Published in final edited form as:

J Am Stat Assoc. 2021 ; 116(535): 1140–1154. doi:10.1080/01621459.2020.1828091.

High-Dimensional Precision Medicine From Patient-Derived Xenografts

Naim U. Rashid^{a,b}, Daniel J. Lockett^a, Jingxiang Chen^a, Michael T. Lawson^a, Longshaokan Wang^g, Yunshu Zhang^g, Eric B. Laber^g, Yufeng Liu^{a,c,d}, Jen Jen Yeh^{b,e,f}, Donglin Zeng^a, Michael R. Kosorok^{a,c}

^aDepartment of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC

^bLineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

^cDepartment of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC

^dDepartment of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC

^eDepartment of Surgery, University of North Carolina at Chapel Hill, Chapel Hill, NC

^fDepartment of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

^gDepartment of Statistics, North Carolina State University, Raleigh, NC, USA

Abstract

The complexity of human cancer often results in significant heterogeneity in response to treatment. Precision medicine offers the potential to improve patient outcomes by leveraging this heterogeneity. Individualized treatment rules (ITRs) formalize precision medicine as maps from the patient covariate space into the space of allowable treatments. The optimal ITR is that which maximizes the mean of a clinical outcome in a population of interest. Patient-derived xenograft (PDX) studies permit the evaluation of multiple treatments within a single tumor, and thus are ideally suited for estimating optimal ITRs. PDX data are characterized by correlated outcomes, a high-dimensional feature space, and a large number of treatments. Here we explore machine learning methods for estimating optimal ITRs from PDX data. We analyze data from a large PDX study to identify biomarkers that are informative for developing personalized treatment recommendations in multiple cancers. We estimate optimal ITRs using regression-based (Q-learning) and direct-search methods (outcome weighted learning). Finally, we implement a superlearner approach to combine multiple estimated ITRs and show that the resulting ITR performs better than any of the input ITRs, mitigating uncertainty regarding user choice. Our results indicate that PDX data are a valuable resource for developing individualized treatment strategies in oncology. Supplementary materials for this article are available online.

CONTACT Michael R. Kosorok kosorok@bios.unc.edu Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

Supplementary Materials

Supplementary Materials: Document containing supplementary figures, supplementary tables, supplementary methods, supplementary results, and R code implementing the methodology from this article. (PDF)

Keywords

Biomarkers; Deep learning autoencoders; Machine learning; Outcome weighted learning; Precision medicine; Q-learning

1. Introduction

The complexity of human cancer is reflected in the molecular and phenotypic diversity of patient tumors (Polyak 2011). This diversity results in heterogeneity in response to treatment, which complicates clinical decision making. The complexity of human cancer is also reflected in the high failure rate of new therapies entering oncology clinical trials, highlighting limitations in the ability of standard preclinical models to evaluate new therapies (Tentler et al. 2012). A recent study utilized patient-derived xenografts (PDXs) to perform a large-scale screening in mice to evaluate a large number of FDA-approved and preclinical cancer therapies (Gao et al. 2015). Genomic information and observed treatment responses were used to identify efficacious therapies that standard cell line model systems had missed, and also validate known associations between genomic biomarkers and differential response to treatment. The results from this PDX study mirrored those seen in human patients. Thus, PDX models can be used to evaluate in vivo therapeutic response and discover novel biomarkers to inform individualized treatment decisions.

PDX models are based on the transfer of primary human tumors directly from the patient into an immunodeficient mouse (Siolas and Hannon 2013). Briefly, pieces of primary solid tumors are collected from patients by surgery or biopsy (Hidalgo et al. 2014). The collected tumor pieces from an individual patient are then implanted into mice subcutaneously to create a PDX line, whereby tumor size and rate of tumor growth after implantation may be measured over time. After the tumor reaches sufficient size, the line may be expanded by passaging directly from the implanted tumor into additional genetically identical mice. Through this expansion, multiple treatments may be applied to mice originating from the same PDX line, allowing for the application of multiple treatments to the same patient tumor. High throughput genomic assays such as RNA sequencing (RNA-seq) and DNA sequencing may be performed on the original patient tumor. Features of the original tumor will be retained throughout line expansions (Hidalgo et al. 2014), making PDX models ideal for learning how to personalize cancer treatment, given observed feature-response associations.

Personalized treatment recommendations in oncology have traditionally centered around the classification of patients into subgroups (Sargent et al. 2005). In some cases, patient subgroups may be derived from predictive models based upon genomic biomarkers (Parker et al. 2009). Yet, significant heterogeneity in treatment response may still be observed within such subgroups (Metzger-Filho et al. 2012; Chen et al. 2014), and assignment of optimal treatment is predicated upon accurate subgroup assignment. An alternative approach to precision medicine is the estimation of an individualized treatment rule (ITR), a map directly from the patient covariate space to the space of allowable treatments that can be used to make treatment decisions. The optimal ITR is defined as the one that maximizes

the mean of a clinical outcome, such as treatment response, when applied in a population. Examples of such covariates may include patient clinical information, such as laboratory assay results, or high dimensional genomic data, such as gene expression or mutation data from a patient's tumor. As such, treatment recommendations based upon an optimal ITR may result in improved clinical outcomes by harnessing individual-specific molecular and clinical features not captured by subgroup-based approaches.

A number of methods have been proposed to estimate an optimal ITR. One approach is to fit a regression model for treatment response given a set of applied treatments and patient covariate information. The optimal treatment is then the one that provides the maximum predicted response in a new patient, given the fitted regression model and the covariate information for that patient (Qian and Murphy 2011). An example of this approach is Q-learning (Murphy 2005; Zhao, Kosorok, and Zeng 2009; Qian and Murphy 2011; Schulte et al. 2014). Direct search methods, including outcome weighted learning (OWL) (Zhao et al. 2012; Liu et al. 2016; Chen et al. 2017; Zhou et al. 2017), doubly robust ITR estimators (Zhang et al. 2012, 2013), and marginal structural models (Robins, Orellana, and Rotnitzky 2008; Orellana, Rotnitzky, and Robins 2010) estimate the optimal ITR using inverse probability weighting rather than regression. In direct search methods, the class of ITRs is prespecified, while in other approaches, the class of ITRs is implied by the modeling process. Recent advances in machine learning for causal inference have produced a number of estimators for the conditional average treatment effect that could be used to estimate an ITR for a binary treatment decision (Imai and Ratkovic 2013; Athey and Imbens 2016; Wager and Athey 2018). Other methods for estimating optimal ITRs include marginal mean models (Murphy et al. 2001) and Bayesian predictive modeling (Ma, Stingo, and Hobbs 2016). Regardless of the approach, many existing methods may directly utilize high-dimensional genomic biomarker data. However, such methods were not designed for PDX studies, where the application of multiple treatments within a subject (PDX line) results in correlated outcomes, and the number of available treatments is large.

In this article, we utilize the wealth of clinical and biomarker data generated by the Novartis PDX study (Gao et al. 2015) to estimate optimal ITRs for treating several human cancers. Given the correlated outcomes within PDX lines, the large number of treatments, and the high dimension of the covariate space, it is difficult to fit a nonparametric model to the conditional mean response without large amounts of data. Thus, we explore a number of ways of imposing structure on the conditional mean response, including reducing the dimension of the covariate space, grouping treatments that result in a similar mean response, and constructing a treatment tree to convert the problem of selecting from a large set of treatments to a sequence of binary comparisons. The result is a multi-step procedure, where each step can be thought of as imposing structure on the model for the conditional mean response in a way that alleviates the challenges posed by PDX data and takes advantage of the unique structure of PDX data. We show that the proposed multi-step procedure achieves improved performance over standard methods in certain settings. We examine the various modeling decisions that are made at each step of the multi-step procedure and demonstrate the use of super-learning (Luedtke and van der Laan 2016) to improve prediction performance by aggregating the proposed models.

In Section 2, we describe the large-scale PDX dataset that motivated this work and use it to highlight the challenges associated with estimating optimal ITRs using PDX data. We present our methodological approaches in Section 3. We present results from our data analyses in Section 4. In Section 5, we conclude with a discussion to compare and contrast the various modeling decisions we made and discuss the clinical implications of our findings. Additional details, including software for reproducing our work, are given in the supplementary materials.

2. Large-Scale PDX Drug Screen for Treatment Response

2.1. Data Overview

The Novartis PDX study (Gao et al. 2015) established a total of 1075 PDX lines corresponding to a variety of human cancers. A subset of these lines were genomically profiled prior to treatment for gene expression (399 lines), copy number analysis (375), and mutations (399). In addition, 281 lines were enrolled in the drug response study. Our study utilizes 190 PDX lines with complete genomic and response data (Supplementary Figure 1). Five types of cancer are represented among these 190 lines (Supplementary Figure 2): breast cancer (BRCA), melanoma (CM), colorectal cancer (CRC), non-small cell lung cancer (NSCLC), and pancreatic cancer (PDAC). A median of 21 treatments were applied per PDX line across different cancers (Supplementary Figure 2). Certain PDX lines had fewer mice and therefore fewer treatments than the total number of treatments available for a particular cancer type. One mouse per PDX line was set as a control line and did not receive treatment.

In total, 3487 mice, expanded from the 190 PDX lines with complete data, were used for our study. Details regarding the biomarker data utilized in this study are given in Section 1 of the supplementary materials. Briefly, each of the 190 PDX lines had 22,665 genes measured for gene expression via RNA-seq, 23,854 genes measured for gene-level copy number estimates via copy number array, and between 159 and 293 mutations (25th and 75th percentiles) identified via DNA sequencing. In total, a union set of approximately 60,000 features are available for ITR estimation. Because each genomic assay was performed on the patient tumor prior to implantation, all mice expanded from the same PDX line share the same set of genomic biomarkers.

2.2. Study Design and Response Variables

A total of 38 unique therapies were applied in this study, administered to mice as either a single agent (36 total administered) or in combination with other agents (26 total combinations administered). Certain treatments were limited to particular cancers, whereas others were applied across cancers. Each mouse was treated for a minimum of 21 days. Tumor size was evaluated twice weekly by caliper measurements, and the approximate volume of the tumor was calculated using the formula $(l \times w \times w) \times (\pi/6)$, where l is the major tumor axis and w is the minor tumor axis.

Two measures were used to summarize response to treatment: best average response (BAR) and time to tumor doubling (TTD). BAR is defined as

$\min_{t: d_t > 10} \left[\{1 / (t + 1)\} \sum_{l=0}^t (V_l - V_0) / V_0 \right] \times 100 \%$, where d_t is the day in which the t th measurement was taken, V_0 is tumor volume at day 0, and V_l is tumor volume at measurement l taken on day d_l (Gao et al. 2015). BAR is a measure of the maximum observed tumor shrinkage from baseline over measurements taken at least 10 days after start of treatment, scaled by time since baseline. More negative values of BAR indicate a better response. We used $-BAR$ in the analysis so that larger values indicate a better response. BAR mirrors similar criteria for assessing response in human clinical trials (RECIST, Therasse et al. 2000). TTD is defined as the number of days from baseline that the tumor doubled in size from its baseline measurement. Due to skewness in the distribution of TTD, we used the natural log for analysis.

2.3. Implications of PDX Data for ITR Estimation

The unique structure of PDX data—the application of multiple treatments to mice implanted with the same tumor—makes PDX data ideally suited for estimating optimal ITRs. We note that comparing treatment responses between mice within the same line does not amount to observing true treatment effects due to the inherent variability that exists across mice; however, the improved precision of observing responses to multiple treatments applied to the same tumor may substantially improve the performance of estimated ITRs, which existing methods are not designed to leverage. Furthermore, existing methods for estimating optimal ITRs have typically been designed for a small number of treatments (e.g., two). In this case, the set of treatments is large (20 or greater). In such a setting, modeling the conditional response is difficult due to the large number of treatment $\times$ feature interaction terms that would be included in the model. Finally, the set of genomic biomarkers is high-dimensional. Approximately 60,000 genomic features are available, and many of the available features may exhibit either low variability or low expression across PDX lines. Combined, these issues present a unique challenge to the estimation of the optimal ITR in PDX studies.

To address these issues, in the next section we propose methods to reduce the dimension of the feature space, adaptively group treatments with similar effects across PDX lines using hierarchical clustering, and develop tree-based extensions of Q-learning and OWL to estimate the optimal ITR. Our proposed approach directly leverages responses to multiple treatments within each tumor, a characteristic that would not be available without PDX data.

3. Methods

3.1. Overview

Let Y denote response and let $\mathbf{X}$ denote a vector of covariates. Let $A = (A_1, \dots, A_J)$ denote treatment, where exactly one of $A_1, \dots, A_J$ is equal to 1 and the rest equal to 0, with $A_j = 1$ indicating that treatment j is given. The conditional mean response can be expressed as

$$E(Y | \mathbf{X} = \mathbf{x}, A = a) = h_0(\mathbf{x}) + \sum_{j=1}^J h_j(\mathbf{x}) a_j. \quad (1)$$

If we were to obtain estimators $\hat{h}_j, j = 1, \dots, J$, an estimator for the optimal ITR would be $\hat{d}(\mathbf{x}) = \arg \max_{j=1, \dots, J} \hat{h}_j(\mathbf{x})$. However, fitting model (1) nonparametrically in the PDX setting would be difficult without large amounts of data, due to the large number of treatments and high dimension of the covariate space. Therefore, we propose a multi-step procedure to impose structure on model (1) to ameliorate these difficulties, described in Sections 3.1.1–3.1.3 and illustrated in Supplementary Figure 3. Specifically, our method involves screening features to find a lower dimensional feature space, constructing a treatment tree that allows for recasting ITR estimation as a sequence of binary decisions, and estimating a decision rule at each split to select the arm that contains the optimal treatment. Various modeling decisions must be made at each step, and some alternative choices for these decisions are discussed in Section 3.3. Because it is not immediately obvious what the optimal set of modeling decisions is, we apply super-learning (Luedtke and van der Laan 2016) to combine variants of the proposed method with different embedded models. For complete details on our methodology, we refer the reader to Section 2 of the supplementary materials.

3.1.1. Preprocessing (Step 1)—In the first step, we preprocess genomic features such that we remove those features without sufficient variance (Section 2.1 in the supplementary materials) or low expression across PDX lines. This is a well-studied technique for dimension reduction prior to analysis (Bourgon, Gentleman, and Huber 2010; Love, Huber, and Anders 2014; Rashid et al. 2014). This screening is performed separately for each cancer type. A summary of the number of genes and features remaining after this step is given in Supplementary Table 2. In addition, treatments applied in less than 90% of PDX lines within a cancer type were filtered out for that cancer type (see Section 2.1 of the supplementary materials). We summarize the number of treatments applied per PDX line per cancer after treatment filtering in Supplementary Table 1.

3.1.2. Supervised Screening (Step 2)—In the second step, we further reduce the dimension of the feature space by selecting likely prognostic and predictive genomic features (Section 2.2 in the supplementary materials). Genes are ranked using Brownian distance covariance (Székely and Rizzo 2009), evaluating the dependence between a vector of gene-level predictors for each PDX line (expression, copy number, and mutation) and the bivariate response (BAR, TTD) for a given treatment (prognostic) or difference in response between a pair of treatments (predictive). After ranking genes, we select the features pertaining to the top L_{SUP} genes ($L_{\text{SUP}} = 50, 100, 500, 1000$, Supplementary Table 2) and use all available platforms for the selected genes, giving us $p_{L_{\text{SUP}}}$ features corresponding to each value of L_{SUP} . Screening in this way imposes structure on model (1) by forcing the $h_j(\mathbf{x})$ to be constant in some of the covariates for certain $j = 0, \dots, J$. This strategy helps to alleviate the difficulties caused by the large number of genomic features and is similar in nature to sure independence screening (Fan and Lv 2008), used in ultra-high-dimensional regression problems to reduce the feature dimension to a more moderate size prior to application of variable selection methods. All of our analyses are repeated for each of $L_{\text{SUP}} = 50, 100, 500, 1000$ top genes to yield insights into the performance of estimated ITRs based on differing numbers of features. Screening is performed separately for each cancer type.

We also applied a further dimension reduction step using deep learning autoencoders (DAE, Wang and Laber 2017), a variant of deep neural networks (Vincent et al. 2010), to evaluate whether utilizing a lower dimensional representation of the selected feature set (indexed by $p_{L_{SUP}}$) improves ITR estimation at the cost of additional computational burden. DAEs build a nonlinear prediction model to predict feature variables from a low-dimensional representation of the original features, with dimension chosen by cross-validation. We find that the reconstruction error of this approach is several orders of magnitude lower than principal components analysis across cancer types, particularly for $L_{SUP} = 50, 100$ (Supplementary Tables 3 and 4). This indicates that linear dimension reduction techniques are not sufficient to capture the information contained in the data. Dimension reduction using DAEs imposes structure on model (1) by forcing the $h_j(\mathbf{x}), j = 0, \dots, J$ to depend on $\mathbf{x}$ only through the low-dimensional summary of the original features. The dimensions for each feature set following application of deep learning autoencoders within each cancer are given in Supplementary Table 5. After applying Steps 1 and 2 to obtain low-dimensional sets of covariates, we may optionally use the low-dimensional sets of covariates to estimate optimal ITRs. See Section 2.3 of the supplementary materials for details.

3.2. Estimation of Treatment Tree-Based ITRs

Here we describe a technique for grouping treatments using hierarchical clustering, allowing us to learn ITRs under reduced treatment spaces and borrow strength from similar treatments. Different groupings of treatments may be created by cutting a “treatment tree,” the dendrogram resulting from the clustering, in different places. We then define a class of ITRs that can be expressed as a sequence of binary decisions, one for each step of the tree. This tree structure is distinct from tree-based regimes which use a tree structure to represent the final estimated ITR (Laber and Zhao 2015; Zhang et al. 2015). A variety of methods could be used to construct decision rules for each step of the tree. We introduce two: a regression-based approach (Section 3.2.3) and a direct search approach (Section 3.2.4).

3.2.1. Treatment Tree Construction—Denote the i th mouse corresponding to the j th PDX line within the k th cancer type with subscript ijk , where $k = 1, \dots, 5, j = 1, \dots, m_k$, and $i = 1, \dots, p_{jk}$, with p_{jk} representing the number of treatments applied to PDX line j in cancer type k , and m_k representing the number of PDX lines for cancer type k (Supplementary Figure 2). We let $P_k = \max_{j=1, \dots, m_k} p_{jk}$, that is, for each PDX line, up to P_k mice were expanded to receive a maximum of P_k treatments in cancer type k , one treatment applied per mouse (Supplementary Figure 4). For certain lines, $p_{jk} < P_k$ treatments may have been applied, as the number of mice per line varied. For ease of notation, we assume that $p_{jk} = P_k$ and that the i th mouse in each PDX line received the same treatment within cancer type.

For treatment i in cancer k , we define the treatment response vector $Y_{ik} = (Y_{i1k}, \dots, Y_{im_kk})^T$, where Y_{ik} is scaled to have standard deviation 1 (rows in Supplementary Figure 5, left panel). Because there are inherent baseline differences in response between PDX lines, we then center the response values of each PDX line (columns in Supplementary Figure 5, left panel) with respect to the “null” response within each PDX line, which we calculate in the following manner. Using PDX lines with P_k mice, we calculate the Euclidean distances between Y_{ik} and $Y_{i'k}, i, i' = 1, \dots, P_k, i \neq i'$ to get a distance between each pair of

treatments. For a fixed constant, c_1 , we group the c_1 nearest neighbors of the “untreated” treatment response vector to form a “null” set of treatments, denoted by $A_{k,c_1,0}$, containing those treatments producing low or no response. Then, for each $i \notin A_{k,c_1,0}$, we compute the $m_k \times 1$ vector of centered observed outcomes, $R_{ik} = Y_{ik} - \bar{Y}_{A_{k,c_1,0}}$, where $\bar{Y}_{A_{k,c_1,0}}$ is the $m_k \times 1$ vector of averaged treatment responses for each of the treatments in $A_{k,c_1,0}$. Thus, R_{ik} is the difference between the response to treatment i and the average response in the null treatment group.

Next, we perform hierarchical clustering on the centered treatment response vectors and take the resulting dendrogram as a treatment tree (Supplementary Figure 5, right panel). For a fixed constant, c_2 , we can construct treatment groups by cutting the tree c_2 steps from the root node. For cancer type k , we label the treatment groups determined by c_1 and c_2 as $A_{k,c_1,1}, \dots, A_{k,c_1,c_2+1}$ and denote the set of all possible treatment groups in cancer k given (c_1, c_2) by $\mathcal{A}_{k,c_1,c_2} = \bigcup_{a=1}^{c_2+1} A_{k,c_1,a} \cup A_{k,c_1,0}$. Although $\mathcal{A}_{k,c_1,c_2}$ depends on c_1 and c_2 , we will write $\mathcal{A}$ in place of $\mathcal{A}_{k,c_1,c_2}$ to simplify notation. The hierarchical clustering groups treatments together when they result in similar mean responses, in contrast to grouping treatments by predefined characteristics such as molecular target or mechanism of action. Grouping treatments in this way forces certain $h_j, j = 1, \dots, J$ in model (1) to be equal, similar to a fusion penalty (Tibshirani et al. 2005). This strategy helps to alleviate the difficulties caused by the large number of treatments and allows the estimated ITRs to select a group of treatments likely to produce similar outcomes (Wu 2016). We average responses within the resulting treatment groups, yielding one response for each PDX line and treatment group in $\mathcal{A}$. An ITR can be estimated for each value of c_1 and c_2 , and the optimal values of c_1 and c_2 can be selected using cross-validation.

3.2.2. Identification of the Optimal ITR—Let $X \in \widehat{\mathcal{X}} \subset \mathbb{R}^{PLSUP}$ be the vector of genomic features, where we write $\widehat{\mathcal{X}}$ to make it clear that the domain of the feature space is data-dependent (see Sections 3.1.1 and 3.1.2). An ITR is a mapping $D: \widehat{\mathcal{X}} \rightarrow \mathcal{A}$. For each treatment $a \in \mathcal{A}$, define the potential outcome $R^*(a)$ to be the outcome that would be observed under treatment a (Rubin 1978). Within cancer type k , the set of potential outcomes consists of $R_{ij}^*(a_i), i = 1, \dots, P_k, j = 1, \dots, m_k$. We make the assumption that $E\{R_{ij}^*(a)\} = E\{R_{i'j}^*(a)\}$ for all $i, i' = 1, \dots, P_k$, all $j = 1, \dots, m_k$, and all $a \in \mathcal{A}$. That is, we assume that the expected value of the potential outcomes for mouse i and i' from the same PDX line would be equal if they both received the same treatment. The standard assumption of positivity is not needed because each PDX line is assigned to every treatment used for that cancer type. On the other hand, while the standard assumption of no unmeasured confounders is still needed in our setting, the primary source of confounding is due to the assignment of mice to treatment within PDX line and not due to the genetic features of the PDX line since each line receives all treatments. Hence, we assume that the process of assigning treatment to mice is exchangeable and homogeneous, and thus the no unmeasured confounding assumption obtains. Define the value of an ITR, D , by $V(D) = E(E[R^*\{D(X)\} | X])$. Let $\mathcal{D}$ be the set of ITRs which can be expressed as a sequence of decision rules, one for each step of the treatment tree, starting from the root node and proceeding until a leaf

node is reached. The optimal ITR associated with the treatment tree is $D^* = \arg \max_{D \in \mathcal{D}} V(D)$,

that is, the ITR that maximizes value within the class.

3.2.3. Treatment Tree-Based Q-Learning—In this section, we propose an extension of Q-learning (Murphy 2005; Zhao, Kosorok, and Zeng 2009; Zhao et al. 2011; Schulte et al. 2014) to estimate the optimal ITR associated with the treatment tree. Let $a_w(t)$, $w = 0, 1$ denote the set of treatment groups downstream to the left ($a_1(t)$) and right ($a_0(t)$) arms of a node at step t , where $t = c_2, \dots, 0$. Starting at the step corresponding to the lowest node, we compute $R_{jw}(c_2) = \left[\sum_{i=1}^{P_k} R_{ij} I \{i \in a_w(c_2)\} \right] / \sum_{i=1}^{P_k} I \{i \in a_w(c_2)\}$ for $j = 1, \dots, m_k$ and $w = 0, 1$. Here $R_{jw}(c_2)$ is the mean observed reward (centered response) in PDX line j across the treatments belonging to treatment group $a_w(c_2)$. We then fit a regression model for $R_{jw}(c_2)$ based on $\mathbf{X}_{jk}$, separately for $w = 0$ and $w = 1$, to obtain $\hat{E}_{c_2}[R^* \{a_w(c_2)\} \mid \mathbf{X} = \mathbf{x}]$, the estimated conditional mean reward in treatment group w at step c_2 given genomic features. A number of techniques could be used to fit the regression models at each stage, representing different ways of imposing structure on the h_j , $j = 0, \dots, J$ in model (1). We discuss various embedded models that could be used in more detail in Section 3.3. If either $R_{j0}(c_2)$ or $R_{j1}(c_2)$ are missing for a given line, then that line did not receive that particular treatment; these observations are removed before fitting the regression model. The estimated optimal decision rule at step c_2 of the tree for an individual with genomic features $\mathbf{x}$ is given by

$$\hat{D}_{c_2}^{\text{QL}}(\mathbf{x}) = \arg \max_{w \in \{0, 1\}} \hat{E}_{c_2}[R^* \{a_w(c_2)\} \mid \mathbf{X} = \mathbf{x}]. \quad (2)$$

We repeat the above process for step $t = c_2 - 1, \dots, 1$, except $R_{jw}(t)$ only uses the observations for each PDX line from the optimal treatment group selected at the previous step.

Step $t = 1$ corresponds to the highest node in the treatment tree consisting of the nonnull treatments, and step $t = 0$ corresponds to the split between the null group (centered by their own means) and the nonnull treatments from the treatment tree. The decision rule is determined in a similar manner at step $t = 0$, evaluating whether any nonnull treatment should be applied. At $t = 0$, the response vector for the null group is a vector of zeros, and the decision rule at this step simplifies to determining whether the expected value of the response under the optimal nonnull treatment is greater than 0.

Given the sequence of estimated decision rules $(\hat{D}_0^{\text{QL}}, \dots, \hat{D}_{c_2}^{\text{QL}})$, the optimal treatment for a new individual given their set of predictors is obtained by following the decision rules sequentially from step $t = 0$ downward until one arrives at a terminal node on the tree. Rather than recommending a sequence of treatments, as in standard Q-learning, the estimated optimal decision rule at each step of the tree directs the user through either the left or right arm at each node, creating a path through the tree to arrive at a terminal node representing the optimal treatment group for that individual.

3.2.4. Treatment Tree-Based OWL—OWL estimates the optimal ITR for selecting between two treatments by maximizing an inverse probability weighted estimator of the value function over a fixed class of decision rules. Thus, unlike Q-learning, OWL does not rely on a fitted regression model. Our extension of OWL adopts the same tree structure as in the previous section. However, instead of fitting a separate regression model for each arm at step t , the optimal decision rule at step t is estimated directly using weighted classification methods (Liu et al. 2016; Chen et al. 2017; Zhou et al. 2017). When treatment is binary, any decision rule can be written as $D(X) = \text{sign}\{f(X)\}$ for some decision function, f . We will assume that the decision boundary at each step is smooth. Thus, at each step, we use a class of decision rules defined by some class of smooth functions $\mathcal{F}$, for example, a reproducing kernel Hilbert space. At the n th step of the treatment tree, let $\hat{D}_t^{\text{OWL}}(x) = \text{sign}\{\hat{f}_t(x)\}$, where

$$\hat{f}_t = \arg \min_{f \in \mathcal{F}} \mathbb{E}_n \left(|R| [I(R \geq 0)\{1 - Af(X)\}_+ + I(R < 0)\{1 + Af(X)\}_+] \right) + \lambda_n J(f). \quad (3)$$

Here, $\mathbb{E}_n$ denotes the empirical measure of the data used in the n th step, $J(f)$ is a penalty term for the decision function f , λ_n is a tuning parameter which is selected using cross-validation, and $\mathcal{F}$ is a space of functions. As in Q-learning, the observations included when computing the minimizer in Equation (3) are all of those corresponding to the estimated optimal treatment group at the previous step. Thus, our extension of OWL also directly uses the multiple outcomes observed per PDX line. We discuss options for selecting $\mathcal{F}$ in Section 3.3; different options for $\mathcal{F}$ impose different specific forms that the $h_j, j = 0, \dots, J$ must take in model (1). A theoretical justification of our tree-based ITR approach is given in Section 2.5 of the Supplemental Material.

3.3. Modeling Decisions

Our proposed framework involves a number of specific modeling decisions that must be made prior to implementation. In this section, we describe different variants of the proposed method that result from making different modeling decisions.

The tree-based ITR estimation method involves an embedded method to estimate the optimal decision rule at each step of the tree. Fitting a regression model at each step results in an extension of Q-learning (see Section 3.2.3), where a number of regression models could be used as the embedded regression model. In our analyses, we used linear models with a LASSO penalty and random forests to perform this regression at each step. In Section 4, these are referred to as QL and QL_{RF}, respectively. At each split, our extension of Q-learning uses the maximum outcome across treatments downstream to the left and right of the split as the observed outcomes when fitting the regression model. Alternatively, we could obtain predicted maximum outcomes using the regression model and use the predicted maximum outcomes to fit the model at the next step. These are analogous to the pseudo-values used in standard versions of Q-learning (Zhao, Kosorok, and Zeng 2009; Zhao et al. 2011), and would allow application of the proposed tree-based Q-learning method to data that do not come from a PDX study. We examine both strategies in Section 4, and we refer to Q-learning with observed outcomes and with pseudo-values as QL₁ and QL₂, respectively.

Our tree-based OWL method alternatively constructs an inverse probability weighted estimator (IPWE) at each step (see Section 3.2.4), rather than fitting a regression model (as in Q-learning). The IPWE is maximized over a class of functions, where we consider both the class of linear functions and the nonlinear reproducing kernel Hilbert space associated with the Gaussian kernel function (Zhao et al. 2012; Chen et al. 2017). In Section 4, these are referred to as OWL_{linear} and OWL_{kernel} , respectively.

The second modeling decision that must be made is the selection of the covariate set. We applied the proposed method using the $p_{L_{\text{SUP}}}$ dimensional set of genomic features resulting from screening using Brownian distance covariance, for each $L_{\text{SUP}} = 50, 100, 500, 1000$ (Supplementary Table 2). We also applied the proposed method to the lower dimensional set of features extracted from L_{SUP} genes using the DAE, for each $L_{\text{SUP}} = 50, 100, 500, 1000$ (Supplementary Table 5). Analyses using the features extracted from the DAE are labeled with the subscript “dl” in Section 4.

A final modeling decision that we explored involves replacing the observed outcomes with model-predicted outcomes prior to estimation (“smoothing”). We fit a random forest to predict outcomes based on covariates alone (features and treatments) and replaced the observed outcomes with predictions based on the fitted model for all later stages of the analysis. This approach acts to denoise the observed outcomes. Analyses using this approach are labeled with the subscript “smoothed.” In Section 4, we report analyses using different combinations of the modeling decisions described above to capture synergistic effects of the various modeling decisions. In addition to estimating tree-based ITRs as proposed in Section 3.2.1, we also estimated ITRs by fitting model 1 using linear models with the LASSO penalty and random forests. These “off-the-shelf” methods were included to compare to the proposed method.

3.3.1. Super-Learning—The optimal choice of methods or modeling decisions may not always be clear a priori, and certain variants may perform better than others in certain settings. A natural analysis to try in this case is to combine a set of input ITRs estimated using various embedded models in the hopes that the resulting ITR performs better than any of the input ITRs. To accomplish this, we apply the super-learning algorithm of Luedtke and van der Laan (2016) to combine a number of existing methods using cross-validation to calculate a linear combination of latent functions to maximize the value function. In our implementation, there is no explicit latent function due to the fact that we are using a sequential treatment tree as our model. To approximate the value of the latent function with respect to a single treatment, we use the predicted reward from one of our sub-models at a given node in the tree whose direct children include our goal treatment. To optimize the superlearner, we use simulated annealing to estimate the coefficients. Multiple chains are used with starting points selected from a set of randomly generated coefficients. In Section 4, variants of super-learning with different sets of input ITRs are referred to with the subscript “SL,” followed by the number of ITRs that are combined to produce the superlearner.

3.4. Performance Measures for ITR

We used 5-fold cross-validation to evaluate the performance of ITRs estimated using the proposed method with the various modeling decisions as described in Section 3.3. Within each cancer type, we divided PDX lines into 5 folds. An optimal ITR was estimated using the training dataset that results from holding out each fold, and the estimated value was calculated on the held-out fold as $\mathbb{E}_n[RI\{A = \hat{D}^*(X)\}] / \mathbb{E}_n[I\{A = \hat{D}^*(X)\}]$, where $\mathbb{E}_n$ denotes the empirical measure taken over mice in the held-out fold. Tuning parameters (such as c_1 and c_2 for estimating tree-based ITRs and the OWL penalty parameter) were selected using cross-validation within each training dataset. The estimated value of the estimated ITR, denoted by $\bar{V}(\hat{D}^*)$ is calculated by averaging the value estimates resulting from each fold (“mean value”). We also calculated the standard deviation of value estimates across folds.

Different cancer types may result in different marginal mean outcomes. To facilitate comparisons across cancer types, we also computed the observed value, V_{obs} , defined as the sample average of centered responses for all mice that received nonnull treatments, and the optimal value, V_{opt} , defined as the sample average across PDX lines of the maximum centered response across treatments, that is, $V_{\text{opt}} = m_k^{-1} \sum_{j=1}^{m_k} \max_{i=1, \dots, P_k} R_{ijk}$. The observed value and optimal value can be used to define two metrics for evaluating estimated ITRs: proportion of optimal value, defined as $P_{\text{opt}}(\hat{D}^*) = \bar{V}(\hat{D}^*) / V_{\text{opt}}$, and ratio to observed value, defined as $P_{\text{obs}}(\hat{D}^*) = \bar{V}(\hat{D}^*) / V_{\text{obs}}$.

4. Results

All analyses were performed separately for each cancer type. We applied various combinations of the modeling decisions outlined in Section 3.3. Each modeling variant was applied to the feature set resulting from screening with different values of L_{SUP} (see Section 3.1.2), using either the original features or the features extracted from the DAE. We present results for BAR and defer results for TTD to Section 3 of the supplementary materials.

Figure 1 illustrates the strong linear relationship between the mean value under each estimated ITR and the optimal value for the associated cancer type and treatment grouping. Each point in Figure 1 represents a particular estimated ITR for a method variant within a particular cancer. Note that the optimal values do not vary significantly within a cancer type. In the few cases where variability exists is due to the fact that we select c_1 and c_2 separately for each method variant, resulting in different optimal values. A similar relationship between the mean and observed values for each estimated ITR is observed. This suggests that the marginal mean outcome differs across cancer types. This is not unexpected, as some cancers are known to be more sensitive to available treatment options (e.g., CM), and others less so (e.g., PDAC): This observation motivates our use of $P_{\text{opt}}(\hat{D}^*)$ and $P_{\text{obs}}(\hat{D}^*)$ to evaluate performance of estimated ITRs. We present results for $P_{\text{opt}}(\hat{D}^*)$ here and defer results for $P_{\text{obs}}(\hat{D}^*)$ to Section 4 of the supplementary materials.

4.1. Relative Performance of Methods Pooling Across Conditions

We summarize $P_{\text{opt}}(\hat{D}^*)$ for each variant of the proposed methods and the “off-the-shelf” methods in Figure 2, pooling results across different cancer types and values of L_{SUP} . We find that the application of data smoothing prior to ITR estimation boosts the overall performance for many methods, such as Q-learning with embedded linear models (light and dark green), and OWL methods (light and dark gray). This pre-smoothing is less beneficial for Q-learning methods using embedded random forests (salmon and red), as the pre-smoothing itself is performed using random forests. Q-learning methods using pseudo-values (red, QL_2) performed similarly to their QL_1 counterparts (pink) across conditions. In addition, using the lower dimensional features extracted from the DAE did not show significant benefit for most approaches with the exception of the OWL methods using the linear kernel, which we will show later to be sensitive to the dimension of the feature space. Q-learning methods with nonlinear embedded models (salmon, red) showed much better robustness to various modeling choices than linear ones, and showed to be especially helpful in OWL (dark gray). Finally, the simpler off-the-shelf methods showed lower performance and much higher variability across conditions compared to methods utilizing the treatment tree approach. Relative to the LASSO, almost all methods utilizing the treatment tree performed better than the simpler off the shelf methods.

We also find that using a weighted combination of ITRs using the superlearner approach (dark blue) resulted in the best overall performance across conditions. Increasing the number of ITRs included in the superlearner had the effect of boosting performance while also reducing variability in performance across conditions. Here, the SL4 combined all four Q-learning methods from Figure 2 utilizing presmoothing, SL6 includes the addition of methods $QL_{1,\text{RF}}$ and $QL_{2,\text{RF}}$, SL8 adds QL_1 and QL_2 , and SL16 add all Q-learning methods using DAE features. OWL methods were excluded from the superlearner due to computational time, and we expect additional performance increases through their inclusion. This suggests that the superlearner approach can boost performance and also mitigate user uncertainty regarding the selection of various modeling approaches in cases where the optimal approach may not be clear beforehand, at the expense of additional computation. Estimated value function standard deviations are given in Supplementary Figure 6.

The examination of pooled overall and relative performance in Figure 2 is informative when the analyst is unsure what approach is best for their data and wants to use the “safest” approach. For example, the results in Figure 2 can be used to determine a sequence of modeling decisions that exhibited good overall performance across a variety of conditions and achieved low variability across conditions. For example, $QL_{1,\text{smoothed}}$, or Q-learning using smoothed outcomes and observed-values, achieved good performance on average and low variability in performance across conditions. Alternatively, one may elect to combine ITRs from multiple methods using the superlearner approach to avoid choosing an individual method, at the cost of additional computational burden. In summary, our results suggest the following general conclusions: linear methods may benefit from the use of data smoothing prior to ITR estimation, further dimension reduction of the predictor space via DAEs is beneficial for OWL methods using the linear kernel, and nonlinear embedded models are more robust to various modeling choices than linear embedded models.

4.2. Impact of L_{SUP} on Performance

We now examine the relative performance of our estimated ITRs by L_{SUP} , the number of genes used in estimation. Figure 3 contains boxplots of $P_{\text{opt}}(\hat{D}^*)$ pooled over cancer types and stratified by estimation method and L_{SUP} (Supplementary Table 1). Prior to pooling we adjust $P_{\text{opt}}(\hat{D}^*)$ in each cancer type and method by the performance at $P_{\text{opt}}(\hat{D}^*) = 50$ to more clearly delineate changes with respect to dimension. Most methods did not show strong trends in performance across L_{SUP} , with the exception of the $\text{OWL}_{\text{linear}}$ class of methods (light gray). In this class, $P_{\text{opt}}(\hat{D}^*)$ generally decreases with increasing numbers of genes. However, the same trend did not appear when utilizing the Gaussian kernel. Slight downward trends were also observed for Q-learning with an embedded linear model (light and dark green). These observations together suggest that methods in which a linear decision rule is estimated at each step of the treatment tree may be sensitive to the dimension of the feature space.

The optimal set of features for ITR estimation (indexed by L_{SUP}) is the set of all those features for which at least one of the $h_j, j = 0, \dots, J$ in model (1) is not constant. While the optimal set of features is unknown, the results in Figure 3 indicate that the proposed treatment tree-based ITR estimation method performs well regardless of the number of features selected, with the exception of linear OWL and Q-learning with an embedded linear model. These results lead to the conclusion that a smaller set of features post-supervised screening may be optimal for methods with an embedded linear model, and an embedded nonlinear model can be used to provide robustness against selecting the “wrong” set of features.

4.3. Overall Performance by Cancer

The performance resulting from specific modeling decisions varies across cancer types. Figure 4 contains boxplots of $P_{\text{opt}}(\hat{D}^*)$ for the best performing set of modeling decisions within the classes defined by the colors in Figure 2. To select the best performing variant in each class, we chose the one with the largest $P_{\text{opt}}(\hat{D}^*)$ averaged across values of L_{SUP} within that particular class. The boxplots in Figure 4 contain $P_{\text{opt}}(\hat{D}^*)$ over values of L_{SUP} . We show the full set of results corresponding to all methods and cancers in Section 5 of the supplementary materials. Within certain cancers, such as breast cancer (BRCA), specific modeling decisions do not result in large differences in performance. This may be due in part to the fact that a small number of treatments appear to work well uniformly across samples in BRCA (Supplementary Figure 7). In contrast, for pancreatic cancer (PDAC) and non-small cell lung cancer (NSCLC), greater heterogeneity in response exists across treatments (Supplementary Figures 8 and 9). In addition, we find that in almost all cancer types, the superlearner tended to perform better than or similar to all other classes of methods, suggesting its use when it is unclear which individual method may be optimal for a particular dataset.

Table 1 lists the best overall set of modeling decisions for each cancer type, along with the genomic features that were most important for selecting treatments using the best

performing estimated ITR. For cancers where the best performing ITR resulted from the DAE predictors or OWL methods using the Gaussian kernel, it is difficult to determine which genes were the most important. For these cases, we selected important features using the second-best ITR (reference method in Table 1). For all cancer types, the best performing ITR resulted from the tree-based approach rather than an “off-the-shelf” method. In addition, despite their sensitivity to dimension, the OWL_{linear} class of methods were represented as the best ITR in two out of five cancers. This suggests that as long as the correct feature dimension is selected beforehand, OWL methods can perform well relative to other methods. In practice, however, the optimal dimension is difficult to ascertain unless one evaluates multiple candidate feature sets, as we have done in this article.

The treatments most frequently recommended by the best performing ITR for each cancer, along with the corresponding values of c_1 and c_2 , are given in Supplementary Table 6. The treatment tree for the optimal ITR varied in the values of c_2 across cancer types, suggesting variability in the amount of response heterogeneity across cancer types. For example, in CM, where response tended to be strong overall (Figure 1), $c_2 = 8$, suggesting that relatively more treatments had similar response profiles across PDX lines. The selected value of c_1 for the best performing ITR was low for each cancer type, indicating that only a small number of treatments were found to be effectively the same as “untreated” based on the hierarchical clustering (Supplementary Table 6). For “off-the-shelf” methods, $c_1 = 0$ and $c_1 = P_k$ by design since no grouping of treatments was performed.

We list the average (unconditional) response for each of the P_k treatments within cancer type in Supplementary Table 7, calculated as the sample average response across mice treated with each of the P_k treatments across PDX lines. In BRCA, the treatment with the larger mean response was also the most recommended treatment (LEE011 + everolimus). In PDAC, however, there is less variability in average response across treatments, and the best performing ITR recommends BKM120 + binimetinib to 18 PDX lines and abraxane + gemcitabine to 18 PDX lines (see Supplementary Table 8).

4.4. ITR Performance When Limiting Features to a Single Genomic Platform

The three genomic platforms used in this study provide a wealth of information for ITR estimation and biomarker discovery. However, the high dimension of the feature space provides practical and computational challenges. Predictors from the same gene may be correlated (e.g., gene expression and gene copy number) and may therefore be redundant. Evaluating three genomic platforms for each PDX line’s original tumor increases cost and amount of tumor tissue required. For these reasons, we also explored the performance of ITRs estimated using only the RNA-seq gene expression platform, a common genomic assay performed by biomedical researchers. We repeated the same process described in Section 3 for a subset of methods, but using only features resulting from RNA-seq.

The overall conclusions are largely similar to those in Figure 2 (see Supplementary Figure 10). When we compare the difference in $P_{\text{opt}}(\hat{D}^*)$ between the ITR estimated using the full feature set and the ITR estimated using RNA-seq only, we find that the performance loss was relatively small when using only RNA-seq data (Supplementary Figure 11). In the case

of OWL methods using the linear kernel, an increase in performance was observed over the prior results, which likely reflects the sensitivity of these methods to the dimension of the feature space. Overall, these results suggest that utilizing a single genomic platform may be a more cost-effective option that will result in estimated ITRs with comparable performance.

4.5. Top Genomic Features in NSCLC

Lung cancer is one of the leading causes of cancer related deaths in the United States, and NSCLC accounts for the majority of clinical cases of lung cancer (Ettinger et al. 2010). Therapeutic agents used to treat NSCLC include paclitaxel, which interferes with cellular microtubular dynamics and cellular division through the targeting of tubulin (Wise, Krahe, and Oakley 2000), and cetuximab, which targets the epidermal growth factor receptor (EGFR) (Pirker et al. 2009). Since the mechanism of action differs between these two treatments, it is not surprising that we observed significant heterogeneity in response between these two treatments in this study (Supplementary Figure 9). Here, we examine the relationship between response and the genomic features found to be the most important for making decisions using the best performing estimated ITR.

The best performing ITR for NSCLC resulted from $OWL_{\text{kernel,smoothed}}$ (see Table 1). Given that that Gaussian kernel for OWL does not allow direct interpretation of its predictors, we utilize the reference method $QL_{2,\text{smoothed}}$ in this cancer to examine the role of each selected predictor with respect to response. We determined the most important genomic features for this ITR using the following approach. For each of the c_2 splits in the associated treatment tree, we computed the absolute value of the regression coefficient for each feature in the model fit at a given node and retained the feature with the largest magnitude value at each node. Cross-validation selected $c_2 = 13$; therefore, we retained a set of 13 top features, where 10 of these were unique (Figure 5). We then calculated Spearman's rank correlation coefficient between each selected feature and response, separately within each nonnull treatment. This resulted in the matrix of feature-treatment pairwise correlation coefficients seen in Figure 5. We then clustered the columns of this matrix, pertaining to the seven selected genomic features, using Euclidean distance between the vector of correlation coefficients in each column. We also clustered the rows of this matrix, pertaining to the nonnull treatments, using the tree structure that was previously constructed for $QL_{2,\text{smoothed}}$, rather than the correlations.

The treatment with one of the strongest correlation to Tubulin Gamma 1 (TUBG1) copy number is the therapy paclitaxel, suggesting that a higher copy number of TUBG1 may potentiate response in patients being targeted with agents such as paclitaxel. This is notable because paclitaxel directly targets tubulin (Kumar 1981), of which TUBG1 plays a major role. Furthermore, cetuximab is the only treatment that is anticorrelated with TUBG1 copy number. This unique relationship with TUBG1 is reflected in the treatment tree, as cetuximab is the only member of a branch furthest away from all other nonnull treatments. Cetuximab also exhibits a unique mechanism of action, as the only monoclonal antibody EGFR inhibitor in our dataset. The relationship between TUBG1 copy number and response to cetuximab and paclitaxel is displayed in Figure 5. These results indicate that that the observed correlations between treatment response and the top features determined from

$QL_{2,smoothed}$ reflect the role these features play as important variables in the best performing ITR.

5. Discussion

In this article, we introduced several approaches to ITR estimation using PDX data. The unique structure of a PDX study, where multiple treatments are applied to samples from the same human tumor implanted into mice, naturally lends itself to precision medicine. The substantially improved precision that results from the PDX structure may result in better performing ITRs. However, PDX data also pose a number of challenges, including a large number of unordered treatments and a high-dimensional feature space. These factors make it difficult to nonparametrically model the conditional mean of the response. The method we propose involves a sequence of steps that alleviate these difficulties. Our method involves screening the features to find a lower dimensional feature space, constructing a treatment tree that allows for recasting ITR estimation as a sequence of binary decisions, and estimating a decision rule at each split to select the arm that contains the optimal treatment. Because we aim to select the arm that contains the optimal treatment at each split rather than the arm with the largest average response, our estimation technique utilizes the maximum response downstream of each arm for each line. Thus, our estimation technique incorporates the unique structure of the PDX data by directly using the multiple responses observed per PDX line. We've shown that the method we propose not only produces high-quality ITRs, but also identifies genes that are known to be associated with response to treatment (e.g., the TUBG1 gene shown in Figure 5).

The methods we propose requires making a number of modeling decisions at various stages of the pipeline, including selecting the dimension of the feature space, choosing embedded models, and selecting tuning parameters, among others. We demonstrated various combinations of these modeling decisions in our analyses. While certain variants performed better than others for certain cancer types, the treatment tree-based approach outperformed “off-the-shelf” methods overall. Reducing noise by using random forest-predicted outcomes (smoothing) improved performance of the estimated ITRs in most settings, and basing ITRs on DAE-extracted features improved performance in the presence of a linear embedded model. We recognize that the models, tuning parameters, and implementation discussed here, despite our best efforts, may not be optimal. The method studied here could potentially be improved through careful tuning. Studying the proposed method further, including through extensive simulation experiments, could yield more concrete recommendations as to which embedded models perform best. Our implementation of a superlearner consisting of multiple estimated ITRs was shown to improve performance above individual ITRs. This approach helps mitigate user uncertainty regarding the best choice of ITR estimation approach or modeling choices for a given problem, where one may combine the results from multiple ITRs using the superlearner for improved performance.

The assumption that ITRs estimated from PDX data are applicable to humans is crucial to this work. This assumption is based on decades of biomedical research on generalizing PDX results to humans. A major conclusion of Gao et al. (2015) was that the responses observed in their PDX lines correlated with the responses observed in the human patients from which

the tumors were taken. Many prior studies have shown that PDX models show stronger correlation with human response compared to traditional cell line models (Rosfjord et al. 2014; Scott, Mackay, and Haluska 2014; Whittle et al. 2015). Prior work has also shown that the tumor microenvironment, consisting of stromal cells and other tissue, may impact tumor activity and response to treatment. In PDX models, the microenvironment surrounding the implanted human tumor is nonhuman. However, several recent studies have suggested that tumor recruitment of mouse stroma in PDX models mirrors that seen in humans (Roife et al. 2016; Wang et al. 2017) and may show similarity in treatment response when specifically targeted.

A key next step for this research will be to plan and conduct validation studies in humans to determine if the biomarkers and ITRs discovered here can be used to improve outcomes in human patients. We note that ITRs based on only one genomic platform (RNA-seq) resulted in comparable performance to ITRs based on three genomic platforms. Since independent validation studies will be more expensive if more genomic platforms are needed, validating the ITRs based only on RNA-seq data using a study of human cancer patients would be a low-cost initial step toward validating the results in this article. One advantage of the proposed method is that, in some settings, we could apply the estimated ITRs to external studies that included only a subset of the treatments applied in this study. Given a sequence of decision rules (pertaining to each step of the treatment tree), one could start at the lowest node that contains all of the treatments of interest and follow the tree to arrive at a recommended treatment, rather than starting from the top of the tree. We also note that, while our results allow for comparing the performance of estimated ITRs across different cancer types and modeling strategies, the data used for this article do not allow for comparing the performance of ITRs estimated from PDX studies to ITRs estimated from human trials. Another important step forward for validating these results will be to compare the treatment rules discovered here to those estimated from human trials to determine the value of PDX studies for precision medicine.

The design of a PDX study plays a key role in ITR estimation, and designing high-quality PDX studies is another key next step for this research. Our results suggest that the observed responses in PDX studies are noisy. Having replicates within PDX line, that is, multiple mice per line assigned to the same treatment, may improve the performance of the estimated ITRs. The design could also be improved by having a larger number of distinct tumor lines to ensure sufficient representation of cancer heterogeneity across a diverse spectrum of cancer patients.

While our research has, in some ways, raised more questions than it has answered, we feel that the treatment rules, biomarkers, and other results we have discovered here are interesting in their own right and merit further research, including validation studies. Future research in this area will allow us to make measurable advances in applying PDX studies in precision medicine and translating the results into clinical practice. Plans for such future research are underway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Athey S, and Imbens G (2016), "Recursive Partitioning for Heterogeneous Causal Effects," *Proceedings of the National Academy of Sciences of the United States of America*, 113, 7353–7360. [2] [PubMed: 27382149]
- Bourgon R, Gentleman R, and Huber W (2010), "Independent Filtering Increases Detection Power for High-Throughput Experiments," *Proceedings of the National Academy of Sciences of the United States of America*, 107, 9546–9551. [3] [PubMed: 20460310]
- Chen J, Fu H, He X, Kosorok MR, and Liu Y (2017), "Estimating Individualized Treatment Rules for Ordinal Treatments," *arXiv no. 1702.04755*. [2,5,6]
- Chen Z, Fillmore CM, Hammerman PS, Kim CF, and Wong K-K (2014), "Non-Small-Cell Lung Cancers: A Heterogeneous Set of Diseases," *Nature Reviews Cancer*, 14, 535–546. [2] [PubMed: 25056707]
- Ettinger DS, Akerley W, Bepler G, Blum MG, Chang A, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Ganti AKP, and Govindan R (2010), "Non-Small Cell Lung Cancer," *Journal of the National Comprehensive Cancer Network*, 8, 740–801. [12] [PubMed: 20679538]
- Fan J, and Lv J (2008), "Sure Independence Screening for Ultrahigh Dimensional Feature Space," *Journal of the Royal Statistical Society, Series B*, 70, 849–911. [4]
- Gao H, Korn JM, Ferretti S, Monahan JE, Wang Y, Singh M, Zhang C, Schnell C, Yang G, Zhang Y, and Balbin OA (2015), "High-Throughput Screening Using Patient-Derived Tumor Xenografts to Predict Clinical Trial Drug Response," *Nature Medicine*, 21, 1318–1325. [1,2,3,14]
- Hidalgo M, Amant F, Biankin AV, Budinská E, Byrne AT, Caldas C, Clarke RB, de Jong S, Jonkers J, Mælandsmo GM, and Roman-Roman S (2014), "Patient-Derived Xenograft Models: An Emerging Platform for Translational Cancer Research," *Cancer Discovery*, 4, 998–1013. [1] [PubMed: 25185190]
- Imai K, and Ratkovic M (2013), "Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation," *The Annals of Applied Statistics*, 7, 443–470. [2]
- Kumar N (1981), "Taxol-Induced Polymerization of Purified Tubulin. Mechanism of Action," *Journal of Biological Chemistry*, 256, 10435–10441. [12]
- Laber E, and Zhao Y (2015), "Tree-Based Methods for Individualized Treatment Regimes," *Biometrika*, 102, 501–514. [4] [PubMed: 26893526]
- Liu Y, Wang Y, Kosorok MR, Zhao Y, and Zeng D (2016), "Robust Hybrid Learning for Estimating Personalized Dynamic Treatment Regimens," *arXiv no. 1611.02314*. [2,5]
- Love MI, Huber W, and Anders S (2014), "Moderated Estimation of Fold Change and Dispersion for RNA-seq Data With DESeq2," *Genome Biology*, 15, 550. [3] [PubMed: 25516281]
- Luedtke AR, and van der Laan MJ (2016), "Super-Learning of an Optimal Dynamic Treatment Rule," *The International Journal of Biostatistics*, 12, 305–332. [2,3,6] [PubMed: 27227726]
- Ma J, Stingo FC, and Hobbs BP (2016), "Bayesian Predictive Modeling for Genomic Based Personalized Treatment Selection," *Biometrics*, 72, 575–583. [2] [PubMed: 26575856]
- Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, Loi S, Bradbury I, Bliss JM, Azim HA Jr., Ellis P, and Di Leo A (2012), "Dissecting the Heterogeneity of Triple-Negative Breast Cancer," *Journal of Clinical Oncology*, 30, 1879–1887. [2] [PubMed: 22454417]
- Murphy SA (2005), "A Generalization Error for Q-Learning," *Journal of Machine Learning Research*, 6, 1073–1097. [2,5] [PubMed: 16763665]
- Murphy SA, van der Laan MJ, Robins JM, and Conduct Problems Prevention Research Group (2001), "Marginal Mean Models for Dynamic Regimes," *Journal of the American Statistical Association*, 96, 1410–1423. [2] [PubMed: 20019887]

- Orellana L, Rotnitzky A, and Robins JM (2010), “Dynamic Regime Marginal Structural Mean Models for Estimation of Optimal Dynamic Treatment Regimes, Part I: Main Content,” *The International Journal of Biostatistics*, 6, 8. [2]
- Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, and Quackenbush JF (2009), “Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes,” *Journal of Clinical Oncology*, 27, 1160–1167. [2] [PubMed: 19204204]
- Pirker R, Pereira JR, Szczesna A, Von Pawel J, Krzakowski M, Ramlau R, Vynnychenko I, Park K, Yu C-T, Ganul V, and Roh JK (2009), “Cetuximab Plus Chemotherapy in Patients With Advanced Non-Small-Cell Lung Cancer (FLEX): An Open-Label Randomised Phase III Trial,” *The Lancet*, 373, 1525–1531. [12]
- Polyak K (2011), “Heterogeneity in Breast Cancer,” *The Journal of Clinical Investigation*, 121, 3786–3788. [1] [PubMed: 21965334]
- Qian M, and Murphy SA (2011), “Performance Guarantees for Individualized Treatment Rules,” *The Annals of Statistics*, 39, 1180. [2] [PubMed: 21666835]
- Rashid NU, Sperling AS, Bolli N, Wedge DC, Van Loo P, Tai Y-T, Shammass MA, Fulciniti M, Samur MK, Richardson PG, and Magrangeas F (2014), “Differential and Limited Expression of Mutant Alleles in Multiple Myeloma,” *Blood*, 124, 3110–3117. [3] [PubMed: 25237203]
- Robins J, Orellana L, and Rotnitzky A (2008), “Estimation and Extrapolation of Optimal Treatment and Testing Strategies,” *Statistics in Medicine*, 27, 4678–4721. [2] [PubMed: 18646286]
- Roife D, Dai B, Kang Y, Perez MVR, Pratt M, Li X, and Fleming JB (2016), “Ex Vivo Testing of Patient-Derived Xenografts Mirrors the Clinical Outcome of Patients With Pancreatic Ductal Adenocarcinoma,” *Clinical Cancer Research*, 22, 6021–6030. [14] [PubMed: 27259561]
- Rosfjord E, Lucas J, Li G, and Gerber H-P (2014), “Advances in Patient-Derived Tumor Xenografts: From Target Identification to Predicting Clinical Response Rates in Oncology,” *Biochemical Pharmacology*, 91, 135–143. [14] [PubMed: 24950467]
- Rubin D (1978), “Bayesian Inference for Causal Effects: The Role of Randomization,” *The Annals of Statistics*, 6, 34–58. [5]
- Sargent DJ, Conley BA, Allegra C, and Collette L (2005), “Clinical Trial Designs for Predictive Marker Validation in Cancer Treatment Trials,” *Journal of Clinical Oncology*, 23, 2020–2027. [1] [PubMed: 15774793]
- Schulte PJ, Tsiatis AA, Laber EB, and Davidian M (2014), “Q- and A-Learning Methods for Estimating Optimal Dynamic Treatment Regimes,” *Statistical Science*, 29, 640–661. [2,5] [PubMed: 25620840]
- Scott CL, Mackay HJ, and Haluska P Jr. (2014), “Patient-Derived Xenograft Models in Gynecological Malignancies,” in *American Society of Clinical Oncology Educational Book/ASCO. American Society of Clinical Oncology. Meeting, NIH Public Access*, p. e258. [14]
- Siolas D, and Hannon GJ (2013), “Patient-Derived Tumor Xenografts: Transforming Clinical Samples Into Mouse Models,” *Cancer Research*, 73, 5315–5319. [1] [PubMed: 23733750]
- Székel GJ, and Rizzo ML (2009), “Brownian Distance Covariance,” *The Annals of Applied Statistics*, 3, 1236–1265. [4]
- Tentler JJ, Tan AC, Weekes CD, Jimeno A, Leong S, Pitts TM, Arcaroli JJ, Messersmith WA, and Eckhardt SG (2012), “Patient-Derived Tumour Xenografts as Models for Oncology Drug Development,” *Nature Reviews Clinical Oncology*, 9, 338–350. [1]
- Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, and Gwyther SG (2000), “New Guidelines to Evaluate the Response to Treatment in Solid Tumors,” *Journal of the National Cancer Institute*, 92, 205–216. [3] [PubMed: 10655437]
- Tibshirani R, Saunders M, Rosset S, Zhu J, and Knight K (2005), “Sparsity and Smoothness via the Fused Lasso,” *Journal of the Royal Statistical Society, Series B*, 67, 91–108. [4]
- Vincent P, Larochelle H, Lajoie I, Bengio Y, and Manzagol P-A (2010), “Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network With a Local Denoising Criterion,” *Journal of Machine Learning Research*, 11, 3371–3408. [4]
- Wager S, and Athey S (2018), “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests,” *Journal of the American Statistical Association*, 113, 1228–1242. [2]

- Wang L, and Laber E (2017), "Sufficient Markov Decision Processes," Submitted. [4]
- Wang X, Mooradian AD, Erdmann-Gilmore P, Zhang Q, Viner R, Davies SR, Huang K.-l., Bomgarden R, Van Tine BA, Shao J, and Ding L (2017), "Breast Tumors Educate the Proteome of Stromal Tissue in an Individualized But Coordinated Manner," *Science Signaling*, 10, eaam8065. [14] [PubMed: 28790197]
- Whittle JR, Lewis MT, Lindeman GJ, and Visvader JE (2015), "Patient-Derived Xenograft Models of Breast Cancer and Their Predictive Power," *Breast Cancer Research*, 17, 17. [14] [PubMed: 25849559]
- Wise DO, Krahe R, and Oakley BR (2000), "The γ -Tubulin Gene Family in Humans," *Genomics*, 67, 164–170. [12] [PubMed: 10903841]
- Wu T (2016), "Set Valued Dynamic Treatment Regimes," PhD thesis, The University of Michigan. [4]
- Zhang B, Tsiatis AA, Davidian M, Zhang M, and Laber E (2012), "Estimating Optimal Treatment Regimes From a Classification Perspective," *Stat*, 1, 103–114. [2] [PubMed: 23645940]
- Zhang B, Tsiatis AA, Laber EB, and Davidian M (2013), "Robust Estimation of Optimal Dynamic Treatment Regimes for Sequential Treatment Decisions," *Biometrika*, 100, 681–694. [2]
- Zhang Y, Laber EB, Tsiatis A, and Davidian M (2015), "Using Decision Lists to Construct Interpretable and Parsimonious Treatment Regimes," *Biometrics*, 71, 895–904. [4] [PubMed: 26193819]
- Zhao Y, Kosorok MR, and Zeng D (2009), "Reinforcement Learning Design for Cancer Clinical Trials," *Statistics in Medicine*, 28, 3294–3315. [2,5,6] [PubMed: 19750510]
- Zhao Y, Zeng D, Rush AJ, and Kosorok MR (2012), "Estimating Individualized Treatment Rules Using Outcome Weighted Learning," *Journal of the American Statistical Association*, 107, 1106–1118. [2,6] [PubMed: 23630406]
- Zhao Y, Zeng D, Socinski MA, and Kosorok MR (2011), "Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer," *Biometrics*, 67, 1422–1433. [5,6] [PubMed: 21385164]
- Zhou X, Mayer-Hamblett N, Khan U, and Kosorok MR (2017), "Residual Weighted Learning for Estimating Individualized Treatment Rules," *Journal of the American Statistical Association*, 112, 169–187. [2,5] [PubMed: 28943682]

Figure 1.

Original (left) and scaled (right) estimated ITR values from all analyses, spanning method variants, cancer types, and L_{SUP} . Optimal values vary significantly by cancer type. The mean value of each estimated ITR is correlated with the optimal value across cancer types (left panel). We normalize the estimated value for each estimated ITR by the optimal value to allow for comparisons between cancers. This metric, called “proportion of optimal,” provides a measure of how close the value of an estimated ITR is to the optimal value.

Figure 2. Overall performance across methods, pooled over cancer types and number of features (top). $P_{\text{opt}}(\hat{D}^*)$ for each method is also normalized to the LASSO in each condition to highlight the relative performance of each approach (bottom). This relative measure was constructed by subtracting the $P_{\text{opt}}(\hat{D}^*)$ pertaining to the LASSO from that of the other methods within each combination of cancer type and L_{SUP} value.

Figure 3. Overall trends in performance for each method across L_{SUP} , aggregated over cancer types.

Figure 4.
Performance of the best estimated ITR in each class across cancer types.

Figure 5.

Top genomic features selected by $QL_{2,smoothed}$ in NSCLC. Cells are colored by the magnitude of their Spearman's correlation between response to treatment (rows) and expression of top features (columns). Genomic features were clustered by their vector of Spearman's correlation to each treatment (using Euclidean distance), whereas treatments were grouped based on the treatment tree constructed for $QL_{2,smoothed}$. The patterns of treatment-feature correlations tended to coincide with the predetermined treatment groupings.

Table 1.

Best performing method and number of predictors for each cancer.

Cancer	Method	Lsup	Reference method	Top 5 predictors
BRCA	owllinearsmoothdl	50	qlr ²	COL1A1.rna,CTCF.L.rna,FMNL3.rna,SRPX.rna,HMCN1.cn
CM	owllineardl	100	ql1smooth	ETV7.rna,DPYSL3.ma,LEPREL1.ma,GABRE.cn,ATP2B1.ma
CRC	ql1smooth	100	ql1smooth	FGG.rna,ALPK1.ma,WDR27.cn,DIDO1.mut,C10orf26.ma
NSCLC	owlkernelsmooth	100	ql2smooth	POFUT2.rna,TNNI3.cn,TUBG1.cn,NAT8L.cn,PTPRE.cn
PDAC	ql1smooth	100	ql1smooth	CTH.rna,DUSP4.cn,TPP2.rna,ACVR1B.ma,ZNF264.cn

NOTE: Top 5 most important predictors are listed. Predictors pertaining to the next best performing method were provided if the top performing method utilized deep learning (reference method). Predictors ending in .rna are from the gene expression data, .cn from the copy number data, and .mut from the mutation dataset.