a prognostic term with coefficient β_1, an average treatment effect term with coefficient β_2, an interaction term between treatment and the covariate X with coefficient β_3, and some independent Gaussian noise ε with variance $\sigma^2 > 0$:

$$Y^t = \beta_0 + \beta_1 X + \beta_2 t + \beta_3 X t + \varepsilon, \quad t = 0, 1. \tag{2}$$

We will suppose, without loss of generality, that X is centered ($\mathbb{E}_X X = 0$), so that β_2 represents the average treatment effect, $\mathbb{E}Y^1 - \mathbb{E}Y^0$. Note that this implies that $0 \in [x_0, x_1]$. It is also assumed that the variance σ^2 is known for simplicity. If we do not make this assumption, we have to use the t-distribution instead of the Gaussian distribution, which adds complication for no additional insights.

For each patient i with covariate X_i, the difference in outcome between the reference treatment and the alternative treatment is simply

$$\Delta(X_i) = \mathbb{E}[Y_i^1|X_i] - \mathbb{E}[Y_i^0|X_i] = \beta_2 + \beta_3 X_i.$$

If the coefficients were known exactly, then we would choose to give the reference treatment 1 to all patients with $\Delta(X_i) \geq 0$ (favoring the treatment with the best outcome on average when $\Delta(X_i) = 0$ but the other choice would be equivalent) and give the alternative treatment 0 to patients with $\Delta(X_i) < 0$. The average gain between this perfect optimal treatment strategy and the reference strategy is then the one given in equation (1). We note that if $\beta_3 = 0$, i.e., if there is no interaction, then $\Delta(X_i)$ is constant, equal to the average treatment effect β_2. We also see that if $\beta_2 < 0$, then there are always some patients who would benefit from personalization as $\Delta(0) = \beta_2 < 0$ and $0 \in [x_0, x_1]$. On the other hand, when $\beta_2 \geq 0$, the presence of an interaction, i.e., $\beta_3 \neq 0$, is not sufficient for the existence of patients benefiting the alternative treatment. Indeed, if $\beta_2 \geq 0$, the existence of $x \in [x_0, x_1]$ such that $\Delta(x) < 0$ is equivalent to $-\beta_2/\beta_3 > x_0$ when $\beta_3 > 0$, and to $-\beta_2/\beta_3 < x_1$ when $\beta_3 < 0$.

In practice, we do not have access to the true values of β_2 and β_3, and have to rely on some estimation procedure. This implies that there will be uncertainty in our estimation, and therefore

uncertainty in the associated strategy. An important contribution of the present paper is to estimate the average gain under the estimated strategy and not the perfect, but unknown, strategy $\texttt{opt}$.

Let us denote by $\hat{\beta} \overset{\text{def}}{=} (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ the standard least square estimators of $\beta \overset{\text{def}}{=} (\beta_0, \beta_1, \beta_2, \beta_3)$. We let $\mathbb{P}_\beta$ be the sampling probability of ϵ or, in other words, the sampling probability of Y given X and T. Under $\mathbb{P}_\beta$, we have $\hat{\beta} \sim \mathcal{N}(\beta, \Sigma)$, where $\Sigma = \sigma^2(Z^\top Z)^{-1}$ and Z is the design matrix

$$Z = \begin{bmatrix} 1 & X_1 & T_1 & X_1T_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_n & T_n & X_nT_n \end{bmatrix}.$$

While $\hat{\beta}_2$ is the estimated effect of the reference treatment versus the alternative treatment, this does not warrant that $\hat{\beta}_2 \geq 0$. As we discussed in the introduction, we could have $\hat{\beta}_2 \leq 0$ if the new treatment improved outcomes compared to the old treatment but not significantly. The reference treatment is then set to the old treatment, and the new treatment is the alternative one.

The presence of an interaction is a necessary condition for improvement due to personalization, and the associated test will be ubiquitous throughout. In order to make the discussion more transparent, we formally define this test below. Notation $\Sigma_{i,j}$ means the element in the i, j position of the covariance matrix Σ, and q_α is the α-quantile of the standard normal distribution.

Definition 1 (Interaction test)**.** *To test against the null hypothesis $H_0^{\text{interact}} : \beta_3 = 0$, the test statistic is $\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$. Its distribution is $\mathcal{N}(0,1)$ under $\mathbb{P}_\beta$. When the alternative hypothesis is $H_1^{\text{interact}} : \beta_3 \neq 0$, the rejection region is $\{|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} > q_{1-\alpha/2}\}$. If the alternative hypothesis is $H_1^{\text{interact}} : \beta_3 > 0$ (respectively, $H_1^{\text{interact}} : \beta_3 < 0$), the rejection region is $\{\hat{\beta}_3/\sqrt{\Sigma_{3,3}} > q_{1-\alpha}\}$ (respectively, $\{\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha\}$).*

Now that we have discussed the interpretation of the coefficients, we can turn to the problem at hand. The natural estimator of $\Delta(x)$ is simply $\hat{\Delta}(x) = \hat{\beta}_2 + \hat{\beta}_3 x$, and an example of estimated strategy is the plug-in estimator of the optimal strategy, say $\hat{\texttt{pol}}_1$, sometimes called the optimal treatment regime (OTR, Brinkley et al., 2010; Zhang et al., 2012; Janes et al., 2014). It is defined by $\hat{\texttt{pol}}_1(x) = \mathbb{1}_{\hat{\Delta}(x) \geq 0}$, and recommends to each patient the treatment maximizing the predicted outcome given the patient's characteristics X. For a general estimator $\hat{\texttt{pol}}(x)$, eventually different from $\hat{\texttt{pol}}_1(x)$, the associated improvement in population averaged outcome is

$$\Theta(\hat{\texttt{pol}}) \overset{\text{def}}{=} \mathbb{E}_X[-\Delta(X)\mathbb{1}_{\hat{\texttt{pol}}(X)=0}],$$

which is unknown and has to be estimated. More importantly, we should also provide a lower confidence bound for this quantity at a predefined level. We will assume that $\mathbb{P}_X$ is known throughout our theoretical discussions, but in practice, we use $\mathbb{P}_n$, the empirical distribution of $X_1, \ldots, X_n$. The distribution $\mathbb{P}_X$ used in the estimation of $\Theta(\hat{\texttt{pol}})$ could be different from the one generating the sample if, for example, we want to estimate the average improvement in a population with a different distribution than the one of the RCT.

We insist that $\Theta(\hat{\texttt{pol}})$ is a random variable, because it depends upon $\hat{\texttt{pol}}$. This dependency complicates the study of $\Theta(\hat{\texttt{pol}})$ under $\mathbb{P}_\beta$. To circumvent this complication, we have chosen to adopt a Bayesian estimation point of view and, conditionally on the sample $\mathscr{D}_n$, put a distribution on the regression parameter β. This allows to separate the uncertainty on β (random under the

posterior distribution) from the uncertainty on the strategy (fixed under the posterior distribution). We will however be interested in the frequentist properties of our estimation, and therefore the use of a prior would be counter-productive. In other words, we use a constant prior on $\mathbb{R}^4$ for β. Therefore, the posterior distribution of β given the sample $\mathcal{D}_n$ is simply $\Pi \overset{\text{def}}{=} \mathcal{N}(\hat{\beta}, \Sigma)$, where Σ is the covariance matrix defined above.

To distinguish the true fixed β, we denote by $\beta^\star \overset{\text{def}}{=} (\beta_1^\star, \beta_2^\star, \beta_3^\star, \beta_4^\star)$ a random variable whose distribution is Π. Similarly, we add a $\star$ superscript on the quantities depending on $\beta^\star$ instead of β. Thus, we set

$$\Theta^\star(\hat{\texttt{pol}}) = \mathbb{E}_X[-\Delta^\star(X)\mathbb{1}_{\hat{\texttt{pol}}(X)=0}],$$

where $\Delta^\star(X) = \beta_2^\star + \beta_3^\star X$. This quantity is easy to study, as shown in the next proposition (proof in the Appendix A1). We let

$$\hat{\Theta}(\hat{\texttt{pol}}) = \mathbb{E}_X[-\hat{\Delta}(X)\mathbb{1}_{\hat{\texttt{pol}}(X)=0}].$$

Proposition 1. *Under* Π, $\Theta^\star(\hat{pol})$ *follows a Gaussian distribution with* $\mathbb{E}_\Pi[\Theta^\star(\hat{pol})] = \hat{\Theta}(\hat{pol})$ *and*

$$\mathrm{Var}_\Pi[\Theta^\star(\hat{pol})] = \mathbb{E}_\Pi\big[\mathbb{E}_X^2[(\Delta^\star(X) - \hat{\Delta}(X))\mathbb{1}_{\hat{pol}(X)=0}]\big].$$

The credible α-quantile of $\Theta(\hat{\texttt{pol}})$—that is, the α-quantile of the posterior distribution—is then

$$\hat{q}_{n,\alpha}(\hat{\texttt{pol}}) = \hat{\Theta}(\hat{\texttt{pol}}) + q_\alpha \mathrm{sd}_\Pi(\Theta^\star(\hat{\texttt{pol}})), \tag{3}$$

where sd denotes the standard deviation. This quantity will be studied thoroughly in the next section, and we will see later that these credible quantiles are in fact confidence bounds with valid frequentist coverage for $\Theta(\hat{\texttt{pol}})$ under $\mathbb{P}_\beta$. In practice, the use of formula (3) requires to compute the standard deviation. It is however simpler to sample $\beta^\star$ from Π J times and compute $\Theta^\star(\hat{\texttt{pol}})$ each time, which gives us an empirical distribution $\{\Theta_j^\star(\hat{\texttt{pol}})\}_{1\le j\le J}$, from which we can then retrieve the quantiles of the posterior distribution of $\Theta(\hat{\texttt{pol}})$.

In addition to estimating the benefit of personalization, it is essential in clinical studies to be able to test the null hypothesis of no gain from personalization, that is, $H_0 : \Theta(\hat{\texttt{pol}}) \le 0$ against the alternative hypothesis of gain from personalization $H_1 : \Theta(\hat{\texttt{pol}}) > 0$. This can naturally be done using our credible α-quantile. Indeed, the natural rejection region of the null hypothesis at level α is simply $\{\hat{q}_{n,\alpha}(\hat{\texttt{pol}}) > 0\}$, as this means that the posterior distribution of $\Theta(\hat{\texttt{pol}})$ has more than $1-\alpha$ of its weight on the positive line. It is important to note that this test is not a proper test because our null hypothesis depends on a random variable instead of a fixed quantity. This means that even when there is an improvement to be found, i.e., $\Theta(\texttt{opt}) > 0$, the null hypothesis $H_0 : \Theta(\hat{\texttt{pol}}) \le 0$ can occur, either because of poor estimation of the strategy or because the estimated strategy defines an empty personalized set. This issue will depend on the choice of the strategy, as illustrated in the next section. However, while this is not a proper test, it is precisely the decision we care about. In fact, we will show in Section 4 that this test needs to be combined with the test for interaction presented in Definition 1. Furthermore, we will see that for the right choice of strategy, the resulting joint test is a valid test for presence of improvement under the optimal strategy.

We close this section by emphasizing that we have three probability measures of interest: $\mathbb{P}_X$ (the distribution of X under which we want to compute the expected gains), $\mathbb{P}_\beta$ (the frequentist

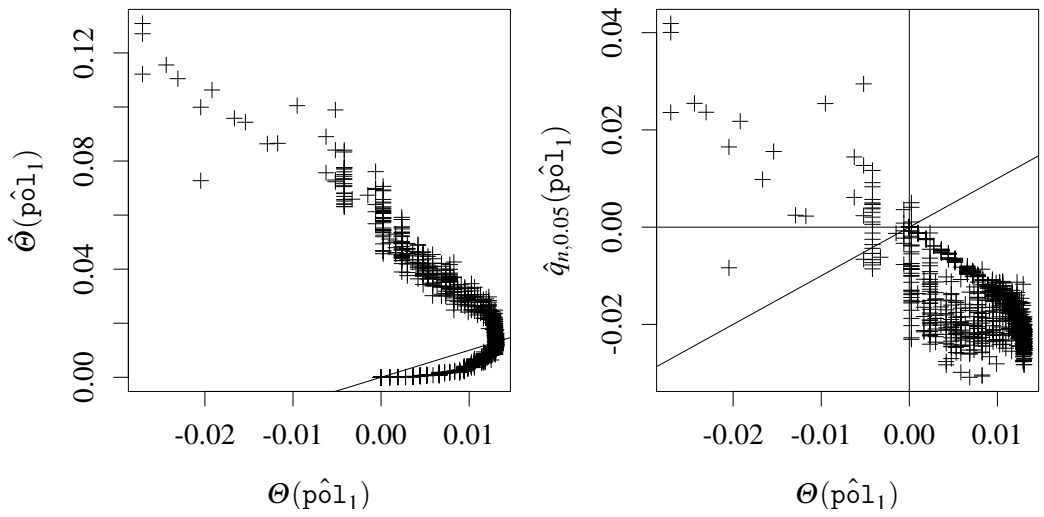

Figure 1: Graph of $\Theta(\hat{\mathtt{pol}}_1)$ against $\hat{\Theta}(\hat{\mathtt{pol}}_1)$ and $\hat{q}_{n,0.05}(\hat{\mathtt{pol}}_1)$. We simulated 10 000 datasets based on the linear model (2), i.e., we sampled ϵ for each simulation. The variable X is sampled once from a uniform distribution between -1 and 1. We take $n = 300$, $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 1.3$, and $\sigma^2 = 1$. Under these simulation settings, optimal personalization affects 10% of the patients, $P_{\text{neg}} = 0.1$, and $\Theta(\mathtt{opt}) = 0.013$. We sampled from Π 10 000 times in order to compute the quantiles. To enhance readability, we have plotted only 1 000 points. The line $y = x$ is also plotted.

probability of ϵ) and Π (the posterior probability of β given $\mathcal{D}_n$). The subscript β in the second one underlines the fact that β is a constant under this probability, whereas it is random under Π. To avoid confusion, we introduced a $\star$ superscript for the random variables under Π.

3 Choice of a treatment strategy

3.1 Problems of estimation under $\hat{\mathtt{pol}}_1$

Our first example of strategy was the plug-in estimator of the optimal treatment strategy, that is $\hat{\mathtt{pol}}_1(x) = \mathbb{1}_{\hat{\Delta}(x)\geq 0}$. We show here that, despite its simplicity, this policy behaves poorly for estimating the benefit of personalization. In particular, it can be overall detrimental even though our estimation predicts that it will be beneficial, i.e., $\mathbb{P}_\beta(\Theta(\hat{\mathtt{pol}}_1) < 0 | \hat{q}_{n,\alpha}(\hat{\mathtt{pol}}_1) > 0)$ is large.

We illustrate this using simulations based on a linear model of the form (2), whose specifics are described in the legend of Figure 1. Briefly, under our simulation settings, 10% of patients should benefit more from the alternative than from the reference treatment (i.e., $P_{\text{neg}} = 0.1$), and under the optimal treatment strategy, the average outcome is improved by 0.013 (i.e., $\Theta(\mathtt{opt}) = 0.013$). We consider the properties of the quantile $\hat{q}_{n,0.05}(\hat{\mathtt{pol}}_1)$, which can be used to provide a test, and the mean of the distribution, $\hat{\Theta}(\hat{\mathtt{pol}}_1)$, which provides an estimation of the benefit of personalizing treatment. As the true parameters of the linear model are known, we can compute the true value of $\Theta(\hat{\mathtt{pol}}_1)$.

The quantile $\hat{q}_{n,0.05}(\hat{\mathtt{pol}}_1)$ is positive, i.e., we are detecting the presence of a benefit of personalization, in 4.7% of the simulations. This is suboptimal as with a different strategy detailed later on, we will be able to detect a benefit of personalization in 34% of cases.

More worrying yet is the behavior of the true parameter $\Theta(\hat{\texttt{pol}}_1)$ against its estimator $\hat{\Theta}(\hat{\texttt{pol}}_1)$. This is shown in the upper panel of Figure 1. It should be noted that most points are close to the optimal $(0.013, 0.013)$, where the strategy is well estimated and the estimated gain is close to its real value. However, in a certain number of simulations the strategy is poorly estimated, and as a consequence the real gain is negative while the estimated gain appears very large. These situations are problematic because a decision based on such estimations will negatively impact health while being claimed to have a large positive impact.

Manual inspection of the problematic points shows that this behavior happens when the ratio $\hat{\beta}_2/\hat{\beta}_3$ is underestimating β_2/β_3. As $x = -\hat{\beta}_2/\hat{\beta}_3$ is the value at which $\hat{\Delta}$ changes sign, this means that $\hat{\texttt{pol}}_1$ will attribute the alternative treatment to too many participants compared to the optimal treatment strategy, and therefore many participants who would have benefited from the reference treatment will receive the alternative treatment. At the same time, since the posterior distribution we use is centered on poorly estimated coefficients, the estimation of benefit will be optimistic. The inverse situation can also be seen under the first bisector: when the ratio $\hat{\beta}_2/\hat{\beta}_3$ is overestimated, not enough participants will be attributed the alternative treatment, leading to a decreased population averaged outcome as compared to what could have been obtained by using the optimum.

The consequence of these problematic situations is even more poignant when we consider the lower quantile, i.e., the one we use to test the presence of an improvement. As we can see in the lower panel of Figure 1, $\hat{q}_{n,0.05}(\hat{\texttt{pol}}_1)$ seems to be positive mostly when $\Theta(\hat{\texttt{pol}}_1)$ is negative. Although $\hat{q}_{n,0.05}(\hat{\texttt{pol}}_1)$ is positive 4.7% of the time, in 90% of those cases $\Theta(\hat{\texttt{pol}}_1)$ is negative. This is the same phenomenon as above: an underestimated ratio leads to a large personalized set and at the same time overestimated bounds on the gain. While the quantile has the right coverage probability, the mistakes it makes are the one we care the most about: claiming a large improvement when personalizing treatment would be detrimental on average.

Fortunately, all these issues can be dealt with by considering a different strategy. The key idea is that if we can identify the patients that bring the most uncertainty to our estimation, then we can choose to exclude them from the personalization. This means that in case of uncertainty, we prefer to give the reference treatment. This asymmetry is analogous to the asymmetry created by defining a hypothesis as the null hypothesis in statistical test theory. This asymmetry is present in our quantity of interest $\Theta(\hat{\texttt{pol}})$ as no uncertainty comes from the patients receiving the reference treatment. Furthermore, this asymmetry is desirable, as a clinician implementing a personalization strategy will want to make sure that the change will be beneficial even if some patients who would have benefited from the alternative treatment are missed by this strategy. As a consequence, the strategies that we will recommend here are different from the ones where the global health impact of a strategy (e.g Qiu et al. (2018)) is considered and not its relative impact compared to a reference treatment.

In order to define a better strategy, we need to understand the uncertainty that patients bring to our aggregated quantities. To reach this goal, it is necessary to make a detour through an analysis of uncertainty at the individual level. This is the topic of the next subsection.

3.2 Uncertainty at the individual level

A patient with covariate $X = x$ has an estimated improvement of $\hat{\Delta}(x) = \hat{\beta}_2 + \hat{\beta}_3 x$. The distribution of improvement under Π is

$$\Delta^\star(x) = \beta_2^\star + \beta_3^\star x \sim \mathcal{N}\big(\hat{\Delta}(x), \mathrm{Var}_\Pi(\Delta^\star(x))\big),$$

where

$$\mathrm{Var}_\Pi(\Delta^\star(x)) = \Sigma_{2,2} + 2\Sigma_{2,3}x + \Sigma_{3,3}x^2.$$

The certainty with which such a patient will benefit from personalization is naturally measured by the quantity $\Pi(\Delta^\star(x) < 0)$. As $\Delta^\star(x)$ is Gaussian, this probability depends only on

$$z_\Delta(x) \stackrel{\text{def}}{=} \frac{\hat{\Delta}(x)}{\mathrm{sd}_\Pi(\Delta^\star(x))} = \frac{\hat{\beta}_2 + \hat{\beta}_3 x}{\sqrt{\Sigma_{2,2} + 2\Sigma_{2,3}x + \Sigma_{3,3}x^2}},$$

as we have $\Pi(\Delta^\star(x) < 0) = \Phi(-z_\Delta(x))$, with Φ the cumulative distribution function of the standard normal distribution.

In this context, we can derive a test for personalization at the individual level. Namely, we consider the null hypothesis $H_0 : \Delta(x) \geq 0$, i.e., the reference treatment is better than the alternative for a patient with covariate x. The alternative hypothesis is then $H_1 : \Delta(x) < 0$, and the rejection region of this test is $\{z_\Delta(x) < q_\alpha\}$. This rejection region is the region of superiority for the alternative treatment and was already defined in Shuster and van Eys (1983). Observe that, for each individual, $\Phi(z_\Delta(x)) = \Pi(\Delta^\star(x) \geq 0)$ is the level at which we would reject the null.

This state of affairs suggests a new, natural, strategy, which personalizes treatment for patients in the region of superiority of the alternative treatment, i.e., patients such that $z_\Delta(x) < q_\alpha$. We call this strategy the individual strategy $\hat{\texttt{ind}}_\alpha$, and notice that it depends on α, the chosen confidence level. Thus, we have

$$\hat{\texttt{ind}}_\alpha : x \mapsto \mathbb{1}_{z_\Delta(x) \geq q_\alpha}. \tag{4}$$

The fundamental difference with the naïve plug-in $\hat{\texttt{pol}}_1$ strategy is that patients with $q_\alpha \leq z_\Delta(x) < 0$, who are predicted to benefit from the alternative treatment under the strategy $\hat{\texttt{pol}}_1$, would still receive the reference one under $\hat{\texttt{ind}}_\alpha$. The function $x \mapsto z_\Delta(x)$ is therefore sufficient to quantify the uncertainty with which a patient would benefit from the alternative treatment, and we need to study it in depth. It is important to keep in mind that we are mainly interested by the behavior of z_Δ when it is negative, as this corresponds to patients who are expected to benefit from personalization. The study of z_Δ variations is detailed in the Appendix B. Results are summarized in the figure 2 and proposition 2.

Proposition 2. *Assume, for simplicity, that $\Sigma_{2,3} = 0$. Then the function z_Δ defines a bijection to $(-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}}, 0]$. The domain of the bijection is*

$$\begin{cases} (-\infty, -\hat{\beta}_2/\hat{\beta}_3] & \text{if } \hat{\beta}_2 > 0 \text{ and } \hat{\beta}_3 > 0, \\ [-\hat{\beta}_2/\hat{\beta}_3, +\infty) & \text{if } \hat{\beta}_2 > 0 \text{ and } \hat{\beta}_3 < 0, \\ (C, -\hat{\beta}_2/\hat{\beta}_3] & \text{if } \hat{\beta}_2 < 0 \text{ and } \hat{\beta}_3 > 0, \text{ with } z_\Delta(C) = -|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}}, \\ [-\hat{\beta}_2/\hat{\beta}_3, C) & \text{if } \hat{\beta}_2 < 0 \text{ and } \hat{\beta}_3 < 0. \end{cases}$$

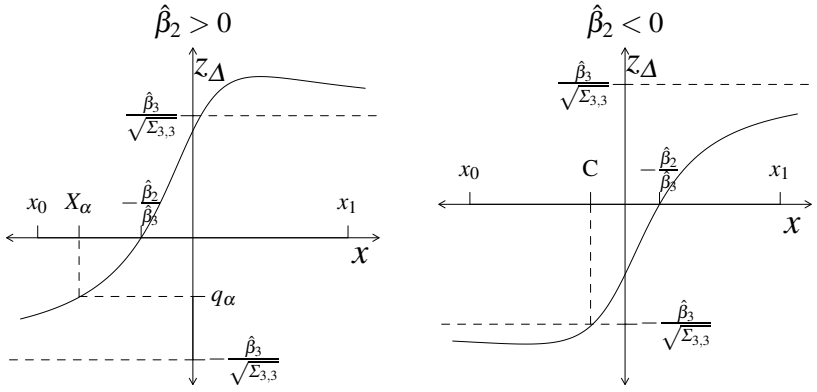

Figure 2: Example of variation of z_Δ depending on the sign of $\hat{\beta}_2$. The interaction coefficient $\hat{\beta}_3$ is positive in both plots. If it were negative, the curve would the symmetric of the plotted curve with respect to the ordinate axis.

Thus, for all $\alpha < 0.5$ such that $q_\alpha > -|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}}$, we can define $X_\alpha \overset{\text{def}}{=} z_\Delta^{-1}(q_\alpha)$, as shown in Figure 2. We have capitalized the X to underline that this is a random variable under $\mathbb{P}_\beta$. This allows to rewrite the rejection region to test superiority of the alternative treatment for a patient with covariate x as $\{z_\Delta(x) < q_\alpha\} = \{x < X_\alpha\}$ if $\hat{\beta}_3 > 0$ and $\{z_\Delta(x) < q_\alpha\} = \{x > X_\alpha\}$ if $\hat{\beta}_3 < 0$. The point of Proposition 2 is to be able to parametrize on the scale of x instead of the scale of z_Δ—this new parameterization will play a key role in the next subsection.

Our detour through the analysis of the uncertainty at the individual level is now over. On the way, we have found another strategy of interest, $\hat{\texttt{ind}}_\alpha$, and a function sufficient to quantify the uncertainty, z_Δ. We are now prepared to study the uncertainty at the aggregated level, and use it to define a strategy with maximal guarantee on its gain.

4 Max lower bound strategy

4.1 Definition

Our goal is to find a strategy that defines a non-empty personalized set as often as possible, while having the most confidence that personalization will be beneficial. To this aim, we first need to restrict the range of possible strategies we will be looking at. The two strategies we have seen up to now, $\hat{\texttt{pol}}_1$ and $\hat{\texttt{ind}}_\alpha$, are both of the form $x \mapsto \mathbb{1}_{z_\Delta(x) \geq \eta}$, with $\eta = 0$ for $\hat{\texttt{pol}}_1$ and $\eta = q_\alpha$ for $\hat{\texttt{ind}}_\alpha$. We have also seen that z_Δ is sufficient to quantify uncertainty at the individual level. It is therefore reasonable to limit our search to strategies of the same form as our two previously defined strategies.

Conditionally on $\mathcal{D}_n$, z_Δ is a fixed function. Assume for now, without loss of generality, that $\hat{\beta}_3 > 0$, and, to fix ideas, that x_0 (the left extremity of the support of X) is in the domain of the bijection defined in Proposition 2. In this case, we can parametrize our class of policies on the scale of x instead of the scale of z_Δ. This suggests to look for our strategy in functions of the form $x \mapsto \mathbb{1}_{x \geq \rho}$, with $\rho \in [x_0, -\hat{\beta}_2/\hat{\beta}_3]$. (Of course, in the opposite situation where $\hat{\beta}_3 < 0$, the corresponding form is $x \mapsto \mathbb{1}_{x \geq \rho}$ for $\rho \in [-\hat{\beta}_2/\hat{\beta}_3, x_1]$.)

In order to achieve our goal, we propose to maximize the test statistic for positive impact of personalization, i.e., $\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})$, over all possible choices of $\rho \in [x_0, -\hat{\beta}_2/\hat{\beta}_3]$. More precisely, let $A_\alpha = \arg\max_\rho \hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})$ be the set of maximizers. To define the strategy unambiguously, we choose the largest such threshold, i.e., we set

$$\rho_{\max,\alpha} = \max A_\alpha.$$

We call the resulting strategy the max lower bound strategy and denote it by $\hat{\texttt{mlb}}_\alpha$. Thus, by definition,

$$\hat{\texttt{mlb}}_\alpha(x) = \mathbb{1}_{x\geq\rho_{\max,\alpha}}.$$

To study the behavior of the random variable $\rho_{\max,\alpha}$, we simply use equation (3), which in this context takes the form

$$\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho}) = \hat{\Theta}(\mathbb{1}_{x\geq\rho}) + q_\alpha \mathrm{sd}_\Pi(\Theta^\star(\mathbb{1}_{x\geq\rho})).$$

As we want to maximize this quantity, we are going to derivate with respect to ρ. Assume, to simplify, that X has a bounded density with respect to the Lebesgue measure on $[x_0, x_1]$, i.e., $\mathbb{P}_X(dx) = f_X(x)dx$. In that case, the variance becomes

$$\mathrm{Var}_\Pi[\Theta^\star(\mathbb{1}_{x\geq\rho})] = \mathbb{E}_\Pi\left[\Big(\int_{x_0}^{\rho}(\Delta^\star(x) - \hat{\Delta}(x))f_X(x)dx\Big)^2\right] \stackrel{\text{def}}{=} v(\rho).$$

Thus, using the Lebesgue dominated convergence theorem, we have

$$\begin{aligned}\frac{dv}{d\rho}(\rho) &= \mathbb{E}_\Pi\left[2f_X(\rho)(\Delta^\star(\rho) - \hat{\Delta}(\rho))\Big(\int_{x_0}^{\rho}(\Delta^\star(x) - \hat{\Delta}(x))f_X(x)dx\Big)\right] \\ &= 2f_X(\rho)\mathrm{Cov}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x\geq\rho})).\end{aligned}$$

We conclude that

$$\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})}{d\rho}(\rho) = f_X(\rho)\left(-\hat{\Delta}(\rho) + q_\alpha \mathrm{sd}_\Pi(\Delta^\star(\rho))\mathrm{Cor}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x\geq\rho}))\right). \tag{5}$$

We are now ready to state our main theorem, which summarizes the connection between what happens at the individual level and at the aggregated level. Its proof is given in the Appendix A2. Observe that the result does not assume that x_0 is in the domain of the bijection.

Theorem 1. *Assume that X is either discrete or has a bounded density with respect to the Lebesgue on $[x_0, x_1]$. Recall that $X_\alpha = z_\Delta^{-1}(q_\alpha)$. If $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha$, then:*

(i) If $x_0 < X_\alpha$, the strategy $\hat{\texttt{mlb}}_\alpha = \mathbb{1}_{x\geq\rho_{\max,\alpha}}$ defines a non-empty personalized set with $\rho_{\max,\alpha} \geq X_\alpha$. In this case, $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0$.

(ii) If, on the contrary, $X_\alpha \leq x_0$, then the personalized set is empty, i.e, $\rho_{\max,\alpha} = x_0$ and $\hat{\texttt{mlb}}_\alpha =$ `ref`. In this case, $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) = 0$.

On the other hand, if $\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha$, then:

(i) If $X_\alpha < x_1$, the strategy $\hat{\texttt{mlb}}_\alpha = \mathbb{1}_{x\leq\rho_{\max,\alpha}}$ defines a non-empty personalized set with $\rho_{\max,\alpha} \leq X_\alpha$. In this case, $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0$.

(*ii*) *If, on the contrary,* $x_1 \leq X_\alpha$*, then the personalized set is empty, i.e,* $\rho_{\max,\alpha} = x_1$ *and* $\hat{\texttt{mlb}}_\alpha = \texttt{ref}$*. In this case,* $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) = 0$*.*

This theorem is important insofar as it connects the behavior of the max lower bound strategy to the presence or absence of patients who benefit individually from the alternative treatment at level α. If there are some such patients, then we can personalize treatment for a slightly larger set of patients while maximizing the confidence that personalization will be beneficial overall. If there are none, then the strategy is but the `ref` strategy, which does not personalize anyone and recommends the reference treatment to everyone.

Theorem 1 relies on the assumption that the unilateral test for interaction is significant (that is, $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha$). If this assumption is not respected, i.e., $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} \geq q_\alpha$, and if we also have $\hat{\beta}_2 < 0$, then $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha)$ can be positive. As discussed in Subsection 3.2, this would mean an undesirable recommendation of the alternative treatment when neither the treatment effect nor the interaction effect are significant. To circumvent this problem, we propose to simply replace the $\hat{\texttt{mlb}}_\alpha$ by the restricted max lower bound strategy $\hat{\texttt{Mlb}}_\alpha$, which reduces to $\hat{\texttt{mlb}}_\alpha$ under the assumptions of the theorem and does not personalize otherwise. This will allow to control the type I error of the companion test, as we will see in Section 4.3. The restriction depends on the alternative hypothesis for interaction:

$$\hat{\texttt{Mlb}}_\alpha = \mathbb{1}_{-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} \geq q_\alpha} \texttt{ref} + \mathbb{1}_{-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha} \hat{\texttt{mlb}}_\alpha$$

when the test for interaction is unilateral with $H_1^{\text{interact}} : \beta_3 > 0$, and

$$\hat{\texttt{Mlb}}_\alpha = \mathbb{1}_{\hat{\beta}_3/\sqrt{\Sigma_{3,3}} \geq q_\alpha} \texttt{ref} + \mathbb{1}_{\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha} \hat{\texttt{mlb}}_\alpha$$

when the test for interaction is unilateral with $H_1^{\text{interact}} : \beta_3 < 0$. If the test is bilateral, i.e., $H_1^{\text{interact}} : \beta_3 \neq 0$, then we let

$$\hat{\texttt{Mlb}}_\alpha = \mathbb{1}_{-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} \geq q_{\alpha/2}} \texttt{ref} + \mathbb{1}_{-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} < q_{\alpha/2}} \hat{\texttt{mlb}}_\alpha.$$

As $q_{\alpha/2} < q_\alpha$, when $-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} < q_{\alpha/2}$ we are under one of the assumptions of Theorem 1.

We conclude this subsection by studying the gap between X_α and $\rho_{\max,\alpha}$. If we go back to equation (5), we see that if $\text{Cor}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x \geq \rho})) = 1$, then $\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{x \geq \rho})}{d\rho}(\rho) = 0$ if and only if $\rho_{\max,\alpha} = X_\alpha$ defined above. In practice, this will be approximately true as long as ρ is close to x_0, as shown in the next proposition.

Proposition 3. *We have, almost surely,* $\lim_{\rho \to x_0^+} \text{Cor}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x \geq \rho})) = 1$.

Proof. By Lemma A1 in the Appendix A,

$$\begin{aligned}\text{Cor}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x \geq \rho})) &= \text{Cor}_\Pi\big(-\Delta^\star(\rho), -\Delta^\star(g(\rho))\big) \\ &\to_{\rho \to x_0^+} \text{Cor}_\Pi(-\Delta^\star(x_0), -\Delta^\star(x_0)) = 1.\end{aligned}$$

□

This means that in many cases the max lower bound strategy will be very close to the individual strategy defined in (4), i.e.,

$$\hat{\texttt{Mlb}}_\alpha \approx \hat{\texttt{ind}}_\alpha.$$

Furthermore, the proximity of $\mathrm{Cor}_\Pi(-\Delta^\star(X_\alpha), \Theta^\star(\mathbb{1}_{x \geq X_\alpha}))$ with 1 is a diagnostic tool to evaluate the quality of this approximation. If $\mathrm{Cor}_\Pi(-\Delta^\star(X_\alpha), \Theta^\star(\mathbb{1}_{x \geq X_\alpha}))$ is not close to 1, then we can personalize more patients than if we were considering uncertainty at the individual level, while controlling type I error. This is a consequence of subadditivity of standard deviation: if we sum variables that have correlation smaller than 1, the standard deviation of the sum is smaller than the sum of the standard deviations. The two possibilities are illustrated in the next subsection.

4.2 Illustration

In this subsection, we illustrate the behavior of the bilateral $\hat{\texttt{Mlb}}_\alpha$ strategy. We simulate two scenarios: the first scenario is the one we used for Figure 1, to exemplify the problems caused by considering $\hat{\texttt{pol}}_1$, the plug-in estimator of the optimal treatment strategy. Our second scenario illustrates the case where there is substantial distance between X_α and $\rho_{\max,\alpha}$.

In both scenarios, we simulate datasets based on the linear model (2). The specifics of the first scenario are detailed in Figure 1. To show the influence of the choice of strategy, we compute our quantities of interest for a grid of policies depending on a threshold η for z_Δ, with η ranging from $q_{0.02}$ to $q_{0.5}$. For each simulation, we select $\eta_{\max,0.05}$ empirically, i.e., we select the threshold for which $\hat{q}_{n,0.05}(\mathbb{1}_{z_\Delta(x) \geq \eta})$ is maximal. This is shown in Figure 3. As we can see, the problematic points in the upper left quadrant under $\hat{\texttt{pol}}_1$ are brought to the upper right quadrant by decreasing η. Only 0.2% of the simulations where $\hat{q}_{n,0.05}(\hat{\texttt{Mlb}}_{0.05}) > 0$ identify a strategy that is not beneficial, i.e., $\Theta(\hat{\texttt{Mlb}}_{0.05}) \leq 0$. Manual inspection of the instances where this is true show that they correspond to extreme underestimation of the ratio β_2/β_3.

As noted at the end of Section 2, the hypothesis we want to test $H_0 : \Theta(\hat{\texttt{Mlb}}_\alpha) \leq 0$ depends on a random variable. We have seen in Subsection 3.1 that for a poor choice of strategy such as $\hat{\texttt{pol}}_1$, $\mathbb{P}_\beta(H_0 | \hat{q}_{n,\alpha}(\hat{\texttt{pol}}_1) > 0)$ can be substantial even though $\Theta(\texttt{opt}) > 0$. However, for the $\hat{\texttt{Mlb}}_\alpha$ strategy, the simulations are reassuring as this same probability is quite small and the strategy identified is not beneficial only if the personalized set is empty.

Power for detection of interaction in this case is 1 and we are therefore always under the assumption of our theorem. We detect a positive improvement due to personalization in 34% of simulations. There is a small difference between the probability of $\min(z_\Delta) < q_{0.05}$ at 34.7% and $\hat{q}_{n,0.05}(\hat{\texttt{Mlb}}_{0.05}) > 0$ at 34.3% that is likely due to Monte-Carlo noise. The proportion personalized P_{neg} is 5% when the personalized set is not empty instead of the 10.6% of the $\texttt{opt}$ strategy.

The factor identified in the derivative of $\rho \mapsto \hat{q}_{n,0.05}(\mathbb{1}_{x \geq \rho})$ as influencing the quality of the approximation of $\rho_{\max,0.05}$ by $X_{0.05}$, i.e., of $\eta_{\max,0.05}$ by $q_{0.05}$, $\mathrm{Cor}_\Pi(-\Delta^\star(X_{0.05}), \Theta^\star(\mathbb{1}_{z_\Delta(x) < q_{0.05}}))$ is always larger than 0.98 in this scenario. This implies that $q_{0.05}$ is a good approximation of $\eta_{\max,0.05}$ and the $\hat{\texttt{Mlb}}_{0.05}$ strategy is almost equal to the $\hat{\texttt{ind}}_{0.05}$ strategy. Since we sample only 10 000 times from Π, there is still some noise and $q_{0.05}$ is selected only 94% of the time when $\hat{q}_{n,0.05}(\hat{\texttt{Mlb}}_{0.05}) > 0$, the remaining occurrences select either $q_{0.04}$ or $q_{0.07}$, the closest values on the grid.

To conclude on this scenario, the use of the $\hat{\texttt{Mlb}}_{0.05}$ strategy has allowed to deal with the problems that were due to the use of $\hat{\texttt{pol}}_1$, i.e., the low probability of identifying a significant improvement and, when a significant improvement was identified, the large probability of the strategy being detrimental.

With our second scenario, we want to illustrate the interest of using $\eta_{\max,0.05}$ instead of $q_{0.05}$. We

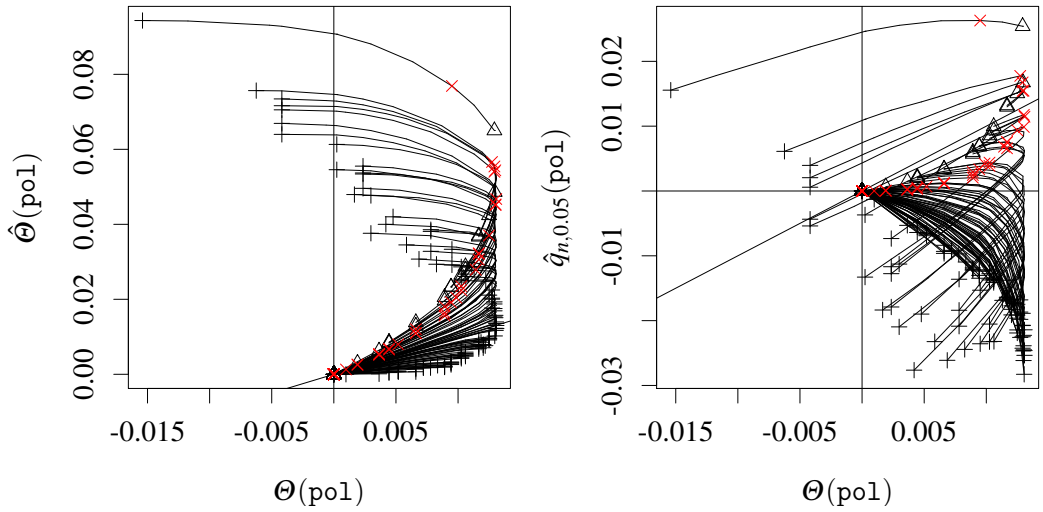

Figure 3: Max lower bound strategy. Graph of $\Theta(\texttt{pol})$ against $\hat{\Theta}(\texttt{pol})$ and $\hat{q}_{n,0.05}(\texttt{pol})$. Each trajectory represents the change in those quantities when η varies. The cross corresponds to $\eta = 0$, i.e., $\texttt{pol} = \hat{\texttt{pol}}_1$. The triangles correspond to the end of the trajectory, i.e., $\rho = q_{0.02} = -2.05$. The red X marks the quantities under the $\hat{\texttt{Mlb}}_{0.05}$ strategy. We plotted only the first 100 trajectories in order to have a readable output. In both plots, the first bisector, i.e., the $y = x$ line, is drawn.

simulate 100 datasets with parameters $\beta_1 = 1$, $\beta_2 = 0$, $\beta_3 = 0.5$, $n = 500$, and $\sigma^2 = 1$. As we still use a uniform distribution on $[-1, 1]$ for X, the optimal treatment strategy would be to personalize half the patients. We sampled 100 000 times from Π in order to minimize Monte Carlo noise and pinpoint precisely $\eta_{\max,0.05}$. We searched for η on a grid with 0.1 increments that contained $q_{0.05}$.

Out of the 70 occurrences where the test for interaction is significant, i.e., $-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} > q_{0.025}$, and $\hat{q}_{n,0.05}(\hat{\texttt{Mlb}}_{0.05}) > 0$, the best choice for η was $q_{0.05}$ only 30 times. All other selected thresholds were larger than $q_{0.05}$, as expected. In these 70 simulations, the average $\text{Cor}_\Pi(\Delta^\star(X_{0.05}), \Theta^\star(\mathbb{1}_{z_\Delta(x)<q_{0.05}}))$ was 0.91. This quantity was 0.99 when $q_{0.05}$ had been selected and 0.86 in the 40 other instances. In the 40 simulations where a larger threshold had been selected, the average P_{neg} was 42% for the threshold $q_{0.05}$ while it was 46% for $\eta_{\max,0.05}$. This shows that the gain from aggregating uncertainty instead of controlling it at the individual level can be substantial.

4.3 Properties

The definition of the restricted max lower bound strategy $\hat{\texttt{Mlb}}_\alpha$ was motivated by the need to combine the rejection region proposed at the end of Section 2, $\{\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0\}$, with the test for interaction presented in Definition 1. One important issue is therefore to show that the combined test controls the type I error when there is no improvement under the optimal treatment strategy, i.e., $\Theta(\texttt{opt}) = 0$. In the Appendix C1, we illustrate that using $\{\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0\}$ as rejection region for our test does not control the type I error rate. On the contrary, combining this rejection region with the one of the interaction test of Proposition 1 allows to control type I error under the null hypothesis $H_0^{\texttt{opt}}$. We have therefore defined a strategy, $\hat{\texttt{Mlb}}_\alpha$, and a companion test such that our strategy personalizes treatment when the null hypothesis of no improvement is rejected and recommends the reference treatment to everyone in the opposite case. The test is also a valid test

for presence of improvement under the optimal treatment strategy. Therefore, $\hat{\texttt{Mlb}}_\alpha$ personalizes treatment whenever we can detect the presence of a theoretical improvement. Janes et al. (2014) propose a test for presence of benefit under the optimal rule but do not derive a rejection region. We expect their test to be identical to ours.

We have defined our estimation procedure using Bayesian arguments, and we have used this estimation to select the strategy $\hat{\texttt{Mlb}}_\alpha$. We checked that the resulting quantiles respect the expected frequentist coverage probabilities, i.e., $\mathbb{P}_\beta(\Theta(\hat{\texttt{Mlb}}_\alpha) \leq \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha)) = \gamma$, were $\gamma \in [0,1]$ is any confidence level, which does not have to be equal to α, the confidence level used in the definition of $\hat{\texttt{Mlb}}_\alpha$. Results of simulations detailed in the Appendix C2 show that coverage is approximately respected. However, there is a clear asymmetry between left and right coverages.

5 Extensions of the method

5.1 Extension to other outcomes

If the outcome Y is binary and not continuous, one option would be to use logistic regression instead of model 2 such as

$$\text{logit}(\mathbb{P}(Y = 1|X, T)) = \beta_0 + \beta_1 X + \beta_2 T + \beta_3 XT,$$

where, for all $p \in [0,1]$, $\text{logit}(p) = \log(\frac{p}{1-p})$. As the maximum likelihood estimator of β is asymptotically normal, we can apply the procedures described in the previous sections. However, the definition of $\Delta(X)$ as $\beta_2 + \beta_3 X$ is not natural as it is unclear what the corresponding Θ would mean. It is more logical to follow Janes et al. (2014), and let

$$\Delta(X) = \mathbb{P}(Y = 1|X, T = 1) - \mathbb{P}(Y = 1|X, T = 0),$$

in which case $\Theta(\texttt{pol})$ will be the difference in event rate between `pol` and `ref`. Thus,

$$\Delta(X) = \frac{\exp(\beta_0 + \beta_1 X + \beta_2 + \beta_3 X)}{1 + \exp(\beta_0 + \beta_1 X + \beta_2 + \beta_3 X)} - \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)},$$

an expression that depends on the whole of β instead of just β_2 and β_3.

We can then, as before, sample $\beta^\star$ from Π, the asymptotic posterior distribution of β, and obtain quantiles of our quantity of interest. Instead of using z_Δ as we did under a Gaussian assumption, a natural idea is to work with $\Pi(\Delta^\star(x) > 0)$. Indeed, in the Gaussian case, we have $z_\Delta(x) < q_\alpha \Leftrightarrow \Pi(\Delta^\star(x) > 0) < \alpha$. We then look for policies of the form $\mathbb{1}_{\Pi(\Delta^\star(x)>0)<\gamma}$, with $\gamma \in [0,1]$. Our threshold is now on the probability scale between 0 and 1. We apply this extension to real data in Section 6.

Similarly, our approach can be extended to a censored outcome. In that case, one should first select a measure of treatment contrast, which could be either the survival probability at a prespecified time τ, or the restricted mean survival up to τ, for instance. A Cox proportional hazards model can then be fitted to the date and then used to derive the survival probability at τ or the restricted mean survival up to τ (Zhao et al., 2013; Li et al., 2016).

5.2 Extension to the multivariate case

In most RCTs, a large number of covariates describe each patient instead of just one, as we previously assumed. In order for personalized medicine to fulfill its promises, it is crucial that the information contained in this data be harnessed.

Let us then assume that $X = (X_1, \ldots, X_d)$ is now d-dimensional. Using a machine learning algorithm, we can predict Y given X and T. In the present paper, we use the random forests algorithm (Breiman, 2001), and denote by $\hat{h}(X, T)$ such a prediction. We can then define, as before,

$$\hat{\Delta}(X) = \hat{h}(X, 1) - \hat{h}(X, 0).$$

The procedure to estimate $\Theta(\texttt{pol})$ we proposed in Section 2 was based on our knowledge of the joint distribution of $(\Delta^\star(x))_x$, with $\Delta^\star(x)$ following the posterior distribution of $\Delta(X)$. In the case of a linear model, this distribution was quite simple, thanks to the straightforward relation between Δ and the coefficients, namely $\Delta^\star(x) = \beta_2^\star + \beta_3^\star x$. Unfortunately, such a simple technique does not immediately transpose to machine learning algorithms. Of course, there are ways to estimate the uncertainty around a prediction, for example using bootstrap methods (Tibshirani, 1996; Wager et al., 2014), but they only concern single predictions. Besides, studying the joint distribution of prediction errors of machine learning algorithms exceeds the scope of this article. Instead, we follow in this section Chernozhukov et al. (2018). This will allow us to extend what we have developed in the univariate case. The idea is to use $\hat{h}$ to build features that will be plugged in a linear model, for which we will be able to apply the procedure detailed in Section 2.

Our univariate linear model has three parts besides the intercept: a prognostic term $\beta_1 X$, a reference treatment average effect term β_2, and an interaction term $\beta_3 X$, with the last two multiplied by treatment assignment T. The prognostic term is the expected value given X when a patient receives the alternative treatment. Therefore, we define our first feature as $Z_1(X) = \hat{h}(X, 0)$ in order to capture the prognosis. Besides, the treatment effect $\Delta(X)$ is estimated by $\hat{h}(X, 1) - \hat{h}(X, 0)$. Since $\Delta(X) = \beta_2 + \beta_3 X$ in the linear case, and $\mathbb{E}_X X = 0$, it is logical to define the interaction term by $Z_3(X) = \hat{h}(X, 1) - \hat{h}(X, 0) - \bar{\Delta}$, where $\bar{\Delta}$ is the empirical mean of $\hat{\Delta}(X)$. All in all, we assume that our outcome follows a linear model, of the form

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 (T - \bar{T}) + \beta_3 Z_3 (T - \bar{T}) + \epsilon.$$

Chernozhukov et al. (2018) show that given $\hat{h}$ the best linear prediction of the individual treatment effect is $\beta_2 + \beta_3 Z_3$. Unlike in the univariate case where we had a main term and an interaction term, here Z_1 is not the main effect of Z_3 and therefore Z_3 might have a main term. Centering T ensures that only the interaction between Z_3 and T will be captured.

Observe that, because of the way we have built our features, we expect the coefficients in the linear regression to be close to natural values that correspond to perfect estimation of Y by $\hat{h}$: $\beta_1 = 1$, $\beta_2 = \bar{\Delta}$, and $\beta_3 = 1$. If the coefficients are far from their expected values, if for example β_3 is negative, then we should be suspicious of the results. This motivates the use of a unilateral test for interaction in the test described in Section 4.3.

Overfitting is also a concern in our setting, because an overfit prediction will always find an interaction term Z_3 by fitting the noise, and the same noise will be present when doing the linear

regression. Chernozhukov et al. (2018) advocate for a sample splitting approach. We prefer to use the out-of-bag estimators of random forests or more generally a prediction obtained in cross-validation in order to have a full sample for the regression.

Following the approach developed above, we can test for presence of an improvement and estimate the improvement under the max lower bound strategy. Let us finally mention that the crux of the approach is to replace the joint distribution of prediction errors by the joint distribution of $(\beta_2^\star + \beta_3^\star Z_3(x))_x$. A reason to expect the procedure to be somewhat conservative is that very different x will lead to similar values of Z_3. If we have two distant points x_1 and x_2 such that $Z_3(x_1) \approx Z_3(x_2)$, then the prediction errors of $Z_3(x_1)$ and $Z_3(x_2)$ might be less correlated than in our procedure, where the correlation is one. Given the subadditivity of standard deviation, taking this into account would lead the aggregate quantity to carry less uncertainty. A promising research perspective is therefore to study the joint distribution of prediction errors of machine learning algorithms to be able to provide tighter confidence bounds and therefore extend personalization to as many patients as possible.

6 Illustration on real data

We illustrate our approach on the data of the Rituximab in ANCA-Associated Vasculitis (RAVE) (Stone et al., 2010). RAVE is multicenter, randomized, double-blind, double-dummy, noninferiority trial of rituximab versus cyclophosphamide for remission induction of severe antineutrophil cytoplasmic antibody associated vasculitis. There were 99 patients assigned to rituximab and 98 assigned to cyclophosphamide. The primary outcome was a measure of remission at 6 months. Overall, the proportion of patients in remission at 6 months was 64% with rituximab vs. 53% for cyclophosphamide ($p < 0.001$ for the non-inferiority test that was the initial aim of the trial and $p = 0.1$ for the usual superiority test). This trial was an interesting candidate for our procedure as the average treatment effect was positive but not significant, but the original article underlined that a prespecified subgroup analysis identified that rituximab patients significantly improved the outcome of patient who presented wit relapsing disease, compared to cyclophosphamide.

We used 27 baseline covariates to model the outcome (see Table A.1 in the Supplementary Material). Missing covariate data were imputed once using an iterative Factorial Analysis for Mixed Data (FAMD) algorithm (Audigier et al., 2016) in order to obtain a dataset with no missing value.

Using this completed dataset, we trained random forests in classification to predict the outcome. Each tree was grown with a bootstrap version of the sample. Each observation therefore had probability around 1/3 to not be used to grow a tree, and such an observation is called out-of-bag in the random forest vocabulary. To avoid overfitting, we used the out-of bag estimate, i.e., only the trees for which the observation is out-of-bag are used to predict. This means that the observation was not used in the model that predicts its outcome. We grew 1 500 trees in order to have around 500 trees for each out-of-bag prediction. For each observation (X, T), we used the out-of-bag forest to predict the treatment effect $Z_1 = \hat{h}(X, 0)$ and $\hat{\Delta}(X) = \hat{h}(X, 1) - \hat{h}(X, 0)$, where $\hat{h}(X, T)$ is the mean vote of the forest.

We then computed $Z_3 = \hat{\Delta}(X) - \bar{\Delta}$, and performed logistic regression of Y on Z_1, $T - \bar{T}$, and $Z_3(T - \bar{T})$. As Z_1 and Z_3 are not on the logit scale, we do not expect for their coefficients to be

close to 1, but the sign of the coefficients should be positive. This was the case for Z_3 but not Z_1. The test for the average treatment effect term yielded $p = 0.1$ which is consistent with the result in the original study. But the test for the interaction coefficient was not significant even at a liberal 0.10 significance threshold. As a consequence the personalized set was empty. Compared to the original article, it has to be noted that the subgroup possibly benefitting from rituximab was identified among eight subgroup analyses with no correction for multiplicity. Our results thus confirm that this analysis likely lacked robustness.

7 Discussion

In this paper, we have shown the importance of considering the uncertainty in the estimated strategy when estimating the benefit of personalization. Indeed, the optimal strategy is never known, and we have to estimate a strategy and its benefit simultaneously. In order to deal with this uncertainty, we chose to prioritize one treatment over the other, just as we prioritize the null hypothesis in significance testing. This has led us to advocate for the max lower bound strategy. It is the strategy for which the α-credible quantile of the improvement under personalization will be maximal. The personalized set it defines is often close to the region of superiority of the alternative treatment in Shuster and van Eys (1983), but can be substantially larger.

The asymmetry between treatments is induced by privileging the reference treatment that would usually be received if no treatment personalization was implemented. This implies in particular that the max lower bound strategy still assigns the reference treatment to some patients for whom it is predicted that the alternative treatment would lead to a more favorable outcome. This could seem unnatural, but we argue it is however necessary to control the risk of implementing a strategy that leads to a worse average outcome than the reference non-personalized strategy. It is therefore crucial to ensure that a change in strategy is beneficial to patients on average, in the spirit of the *primum non nocere* principle in medicine.

Our estimation of benefit is based on Bayesian arguments. The interest of considering Bayesian estimation is that it allows us to consider the estimated strategy as fixed. Nevertheless, we are interested in frequentist properties of our estimation such as type I error and coverage probabilities. Using theoretical arguments and simulations, we saw that these statistical properties appear to be respected. Unfortunately, the positivity of the credible quantile $\hat{q}_{n,\alpha}(\hat{\mathtt{mlb}}_\alpha)$ does not offer a stand-alone test because of somewhat pathological behavior when the estimated treatment effect $\hat{\beta}_2$ is negative. However, it can be combined with the test for presence of interaction, which allows to deal with this issue while not affecting the result of the test when $\hat{\beta}_2$ is positive. The resulting test is a valid test for presence of improvement under the optimal strategy.

In a more speculative part of our paper, we proposed to extend our approach to the multivariate case, by creating features from the prediction of a machine learning algorithm, and applying what we have developed in the univariate case. An important consideration underlying this strategy was to avoid overfitting. A more formal extension of our approach to the multivariate case may involve the derivation of the joint distribution of prediction errors for machine learning algorithms, which needs further work. Also, it may be interesting to study the gain of using recently proposed X-learners (Künzel et al., 2019) to estimate the predictions $\hat{\mathbb{E}}[Y^1 - Y^0|X]$ (also termed conditional

average treatment effects) instead of random forests, for instance. The issue of estimating credible quantiles with X-learner however also remains unsolved.

In the multivariate case, other authors have proposed different approaches to control for overfitting and uncertainty in model-based predictions of individual treatment effects $\hat{\mathbb{E}}[Y^1 - Y^0|X]$. For instance Li et al. (2016) have proposed a two- or three-step procedure. If there are sufficient data, the dataset is divided into three independent subsets. The first subset serves for estimating candidate models. The second subset is used to define the personalized set, for instance by using the lower one-sided 95% confidence bound of predictions. Then, the properties of the resulting strategy are estimated in the third subset. If data are not sufficient, the first two stages are replaced by a cross-training stage, where one subset is iteratively randomly divided into a training and an evaluation set, final predictions being then averaged over the repetitions. While this approach has been shown to have good properties in simulation studies, in particular in controlling the type I error rate, it necessitates a fair amount of data for splitting. It would be interesting to investigate how it compares with the max lower bound strategy.

It could also be interesting to study how the max lower bound strategy compares to (or could be used in conjunction with) residual weighted learning approaches, where the issue of determining the optimal strategy is directly handled as a weighted classification problem (Zhao et al., 2012; Zhou et al., 2017).

All aforementioned methods aiming at identifying treatment strategies are making use of already collected data. Once an "optimal" strategy has been determined, it should therefore also be itself evaluated in a randomized trial. To approach the optimal treatment strategy and maximize outcomes in the long term, it could also be recommended to conduct randomized trials focusing on patients in the region of no superiority of any treatment. Such trials may provide adequate data to refine treatment strategies.

Acknowledgements

This article contains information from the IST dataset which is made available under the ODC Attribution License (`https://datashare.is.ed.ac.uk/bitstream/handle/10283/128/license_text?sequence=12&isAllowed=y`). The authors gratefully acknowledge the collaborative group that collected the data and funding that allowed making these data available.

References

Athey, S. and Wager, S. (2017). Efficient policy learning. *arXiv preprint arXiv:1702.02896*.

Audigier, V., Husson, F., and Josse, J. (2016). A principal component method to impute missing values for mixed data. *Advances in Data Analysis and Classification*, 10:5–26.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

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.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2018). Generic machine learning inference on heterogenous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research.

Freidlin, B. and Korn, E. (2014). Biomarker enrichment strategies: Matching trial design to biomarker credentials. *Nature Reviews Clinical Oncology*, 11:81–90.

Hamburg, M. and Collins, F. (2010). The path to personalized medicine. *New England Journal of Medicine*, 363:301–304.

Janes, H., Brown, M., Huang, Y., and Pepe, M. (2014). An approach to evaluating and comparing biomarkers for patient treatment selection. *The International Journal of Biostatistics*, 10:99–121.

Kang, C., Janes, H., and Huang, Y. (2014). Combining biomarkers to optimize patient treatment recommendations. *Biometrics*, 70:695–707.

Khoury, M. (2010). Dealing with the evidence dilemma in genomics and personalized medicine. *Clinical Pharmacology and Therapeutics*, 87:635–638.

Künzel, S., Sekhon, J., Bickel, P., and Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165.

Li, J., Zhao, L., Tian, L., Cai, T., Claggett, B., Callegaro, A., Dizier, B., Spiessens, B., Ulloa-Montoya, F., and Wei, L. (2016). A predictive enrichment procedure to identify potential responders to a new therapy for randomized, comparative controlled clinical studies. *Biometrics*, 72:877–887.

Mandrekar, S. and Sargent, D. (2009). Clinical trial designs for predictive biomarker validation: Theoretical considerations and practical challenges. *Journal of Clinical Oncology*, 27:4027–4034.

Qian, M. and Murphy, S. (2011). Performance guarantees for individualized treatment rules. *The Annals of Statistics*, 39:1180–1210.

Qiu, X., Zeng, D., and Wang, Y. (2018). Estimation and evaluation of linear individualized treatment rules to guarantee performance. *Biometrics*, 74(2):517–528.

Shen, Y. and Cai, T. (2016). Identifying predictive markers for personalized treatment selection. *Biometrics*, 72:1017–1025.

Shuster, J. and van Eys, J. (1983). Interaction between prognostic factors and treatment. *Controlled Clinical Trials*, 4:209–214.

Stone, J. H., Merkel, P. A., Spiera, R., Seo, P., Langford, C. A., Hoffman, G. S., Kallenberg, C. G., St. Clair, E. W., Turkiewicz, A., Tchao, N. K., et al. (2010). Rituximab versus cyclophosphamide for anca-associated vasculitis. *New England Journal of Medicine*, 363(3):221–232.

Tibshirani, R. (1996). A comparison of some error estimates for neural network models. *Neural Computation*, 8:152–163.

van der Vaart, A. (1998). *Bayes Procedures*, pages 138–152. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

van Klaveren, D., Balan, T. A., Steyerberg, E. W., and Kent, D. M. (2019). Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. *Journal of Clinical Epidemiology*.

Vivot, A., Boutron, I., Ravaud, P., and Porcher, R. (2015). Guidance for pharmacogenomic biomarker testing in labels of FDA-approved drugs. *Genetics in Medicine*, 17:733–738.

Vivot, A., Li, J., Zeitoun, J., Mourah, S., Crequit, P., Ravaud, P., and Porcher, R. (2016). Pharmacogenomic biomarkers as inclusion criteria in clinical trials of oncology-targeted drugs: A mapping of ClinicalTrials.gov. *Genetics in Medicine*, 18:796–805.

Wager, S., Hastie, T., and Efron, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *Journal of Machine Learning Research*, 15:1625–1651.

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

Zhao, L., Tian, L., Cai, T., Claggett, B., and Wei, L. (2013). Effectively selecting a target population for a future comparative study. *Journal of the American Statistical Association*, 108:527–539.

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.

Zhou, X., Mayer-Hamblett, N., Khan, U., and Kosorok, M. (2017). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*, 112:169–187.

Supplementary Material

Appendix A: Proofs

A1 Proof of Proposition 1

We have

$$\Theta^\star(\hat{\texttt{pol}}) = \mathbb{E}_X[\Delta^\star(X)\mathbb{1}_{\hat{\texttt{pol}}(X)=0}] = \beta_2^\star \mathbb{P}_X(\hat{\texttt{pol}}(X)=0) + \beta_3^\star \mathbb{E}_X[X\mathbb{1}_{\hat{\texttt{pol}}(X)=0}],$$

which is a linear combination of a Gaussian vector and is therefore Gaussian. Using Fubini's theorem, we may write

$$\mathbb{E}_\Pi[\Theta^\star(\hat{\texttt{pol}})] = \mathbb{E}_X\left[\mathbb{E}_\Pi[-\Delta^\star(X)]\mathbb{1}_{\hat{\texttt{pol}}(X)=0}\right] = \mathbb{E}_X[-\hat{\Delta}(X)\mathbb{1}_{\hat{\texttt{pol}}(X)=0}] = \hat{\Theta}(\hat{\texttt{pol}}).$$

Besides,

$$\mathrm{Var}_\Pi[\Theta^\star(\hat{\texttt{pol}})] = \mathbb{E}_\Pi\left[(\Theta^\star(\hat{\texttt{pol}}) - \hat{\Theta}(\hat{\texttt{pol}}))^2\right] = \mathbb{E}_\Pi\left[\mathbb{E}_X^2[(\Delta^\star(X) - \hat{\Delta}(X))\mathbb{1}_{\hat{\texttt{pol}}(X)=0}]\right].$$

A2 Proof of Theorem 1

For simplicity, it is assumed throughout that X has a bounded density with respect to the Lebesgue measure on $[x_0, x_1]$. Proofs are similar in the discrete case and left to the reader. We only consider the case where $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha$. The proof is analogous in the other case.

We first begin by stating a lemma.

Lemma A1. *There exists a non-decreasing function g from $[x_0, x_1]$ such that $g(x_0) = x_0$, for all $\rho > x_0$, $x_0 < g(\rho) < \rho$ and*

$$\Theta^\star(\mathbb{1}_{x\geq\rho}) = -\mathbb{P}_X(X<\rho)\Delta^\star(g(\rho)).$$

Proof. Observe that

$$\Theta^\star(\mathbb{1}_{x\geq\rho}) = \mathbb{P}_X(X<\rho)(-\beta_2^\star - \beta_3^\star \mathbb{E}_X[X|X<\rho])$$

and let, for $\rho > x_0$, $g(\rho) = \mathbb{E}_X[X|X<\rho]$.

To see that g is non-decreasing, we derivate for $\rho > x_0$ and obtain

$$g'(\rho) = \frac{f_X(\rho)}{\mathbb{P}_X^2(X<\rho)} \int_{x_0}^{\rho} (\rho - x) f_X(x) dx \geq 0.$$

The inequality is strict except when $f_X(\rho) = 0$.

For all $\rho > x_0$, the inequalities $x_0 < g(\rho) < \rho$ are straightforward from the definition of g. We then have $\lim_{\rho \to x_0^+} g(\rho) = x_0$ and we prolongate by continuity so that $g(x_0) = x_0$. □

We are now ready to prove the theorem.

Proof of Theorem 1. The assumption $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha$ guarantees that q_α is in the image set of the bijection defined in Proposition 2 and that we can safely define $X_\alpha = z_\Delta^{-1}(q_\alpha)$. The function $\rho \mapsto \hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})$ attains a maximum in $[x_0, x_1]$ as a continuous function in a compact set.

In the case $X_\alpha \leq x_0$, we use Lemma A1. Then, for all $\rho \in (x_0, -\hat{\beta}_2/\hat{\beta}_3)$, we may write

$$\Theta^\star(\mathbb{1}_{x\geq\rho}) = -\mathbb{P}_X(X < \rho)\Delta^\star(g(\rho))$$

and

$$\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho}) = \mathbb{P}_X(X < \rho)\left(-\hat{\Delta}(g(\rho)) + q_\alpha \mathrm{sd}_\Pi(\Delta^\star(g(\rho)))\right).$$

As $g(\rho) > x_0 \geq X_\alpha$, we have

$$-\hat{\Delta}(g(\rho)) + q_\alpha \mathrm{sd}_\Pi(\Delta^\star(g(\rho))) < 0,$$

and therefore $\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho}) < 0$, $\forall \rho \in (x_0, -\hat{\beta}_2/\hat{\beta}_3)$, and the maximum is 0 attained in x_0.

If $x_0 < X_\alpha$ and x_0 is in the domain of the bijection defined in Proposition 2, we can write using equation (5), for all $\rho \in [x_0, X_\alpha]$,

$$\begin{aligned}\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})}{d\rho}(\rho) = f_X(\rho)\big(&- \hat{\Delta}(\rho) + q_\alpha \mathrm{sd}_\Pi(\Delta^\star(\rho)) \\ &+ q_\alpha \mathrm{sd}_\Pi(\Delta^\star(\rho))(\mathrm{Cor}_\Pi(-\Delta^\star(\rho), \Theta^\star(\mathbb{1}_{x\geq\rho})) - 1)\big).\end{aligned}$$

The sum of the first two terms in the parenthesis is positive, except in X_α where it is 0, as $z_\Delta(\rho) > q_\alpha$ $\forall \rho \in [x_0, X_\alpha)$, thanks to Proposition 2. The last term is positive, except in x_0 where it is 0, as the product of two negative factors ($\alpha < 0.5$ and the correlation is smaller than 1). The parenthesis is therefore strictly positive. As $f_X(\rho)$ can be 0, we have $\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})}{d\rho}(\rho) \geq 0$, $\forall \rho \in [x_0, X_\alpha]$, and $\rho \mapsto \hat{q}_{n,\alpha}(\mathbb{1}_{x\geq\rho})$ increases on that interval. It follows that the maximum is attained for $\rho_{\max,\alpha} \geq X_\alpha$ and is positive, since $\hat{q}_{n,\alpha}(\mathbb{1}_{x\geq x_0}) = 0$ and $f_X(\rho)$ puts mass in a neighborhood of x_0.

If $x_0 < X_\alpha$ and x_0 is not in the domain of the bijection defined in Proposition 2, then the parametrization of our class of policies on the x scale is not valid. However, once parametrization is dealt with, the same arguments lead to the same result. We now detail the parametrization in this case.

If $x_0 > x_2 \stackrel{\text{def}}{=} \arg\min z_\Delta$, then we can still define a bijection with domain $[x_0, -\hat{\beta}_2/\hat{\beta}_3]$ and the calculations above apply.

However, if $x_0 < x_2$, we cannot replace $x \mapsto \mathbb{1}_{z_\Delta(x)\geq\eta}$ by $x \mapsto \mathbb{1}_{x\geq\rho}$. We have to define two bijections g_1 and g_2 with image set $[\eta_0 \stackrel{\text{def}}{=} \min(z_\Delta), z_\Delta(x_0)]$, as shown in Figure A1. The function g_1 is decreasing while g_2 is increasing, and we have, for all $\eta \leq z_\Delta(x_0)$, $z_\Delta(x) < \eta \Leftrightarrow g_1(\eta) < x < g_2(\eta)$.

The mean improvement of this strategy is then

$$\hat{\Theta}(\mathbb{1}_{z_\Delta(x)\geq\eta}) = -\int_{g_1(\eta)}^{g_2(\eta)} \hat{\Delta}(x) f_X(x) dx,$$

and its derivative

$$\frac{d\hat{\Theta}}{d\eta}(\eta) = -g_2'(\eta)\hat{\Delta}(g_2(\eta)) f_X(g_2(\eta)) + g_1'(\eta)\hat{\Delta}(g_1(\eta)) f_X(g_1(\eta)).$$

The variance under Π is

$$\mathrm{Var}_\Pi[\Theta^\star(\mathbb{1}_{z_\Delta(x)\geq\eta})] = \mathbb{E}_\Pi\Big[\Big(\int_{g_1(\eta)}^{g_2(\eta)} (\Delta^\star(x) - \hat{\Delta}(x)) f_X(x) dx\Big)^2\Big] \stackrel{\text{def}}{=} v(\eta).$$

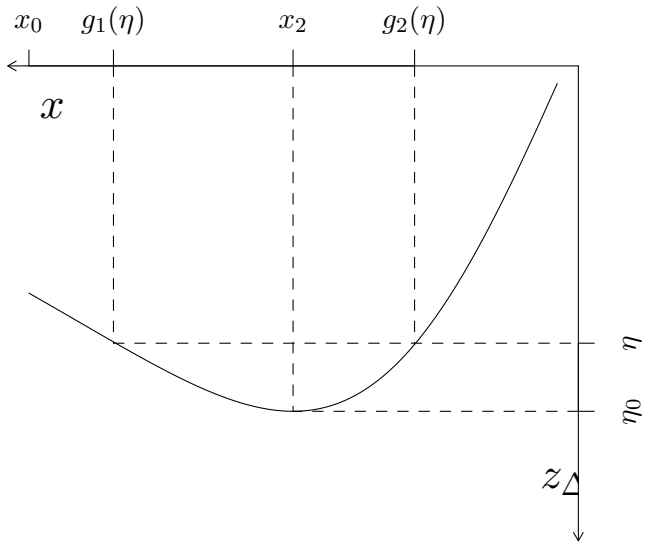

Figure A1: Double parametrization of z_Δ.

Thus,

$$\begin{aligned}\frac{dv}{d\eta}(\eta) = {} & 2g_2'(\eta) f_X(g_2(\eta)) \mathrm{Cov}_\Pi(-\Delta^\star(g_2(\eta)), \Theta^\star(\mathbb{1}_{z_\Delta(x)\geq\eta})) \\ & - 2g_1'(\eta) f_X(g_1(\eta)) \mathrm{Cov}_\Pi(-\Delta^\star(g_1(\eta)), \Theta^\star(\mathbb{1}_{z_\Delta(x)\geq\eta})).\end{aligned}$$

From this, we obtain the derivative of our quantile with respect to η, that is,

$$\begin{aligned}&\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{z_\Delta(x)\geq\eta})}{d\eta}(\eta) \\ &\quad = g_2'(\eta) f_X(g_2(\eta)) \Big(- \hat{\Delta}(g_2(\eta)) \\ &\qquad\qquad + q_\alpha \mathrm{sd}_\beta\big(\hat{\Delta}(g_2(\eta))\big) \mathrm{Cor}_\Pi(-\Delta^\star(g_2(\eta)), \Theta^\star(\mathbb{1}_{z_\Delta(x)\geq\eta})) \Big) \\ &\quad\quad - g_1'(\eta) f_X(g_1(\eta)) \Big(- \hat{\Delta}(g_1(\eta)) \\ &\qquad\qquad + q_\alpha \mathrm{sd}_\beta\big(\hat{\Delta}(g_2(\eta))\big) \mathrm{Cor}_\Pi(-\Delta^\star(g_1(\eta)), \Theta^\star(\mathbb{1}_{z_\Delta(x)\geq\eta})) \Big).\end{aligned}$$

As g_2 is increasing and g_1 is decreasing, both terms have the same sign, and the same argument as above shows that $\frac{d\hat{q}_{n,\alpha}(\mathbb{1}_{z_\Delta(x)\geq\eta})}{d\eta}(\eta) > 0$ for $\eta \leq q_\alpha$. □

Appendix B: Study of z_Δ variations

Note first that $z_\Delta(x) = 0 \Leftrightarrow x = -\hat{\beta}_2/\hat{\beta}_3$. The limits of z_Δ in $-\infty$ and $+\infty$ are, respectively, $\lim_{-\infty} z_\Delta = -\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$ and $\lim_{+\infty} z_\Delta = \hat{\beta}_3/\sqrt{\Sigma_{3,3}}$. This is the z-statistic of the test for presence of an interaction seen in Definition 1. In addition,

$$\frac{dz_\Delta}{dx}(x) = \frac{Ax + B}{\mathrm{sd}_\Pi(\Delta^\star(x))^{3/2}},$$

where $A = -\hat{\beta}_2\Sigma_{3,3} + \hat{\beta}_3\Sigma_{2,3}$ and $B = -\Sigma_{2,3}\hat{\beta}_2 + \hat{\beta}_3\Sigma_{2,2}$.

The signs of A and B decide the shape of z_Δ. Since X and T are independent, $\Sigma_{2,3}$ will be approximately 0. If we set it equal to 0 in the previous expressions, we obtain $A = -\hat{\beta}_2\Sigma_{3,3}$ and $B = \hat{\beta}_3\Sigma_{2,2}$. For simplicity, we will assume that $\Sigma_{2,3} = 0$ throughout, as this allows to discuss more easily interpretable cases. For example, instead of $-\hat{\beta}_2\Sigma_{3,3} + \hat{\beta}_3\Sigma_{2,3} < 0$, we will have $\hat{\beta}_2 > 0$ and instead of $-\Sigma_{2,3}\hat{\beta}_2 + \hat{\beta}_3\Sigma_{2,2} > 0$, we will have $\hat{\beta}_3 > 0$. As $\Sigma_{2,3}$ is close to 0, the probability of the sign of A (respectively, of the sign of B) being different than minus the sign of $\hat{\beta}_2$ (respectively, than the sign of $\hat{\beta}_3$) is small. Note however that if we want to apply our work to observational data, where X and T are not independent, then the more complicated inequalities should be used.

Figure 2 in the main manuscript shows how z_Δ typically varies, depending on the sign of $\hat{\beta}_2$ when $\hat{\beta}_3 > 0$. The sign of $\hat{\beta}_3$ does not affect the behavior: if we change its sign, the curve is simply the symmetric of the original curve with respect to the ordinate axis. If $\hat{\beta}_2 > 0$, then z_Δ is increasing on $(-\infty, -\hat{\beta}_2/\hat{\beta}_3]$ and therefore defines a bijection on its image set $(-\hat{\beta}_3/\sqrt{\Sigma_{3,3}}, 0]$. As $-\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$ is the z-statistic for the test of presence of an interaction, for no individual can the alternative treatment be recommended at a level α smaller than the p-value of the test for presence of interaction.

However, if $\hat{\beta}_2 < 0$, then z_Δ decreases before increasing. This means that for C characterized by $z_\Delta(C) = -\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$, we have $\forall x < C$, $z_\Delta(x) < -\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$, i.e., we can have more evidence for the use of the alternative treatment in some people than for the presence of an interaction. The condition $\hat{\beta}_2 < 0$ means that the observed treatment effect is negative. If we refer back to our definition of the reference and the alternative treatment in the introduction, $\hat{\beta}_2 < 0$ can happen only when the new treatment was superior to the old treatment but not significantly. It is therefore not desirable to recommend the alternative treatment to patients with $z_\Delta(x) < q_\alpha < -\hat{\beta}_3/\sqrt{\Sigma_{3,3}}$ as this means that we recommend an alternative treatment when neither the treatment effect nor the interaction effect are significant. In Section 4 of the main manuscript, we advocate to combine the interaction test with the test for improvement in order to exclude this situation.

Nevertheless, as long as we consider levels α larger than the p-value attained by the interaction, we are in a similar position to the one we had in the case $\hat{\beta}_2 > 0$. Indeed, in this case z_Δ defines a bijection from $(C, -\hat{\beta}_2/\hat{\beta}_3]$ to $(-\hat{\beta}_3/\sqrt{\Sigma_{3,3}}, 0]$. This means that we consider only levels α larger than the p-value for interaction.

Appendix C: Additional properties of the $\hat{\texttt{Mlb}}_\alpha$ strategy

C1 Testing for benefit of personalization

We develop here on the need to combine the rejection region we proposed at the end of Section 2, $\{\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0\}$, with the test for interaction presented in Definition 1.

Let us recall that we want to test the null hypothesis of no benefit of personalization under the $\hat{\texttt{mlb}}_\alpha$ strategy, i.e., $H_0 : \Theta(\hat{\texttt{mlb}}_\alpha) \leq 0$, against the alternative hypothesis that there is a gain from personalization, i.e., $H_1 : \Theta(\hat{\texttt{mlb}}_\alpha) > 0$. As noted above, these hypotheses depend on random variables. However, when the optimal strategy defines an empty personalized set, i.e., $\Theta(\texttt{opt}) = 0$, the estimated strategy can never be beneficial, since $\Theta(\hat{\texttt{mlb}}_\alpha) \leq \Theta(\texttt{opt}) = 0$. Therefore, when $\Theta(\texttt{opt}) = 0$, $\mathbb{P}_\beta(H_0) = 1$. Thus, letting $H_0^{\texttt{opt}} : \Theta(\texttt{opt}) = 0$, we have $H_0^{\texttt{opt}} \subset H_0$, with the added advantage that $H_0^{\texttt{opt}}$ is a traditional null hypothesis that depends only on fixed quantities. In the sequel, we therefore study the well-defined type I error under the null hypothesis $H_0^{\texttt{opt}}$.

We now provide motivation for the combination of the interaction test with the test we proposed based on $\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha)$. We will show that $\{\hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0\}$ does not control type I error under $H_0^{\texttt{opt}}$ and that this is linked with the pathological behavior of z_Δ when $\hat{\beta}_2 < 0$. As we have seen in Subsection 3.2, when $\hat{\beta}_2 < 0$ the function z_Δ has a downward bump and can therefore attain values smaller than $-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}}$. This situation can occur quite frequently. Assume, for example, that $\beta_3 = \beta_2 = 0$, and let us illustrate this scenario with simulations following the protocol in the legend of Figure C2. As $\hat{\beta}_2$ is centered and Gaussian, we have $\mathbb{P}_\beta(\hat{\beta}_2 < 0) = 1/2$. Furthermore, the scenario $\beta_3 = \beta_2 = 0$ falls under $H_0^{\texttt{opt}}$. The simulations show that the quantile we have focused on, $\hat{q}_{n,0.05}(\hat{\texttt{mlb}}_{0.05})$, is positive 1 time out of 4 when $\hat{\beta}_2 < 0$. As a consequence, if we use $\{\hat{q}_{n,0.05}(\hat{\texttt{mlb}}_{0.05}) > 0\}$ as the rejection region of our test, the type I error rate will be 14%. The excess in type I error rate above its nominal level comes exclusively from the simulations where $\hat{\beta}_2 < 0$. As we see next, combining this rejection region with the one of the interaction test of Proposition 1 allows to control type I error under the null hypothesis $H_0^{\texttt{opt}}$.

Depending on the clinical context, the test for interaction can be unilateral or bilateral, and we advocate for the use of the corresponding (unilateral or bilateral) $\hat{\texttt{Mlb}}_\alpha$. When the alternative hypothesis for the interaction test is $H_1^{\text{interact}} : \beta_3 > 0$, i.e., the test is unilateral, then the border between $H_0^{\texttt{opt}}$ and the alternative hypothesis $H_1^{\texttt{opt}} : \Theta(\texttt{opt}) > 0$ and $\beta_3 > 0$ is

$$B = \{(\beta_2, \beta_3), \beta_3 \geq 0, \Delta(x_0) = 0\} = \{(\beta_2, \beta_3), \beta_3 \geq 0, \beta_2 = -x_0\beta_3\}.$$

The rejection region of the joint test is

$$R = \{-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha, \hat{q}_{n,\alpha}(\hat{\texttt{mlb}}_\alpha) > 0\}.$$

Combining the test based on our quantile with the interaction test means that we always respect the assumption of Theorem 1 in the rejection region, and we can therefore apply the theorem to write $R = \{-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha, x_0 < X_\alpha\}$. As we have $x_0 < X_\alpha \Leftrightarrow z_\Delta(x_0) < q_\alpha$, the rejection region then becomes:

$$R = \{-\hat{\beta}_3/\sqrt{\Sigma_{3,3}} < q_\alpha, z_\Delta(x_0) < q_\alpha\}.$$

On the border B, $\Delta(x_0) = 0$, and therefore, under $\mathbb{P}_\beta$, $z_\Delta(x_0) \sim \mathcal{N}(0,1)$. As $\mathbb{P}_\beta(z_\Delta(x_0) < q_\alpha) = \alpha$, we have control of the type I error rate:

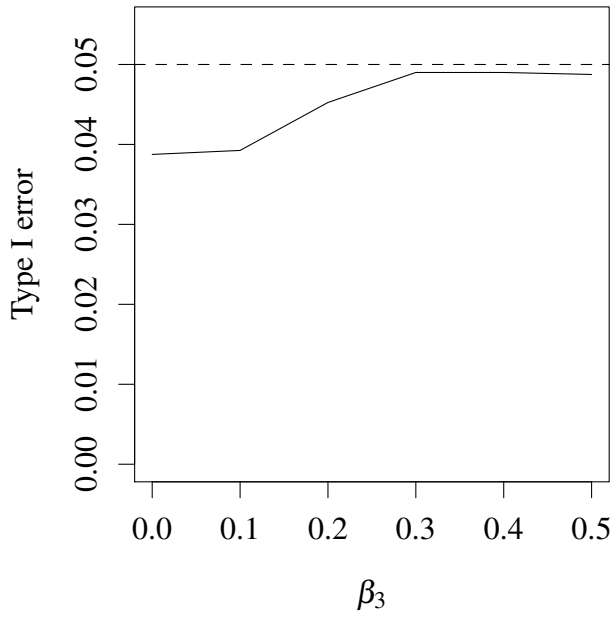

Figure C2: Control of type I error for the joint test with bilateral test for interaction. We simulated 4 000 datasets based on the linear model (2), i.e., we sampled ϵ for each simulation. The variable X is sampled once from a uniform distribution between -1 and 1. We take $n = 300$, $\beta_0 = 0$, $\beta_1 = 1$, and $\sigma^2 = 1$. We sampled 4 000 times from Π. The interaction coefficient β_3 varies between 0 and 0.5 by 0.1 increments with $\beta_2 = -\beta_3 \min_i(X_i)$. The dotted line shows $\alpha = 0.05$.

Proposition 4. *If we consider a unilateral test for interaction in the joint test for presence of improvement, then for all* $(\beta_2, \beta_3) \in B$ *we have* $\mathbb{P}_\beta\,(\text{Rejection}) \leq \alpha$.

The situation is not so straightforward in the bilateral case and we will not formally prove that we have control over type I error. Indeed, when the alternative hypothesis for the interaction is $H_1^{\text{interact}} : \beta_3 \neq 0$, the rejection region changes with the sign of $\hat{\beta}_3$. We have $R = \{-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} < q_{\alpha/2}, z_\Delta(x_0) < q_\alpha\}$ if $\hat{\beta}_3 > 0$ and $R = \{-|\hat{\beta}_3|/\sqrt{\Sigma_{3,3}} < q_{\alpha/2}, z_\Delta(x_1) > q_\alpha\}$ if $\hat{\beta}_3 < 0$. As $\hat{\beta}_3$ is not always of the sign of β_3, a formal proof would be tedious. We can nevertheless expect type I error to be controlled, and we give some arguments to support this statement. Indeed, the border region between the null hypothesis and the alternative becomes $B = \{(\beta_2, \beta_3), \beta_3 \geq 0, \beta_2 = -x_0\beta_3\} \cup \{(\beta_2, \beta_3), \beta_3 \leq 0, \beta_2 = -x_1\beta_3\}$. In the case where $\beta_3 = 0$, the test for interaction is clearly sufficient to control the type I error. If β_3 is large and positive, we have with great probability that $\hat{\beta}_2 > 0$ and $\hat{\beta}_3 > 0$. In this case, $z_\Delta(x_0) \sim \mathcal{N}(0, 1)$, and

$$\mathbb{P}_\beta\,(\text{Rejection}) \approx \mathbb{P}_\beta(z_\Delta(x_0) < q_\alpha) = \alpha.$$

A similar reasoning applies when β_3 is large and negative. Using simulations, we show in Figure C2 that this control is still attained for intermediate values between $\beta_3 = 0$ and β_3 positive and large.

C2 Coverage probabilities

We here check that the resulting quantiles respect the expected frequentist coverage probabilities, i.e., $\mathbb{P}_\beta(\Theta(\hat{\texttt{Mlb}}_\alpha) \leq \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha)) = \gamma$. Here $\gamma \in [0,1]$ is any confidence level, which does not have to be equal to α, the confidence level used in the definition of $\hat{\texttt{Mlb}}_\alpha$. This subsection will use simulations to check this, keeping in mind that the question of frequentist validity of Bayesian credible bounds is an active field of theoretical research (e.g., van der Vaart, 1998).

Coverage probabilities are complicated in our setting by the fact that with positive probability, $\hat{\texttt{Mlb}}_\alpha = \texttt{ref}$. In this case, the set of patients for whom the recommended treatment is the alternative treatment (i.e., the personalized set) is empty. As $\Theta^\star$ is an integral on the personalized set, we have $\Theta^\star(\texttt{ref}) = 0$, and the distribution of $\Theta^\star$ is a Dirac mass at 0. As $\Theta(\texttt{ref}) = 0$ as well, all quantiles are therefore correct in this case. This means that $\mathbb{P}_\beta(\Theta(\hat{\texttt{Mlb}}_\alpha) = \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha)) > 0$, and the best we can hope for is to have control over strict left coverage $\mathbb{P}_\beta(\Theta(\hat{\texttt{Mlb}}_\alpha) < \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha))$ and over strict right coverage $\mathbb{P}_\beta(\Theta(\hat{\texttt{Mlb}}_\alpha) > \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha))$. In the frequentist world, valid coverage translates to

$$\mathbb{P}_\beta\big(\Theta(\hat{\texttt{Mlb}}_\alpha) < \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha)\big) \leq \gamma$$

and

$$\mathbb{P}_\beta\big(\Theta(\hat{\texttt{Mlb}}_\alpha) > \hat{q}_{n,\gamma}(\hat{\texttt{Mlb}}_\alpha)\big) \leq 1 - \gamma.$$

To investigate if coverage is respected by our estimation, we simulated data in the same fashion as above. We simulated 10 000 datasets based on the linear model (2), i.e., we sampled ϵ for each simulation. The variable X is sampled once from a uniform distribution between -1 and 1. We take $n = 300$, $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 0.3$, $\sigma^2 = 1$, and $\beta_3 \in \{0.5, 0.8, 1\}$. We select 0.05 for α and we use $\hat{\texttt{Mlb}}_{0.05}$. We sampled $\beta^\star$ 10 000 times from Π.

Figure C3 shows the results for four thresholds, including $\alpha = 0.05$. Each bar corresponds to a set of simulations with fixed parameters and a choice of γ. It is divided in 3 with the length of each color corresponding (from bottom to top) to the probability of left coverage (blue), equality (green), or right coverage (red). Naturally, the total length of the bar is 1 as those three probabilities sum to 1. The line γ is drawn and the coverage inequalities hold if it is in the green. The main message of Figure C3 is that coverage is approximately respected. However, there is a clear asymmetry between left and right coverages. When β_3 decreases, the green bar corresponding to the probability of an empty personalized set increases. As that probability increases, the probability of right coverage (red) decreases while the probability of left coverage (blue) stays constant. It is only after the probability of right coverage is 0 that the probability of left coverage decreases with β_3.

C3 Influence of parameters

To have some grasp on the situation, we propose the visualization in Figure C4 in the parameters' plane. The confidence ellipse around $(\hat{\beta}_2, \hat{\beta}_3)$ represents Π, the posterior distribution of (β_2, β_3) from which we sample $(\beta_2^\star, \beta_3^\star)$. To each X_i corresponds a line $\Delta^\star(X_i) = \beta_2^\star + \beta_3^\star X_i = 0$. The lines that are under the true value of (β_2, β_3) correspond to participants for whom personalization is beneficial.

The ellipse controls uncertainty in two dimensions instead of one and therefore leads to wider confidence region when projected in one dimension. This is why some of the dotted lines intersect the 90% confidence ellipse despite corresponding to points for which $z_\Delta(X) < q_{0.05}$.

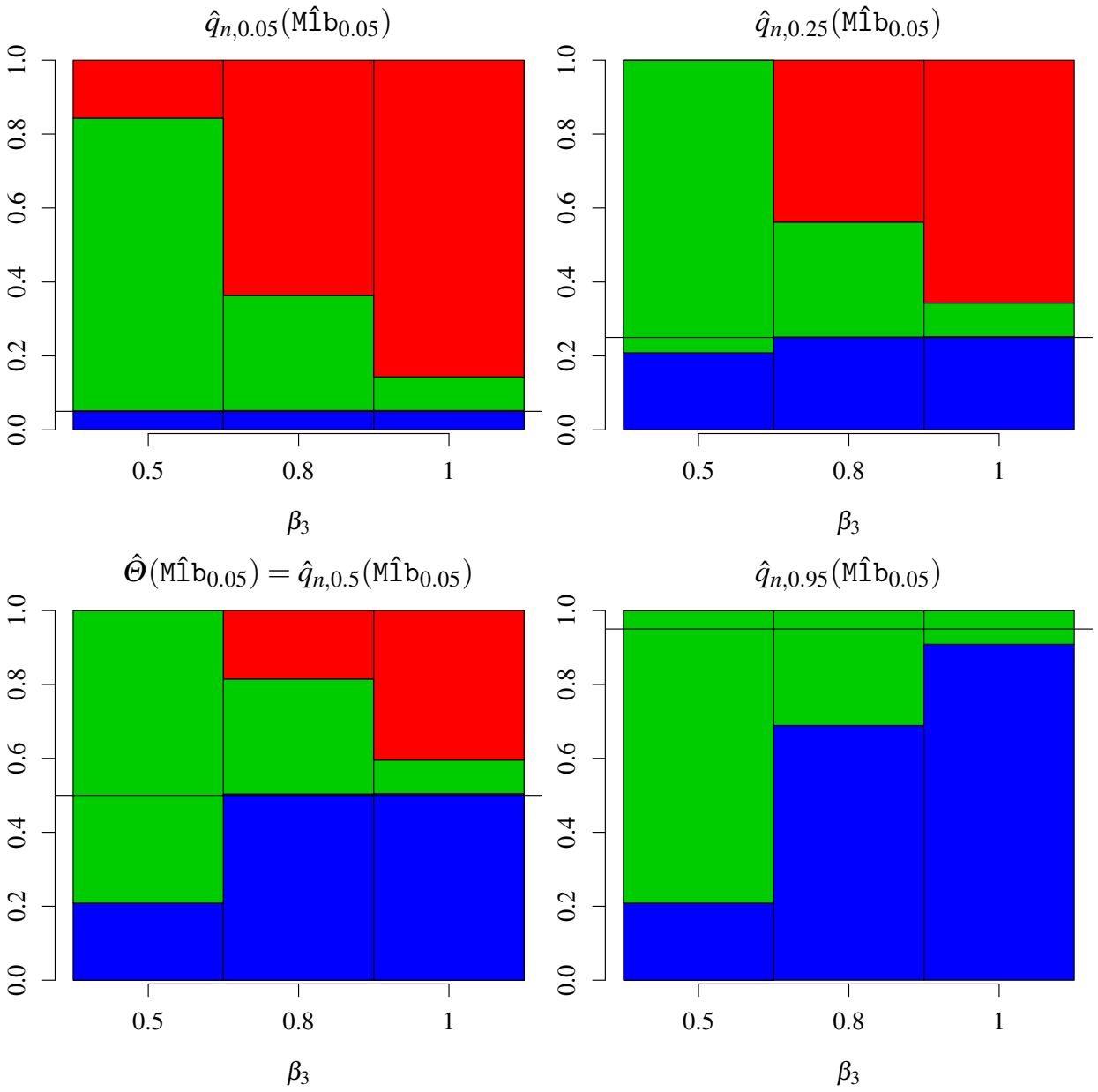

Figure C3: Coverage of $\Theta(\hat{\mathtt{Mlb}}_{0.05})$ by $\hat{q}_{n,\gamma}(\hat{\mathtt{Mlb}}_{0.05})$ for $\gamma \in \{0.05, 0.25, 0.5, 0.95\}$. The bar length corresponds to the probabilities of different events depending on the value of β_3. In each bar, from bottom to top, the length of the blue bar is the empirical probability of left coverage $\mathbb{P}_\beta(\Theta(\hat{\mathtt{Mlb}}_{0.05}) < \hat{q}_{n,\gamma}(\hat{\mathtt{Mlb}}_{0.05}))$. The green bar corresponds to $\mathbb{P}_\beta(\Theta(\hat{\mathtt{Mlb}}_{0.05}) = \hat{q}_{n,\gamma}(\hat{\mathtt{Mlb}}_{0.05}))$, i.e., $\mathbb{P}_\beta(\hat{\mathtt{Mlb}}_{0.05} = \mathtt{ref})$. The red bar corresponds to the probability of right coverage $\mathbb{P}_\beta(\Theta(\hat{\mathtt{Mlb}}_{0.05}) > \hat{q}_{n,\gamma}(\hat{\mathtt{Mlb}}_{0.05}))$. The γ line is drawn. As long as the line is in the green, coverage is respected. Note that the median $\hat{q}_{n,0.5}(\hat{\mathtt{Mlb}}_{0.05})$ corresponds to $\hat{\Theta}(\hat{\mathtt{Mlb}}_{0.05})$.

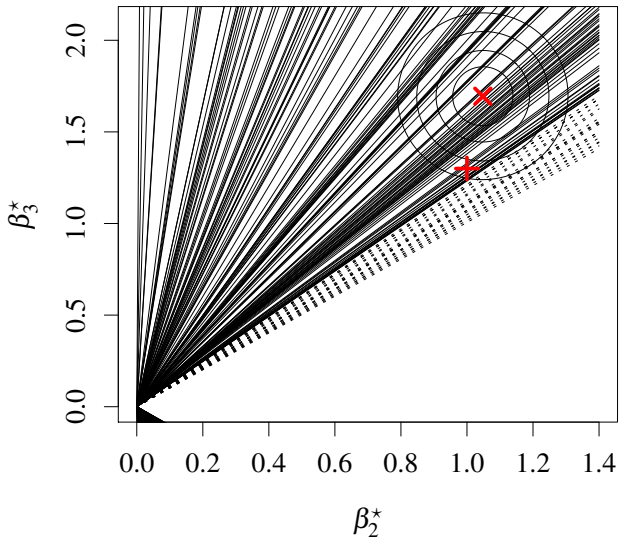

Figure C4: Confidence ellipse around $(\hat{\beta}_2, \hat{\beta}_3)$ (marked by a red x) at confidence level 25%, 50%, 75%, and 90%. The red cross marks the true value of the parameters (β_2, β_3). The lines have equation $x + yX_i = 0$ for all X_i in the sample. The dotted lines correspond to the observations that verify $z_\Delta(X_i) < \eta_{\max,0.05}$, i.e., the set of patients for which the alternative treatment is recommended under strategy $\hat{\texttt{ind}}_{0.05}$. The axis of the ellipse are parallel to the x-axis and y-axis because of the independence between X and T.

With this representation, it is quite easy to infer the influence of parameters on the discovery rate. As the sample size n increases, the ellipse will shrink and the estimated parameters will converge towards the true parameters at rate $1/\sqrt{n}$, which will lead to more frequent identification of a non-empty personalized set. When the treatment effect β_2 increases, there are less patients who benefit from personalization and therefore it becomes harder to identify them. If β_3 increases, the opposite happens. If both parameters increase while their ratio remains fixed, it will become easier to identify a personalized set because the lines will be more spread out.

Appendix D: Covariates used for the analysis of the RAVE data

Table D1: Covariates used for analysis of the RAVE data.

Variable
Delay between stroke and randomization in hours
Conscious state at randomization
Patient sex
Age in years
Symptoms noted on waking
Atrial fibrillation
CT before randomization
Infarct visible on CT
Heparin within 24 hours prior to randomization
Aspirin within 3 days prior to randomization
Systolic blood pressure at randomization (mmHg)
Face deficit
Arm/hand deficit
Leg/foot deficit
Dysphasia
Hemianopia
Visuospatial disorder
Brainstem/cerebellar signs
Other deficit
Stroke subtype
Local time - hours of randomization