Published in final edited form as:

Stat (Int Stat Inst). 2016 ; 5(1): 185–199. doi:10.1002/sta4.114.

Flexible functional regression methods for estimating individualized treatment regimes

Adam Ciarleglio^{a,*}, Eva Petkova^{a,b}, Thaddeus Tarpey^c, and R. Todd Ogden^d

^aDepartment of Child and Adolescent Psychiatry, NYU School of Medicine, New York, NY 10016, USA

^bNathan S. Kline Institute for Psychiatric Research, Orangesburg, NY 10962, USA

^cDepartment of Mathematics and Statistics, Wright State University, Dayton, OH 45435, USA

^dDepartment of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY 10032, USA

Abstract

A major focus of personalized medicine is on the development of individualized treatment rules. Good decision rules have the potential to significantly advance patient care and reduce the burden of a host of diseases. Statistical methods for developing such rules are progressing rapidly, but few methods have considered the use of pre-treatment functional data to guide in decision-making. Furthermore, those methods that do allow for the incorporation of functional pre-treatment covariates typically make strong assumptions about the relationships between the functional covariates and the response of interest. We propose two approaches for using functional data to select an optimal treatment that address some of the shortcomings of previously developed methods. Specifically, we combine the flexibility of functional additive regression models with Q -learning or A -learning in order to obtain treatment decision rules. Properties of the corresponding estimators are discussed. Our approaches are evaluated in several realistic settings using synthetic data and are applied to real data arising from a clinical trial comparing two treatments for major depressive disorder in which baseline imaging data are available for subjects who are subsequently treated.

Keywords

Functional data; Additive models; Q -learning; A -learning; Treatment regime; imaging data

1. Introduction and Motivation

An individualized treatment rule (ITR) is a rule that assigns treatment based on a subject's characteristics. Ideally an ITR is constructed so that each subject receives a treatment that is

* Adam.Ciarleglio@nyumc.org.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site. This includes a pdf with Appendices A, B, and C referenced in the article as well as the R code that was used to perform the simulations in Section 4.

optimal with respect to some response of interest. Much work has been done to develop methods for estimating ITRs at one time point using data from either randomized controlled trials (RCTs) or observational studies. Qian and Murphy (2011), Lu *et al.* (2013), and Tian *et al.* (2014) have proposed regression-based methods in which the response of interest is modeled as a function of treatment assignment and pre-treatment covariates. Zhang *et al.* (2012b), Zhang *et al.* (2012a), Zhao *et al.* (2012), and Zhou *et al.* (2015) have proposed policy-search methods that seek out the best ITR within a large class of possible rules. Recently, Laber and Zhao (2015) proposed a method for estimating easily interpretable ITRs based on decision trees.

Each of the methods mentioned above assume the ITR depends upon subject characteristics that can be captured completely by scalar quantities. Ideally we would like to have simple and interpretable ITRs based on scalar measures that can be easily collected from subjects before starting treatment. However, such simplistic rules may not be realistic. Furthermore, with advancements in technology, we have the ability to collect complex and high dimensional data on each individual, thus expanding the pool of potential patient characteristics to use for developing ITRs. For example, in the search for quality rules for assigning treatment for major depressive disorder (MDD), investigators are turning to measures derived from various brain imaging modalities such as magnetic resonance imaging (MRI) or electroencephalography (EEG) in search for characteristics that can account for differential treatment response. These high dimensional patient characteristics can be viewed as functional data (Ramsay and Silverman, 2005). Methods that incorporate these increasingly prevalent functional data into ITRs may help to significantly advance patient care.

There is scarce research on methods for developing ITRs based on pre-treatment functional covariates. McKeague and Qian (2014) propose methods for estimating and evaluating treatment regimes that depend upon one baseline functional covariate. Ciarleglio *et al.* (2015) propose a method that extends advantage learning (*A*-learning) (Robins, 2004) to allow the ITR to depend on multiple scalar and functional covariates while Ciarleglio *et al.* (2016) developed a method that performs both estimation and variable selection for ITRs when the pre-treatment covariates consist of both scalars and functions. Although different in many respects, each of these regression-based approaches estimate ITRs from among classes of functions that assume simple relationships between the pre-treatment covariates and treatment effect. For example, the *A*-learning approach proposed by Ciarleglio *et al.* (2015) assumes a partial functional linear model for the treatment effect. However if this model is mis-specified then the corresponding ITR may be of poor quality.

This has prompted us to consider alternative flexible functional regression-based methods for estimating treatment decision rules. Here we combine two common approaches for estimating ITRs, *Q*-learning (Watkins and Dylan, 1992) and *A*-learning (for estimating a decision rule at one time point), with functional generalized additive models (fGAMs) (McLean *et al.*, 2014). One advantage of using fGAMs is that it can help to reduce the risk of model mis-specification by allowing for the consideration of ITRs from a large class of functions. Moodie *et al.* (2014) take a similar approach for developing dynamic treatment regimes with a limited number of scalar tailoring variables in which they model *Q*-functions

via generalized additive models (GAMs). In our approaches we employ penalized B -splines with different smoothing parameters allowed for each pre-treatment covariate. In addition to expanding the class of rules from which ITRs can be constructed, the use of fGAMs is attractive since tuning parameter selection and fitting can be done with existing software.

In this paper, we describe our proposed Q and A -learning fGAM approaches and demonstrate their use. The remainder of the paper is organized as follows. In Section 2 we establish a formal framework from which to discuss estimation of ITRs. In Section 3 we present the rationale for the Q and A -learning fGAM approaches as well as how to estimate the corresponding ITRs. Section 4 provides results from numerical investigations that demonstrate the benefits of our proposed methods and Section 5 shows the results when our methods are applied to data arising from a RCT of treatment for MDD. We conclude with a discussion in Section 6 including limitations and extensions of our proposed methods.

2. Framework and Optimal ITRs

Consider data arising from a RCT or observational study in which n subjects receive one of two possible treatment options denoted by either 0 or 1. Let $A \in \mathcal{A} = \{0, 1\}$ be the indicator for the observed treatment received. For each subject, we observe a pre-treatment functional covariate, X , which is a one-dimensional square-integrable functional random variable ($X: D \subset \mathbb{R} \rightarrow \mathbb{R}$). Without loss of generality, we take D to be the interval $[0, 1]$. Although we present only one-dimensional functional predictors here, it is possible to extend our approach to higher dimensional functional random variables such as images.

We also observe a post-treatment response, $Y \in \mathbb{R}$. Without loss of generality, we assume that larger values of Y are desirable. Hence the observed data are given by (Y_i, A_i, X_i) for $i = 1, \dots, n$. Our goal is to use these data to estimate a rule for assigning treatment based on X in order to obtain the best possible response, Y , under the competing treatments in $\mathcal{A}$. Let $\pi(X) = P(A = 1|X)$ denote the propensity score. In the RCT setting $A \perp X$ and we assume that $\pi(X) = 0.5$.

Following the potential outcomes framework of Rubin (1978), let $Y^*(a)$, with $a \in \mathcal{A}$, be the value of the outcome that would be observed if a subject actually received treatment $A = a$. The following assumptions are typically made when computing the expected value of the potential outcome: (A1) Consistency assumption: $Y = Y^*(1)A + Y^*(0)(1 - A)$; (A2) No unmeasured confounders assumption: $A \perp Y^*(a)|X$; (A3) Positivity assumption: $0 < \pi(X) < 1$ for all X .

$E\{Y^*(a)\}$ corresponds to the population mean response under treatment a . An ITR is a function, d , that maps the baseline covariates X to $\mathcal{A}$ such that a patient with pre-treatment covariate $X = x$ will receive treatment 1 if $d(x) = 1$ and will receive treatment 0 if $d(x) = 0$.

The optimal ITR (OITR), $d_{\mathcal{D}}^{opt}$, is the function that maximizes $E\{Y^*(a)\}$ among some class of functions $\mathcal{D}$, so that $d_{\mathcal{D}}^{opt}(X) = \arg\max_{d \in \mathcal{D}} E[Y^*\{d(X)\}]$. In the following, we drop the subscript $\mathcal{D}$. Using the assumptions above, it can be shown that $E[Y^*\{d(X)\}] = E_X[E(Y|A = 1, X)d(X) + E(Y|A = 0, X)\{1 - d(X)\}]$, where $E_X(\cdot)$ denotes the expectation with respect to X . Thus it is easy to show that $d^{opt}(X) = I\{E(Y|A = 1, X) - E(Y|A = 0, X) > 0\}$, where $I(\cdot)$

is the indicator function. Hence, the optimal decision rule corresponding to the class $\mathcal{D}$ can be estimated using the observed data.

Regression-based approaches for constructing ITRs require the user to posit a model for either $E(Y|A, X)$ or for the difference $E(Y|A = 1, X) - E(Y|A = 0, X)$. The class of functions $\mathcal{D}$ is directly related to the model selected for either of these quantities. For example, if we use the functional linear model

$E(Y|A, X) = \alpha_0 + \int_0^1 \omega_1(s)X(s)ds + \alpha_1 A + A \int_0^1 \omega_2(s)X(s)ds$ then $\mathcal{D}$ consists of rules of the form $I\{\alpha_1 + \int_0^1 \omega_2(s)X(s)ds > 0\}$ which assumes a linear relationship between the pre-treatment functional covariate and treatment effect. This is the class of functions from which ITRs are estimated in Ciarleglio *et al.* (2015) and Ciarleglio *et al.* (2016). The class $\mathcal{D}$ can be extended to include a broader collection of ITRs by using a more flexible model for either $E(Y|A, X)$ or $E(Y|A = 1, X) - E(Y|A = 0, X)$. We will use fGAMs to model these quantities.

3. Methods

3.1. Q-Learning with fGAM (Q-fGAM)

In the statistical literature on ITRs, $E(Y|A = a, X)$ is often referred to as the Q -function and denoted by $Q(X, A)$. To estimate an ITR using Q -learning, one must posit a model for Q . Here we allow for the Q -function to be a member of the class of fGAMs. The class of fGAMs provides a highly flexible space from which to model the Q -function and therefore allows for a correspondingly rich and flexible class of functions, $\mathcal{D}$, from which d^{opt} can be estimated. Specifically, we model the Q -function as

$$Q(X, A) = \theta_0 + \int_0^1 F\{X(s), s\}ds + \theta_1 A + A \int_0^1 G\{X(s), s\}ds. \quad (1)$$

In (1), F and G are unspecified, smooth bivariate functions of $X(s)$ and s . Note that the functional linear model is a special case of (1) in which $F\{X(s), s\} = \beta_1(s)X(s)$ and $G\{X(s), s\} = \beta_2(s)X(s)$ for smooth square-integrable functions β_1 and β_2 . Under the parameterization in (1), $\int_0^1 G\{X(s), s\}ds$ corresponds to the differential treatment effect depending on X and it is clear that

$$d^{opt}(X) = I\{Q(X, 1) - Q(X, 0) > 0\} = I\left\{\theta_1 + \int_0^1 G\{X(s), s\}ds > 0\right\}. \quad (2)$$

The quantity $Q(X, 1) - Q(X, 0)$ (and therefore $\theta_1 + \int_0^1 G\{X(s), s\}ds$) is often referred to as the “contrast.” We will follow this convention and we let $C(X) = Q(X, 1) - Q(X, 0)$. We can use the data to fit model (1) and use the estimates for θ_1 and G , given by $\hat{\theta}_1$ and $\hat{G}$ respectively, to obtain an estimate for the OITR given by

$\hat{d}^{opt}(X) = I \left\{ \hat{\theta}_1 + \int_0^1 \hat{G}\{X(s), s\} ds > 0 \right\}$, where the contrast is in the class of fGAMs, which can be used to assign new subjects to treatment.

Estimation of models such as those given in (1) as well as information on necessary identifiability constraints are discussed in depth in McLean *et al.* (2014). Here and in the Appendices we briefly highlight some aspects of the fitting procedure so that we may comment on how certain choices affect the estimated decision rule.

We use tensor products of B -splines to represent the surfaces F and G in model (1). Specifically, we have

$$F(x, s) \approx \sum_{j=1}^{B_x} \sum_{k=1}^{B_s} \eta_{j,k} b_j^X(x) b_k^S(s) \quad \text{and} \quad G(x, s) \approx \sum_{j=1}^{C_x} \sum_{k=1}^{C_s} \gamma_{j,k} c_j^X(x) c_k^S(s). \quad (3)$$

where $\{b_j^X(x): j=1, \dots, B_x\}$, $\{b_k^S(s): k=1, \dots, B_s\}$, $\{c_j^X(x): j=1, \dots, C_x\}$, and $\{c_k^S(s): k=1, \dots, C_s\}$ are sets of spline basis functions on $[0, 1]$ with B_x , B_s , C_x and C_s all chosen large enough (often much larger than needed) to allow for adequate representation of the unknown functions F and G . With these representations we can express the Q -function in

$$(1) \text{ for subject } i \text{ as } Q(X_i, A_i) \approx \theta_0 + \sum_{j=1}^{B_x} \sum_{k=1}^{B_s} \eta_{j,k} U_{i,j,k} + \theta_1 A_i + \sum_{j=1}^{C_x} \sum_{k=1}^{C_s} \gamma_{j,k} A_i V_{i,j,k},$$

where $U_{i,j,k} = \int_0^1 b_j^X\{X_i(s)\} b_k^S(s) ds$ and $V_{i,j,k} = \int_0^1 c_j^X\{X_i(s)\} c_k^S(s) ds$ are computed via numerical integration (e.g., using Simpson's rule). Note that estimating the Q -function corresponds to estimating $2 + B_x B_s + C_x C_s$ parameters which can be arranged into the vector $\xi = (\theta_0, \theta_1, \eta_{1,1}, \dots, \eta_{B_x, B_s}, \gamma_{1,1}, \dots, \gamma_{C_x, C_s})^\top$. Ideally, the parameters in ξ would be estimated by minimizing the penalized loss function given by

$$\ell_Q(\xi; \lambda_r^\eta, \lambda_c^\eta, \lambda_r^\gamma, \lambda_c^\gamma) = \sum_{i=1}^n (Y_i - \mathbf{W}_i^\top \xi)^2 + \lambda_c^\eta P_c(\boldsymbol{\eta}) + \lambda_r^\eta P_r(\boldsymbol{\eta}) + \lambda_c^\gamma P_c(\boldsymbol{\gamma}) + \lambda_r^\gamma P_r(\boldsymbol{\gamma}), \quad (4)$$

where $\mathbf{W}_i = (1, A_i, U_{i,1,1}, \dots, U_{i,B_x, B_s}, A_i V_{i,1,1}, \dots, A_i V_{i,C_x, C_s})^\top$, $\boldsymbol{\eta} = (\eta_{1,1}, \dots, \eta_{B_x, B_s})^\top$, $\boldsymbol{\gamma} = (\gamma_{1,1}, \dots, \gamma_{C_x, C_s})^\top$, and $P_q^p(\boldsymbol{\rho})$ ($\boldsymbol{\rho} = \boldsymbol{\eta}$ or $\boldsymbol{\gamma}$; $q=r$ or c) is a penalty term with corresponding non-negative penalty parameter λ_q^p . These penalty terms are described in Appendix A. Briefly, the structure of the penalty terms is motivated by the fact that the spline coefficients comprising $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ can be meaningfully arranged into separate matrices such that the act of forcing the values of adjacent entries to be similar will promote smoothness in the estimated surfaces. The penalty terms included in (4) correspond to row ($q=r$) and column ($q=c$) penalties for these matrices. The estimated F and G surfaces will be approximately linear in both the X and s dimensions for large values of the tuning parameters. On the other hand, allowing the tuning parameters to be close to zero allows for potentially complex surfaces that may be overfit, resulting in estimated decision rules that perform poorly on independent data.

Because of identifiability issues, the loss in (4) is not directly minimized. Instead, a loss function of the same form but involving suitable transformations of the covariates and parameters is minimized such that identifiability constraints are satisfied. The identifiability constraints and corresponding loss function are described in Appendix B.

Clearly, the tuning parameters need to be determined prior to minimizing the loss in (4). A common approach, which we take here, for selecting the smoothing tuning parameters is to use generalized cross validation (GCV). The fitting procedure, including tuning parameter selection, can be done easily using the `pfr` function from the R package `refund` (Crainiceanu *et al.*, 2014). Further information on the penalty terms and tuning parameter selection is discussed in McLean *et al.* (2014).

The quality of the estimated decision rule clearly depends on the quality of the estimated Q -function. As noted, using a fGAM for the Q -function allows for increased flexibility and increased protection against model mis-specification. However, this flexibility comes at the cost of having to estimate a large number of spline coefficient parameters. With this large number of parameters comes increased variability in the estimates which can result in poor performance of the ITR in realistic settings in which sample sizes are not large. Therefore, although the Q -learning approach offers improvement with respect to flexibility over more restrictive modeling, it may not be practical if the sample size is small. This prompts us to consider an alternative method that allows for flexible estimation of the contrast but requires estimation of fewer parameters.

3.2. A-Learning with fGAM (A-fGAM)

One commonly employed alternative to Q -learning in “classical” settings (i.e., settings with only scalar pre-treatment covariates) is A -learning. A -learning, also known as G -estimation (Robins, 2004), focuses on finding quality estimates of the parameters involved in the contrast function, $C(X)$, since these play a direct role in the OITR. The A -learning approach consists of positing a semiparametric model for the mean of the response and solving a set of appropriately selected estimating equations involving the propensity score. Solutions to these equations yield estimates of the contrast parameters that are consistent and asymptotically normal provided that the contrast is modeled correctly and that the propensity score is either known or is consistently estimated. Here we describe a parallel approach in settings with functional pre-treatment covariates in which we employ fGAMs and discuss similarities and differences between ITRs derived in this setting and those derived in the classical setting.

In the A -learning framework, the model that we consider is given by

$$E(Y|X, A) = h_0(X) + A \left\{ \theta_1 + \int_0^1 G\{X(s), s\} ds \right\}, \quad (5)$$

where h_0 is an unspecified “baseline” function of the of the pre-treatment covariate. Using the same representation for G given in Section 3.1, (5) can be expressed as

$$E(Y|X, A) \approx h_0(X) + A \left\{ \theta_1 + \sum_{j=1}^{C_x} \sum_{k=1}^{C_s} \gamma_{j,k} V_{i,j,k} \right\}.$$

As we are primarily interested in obtaining estimates for θ_1 and γ , we ideally seek to minimize the loss function given by

$$\ell_A(\xi; \lambda_r^\gamma, \lambda_c^\gamma) = \sum_{i=1}^n \left[Y_i - \phi(X_i) - \left\{ \theta_1 + \sum_{j=1}^{C_x} \sum_{k=1}^{C_s} \gamma_{j,k} V_{i,j,k} \right\} \cdot \{A_i - \pi(X_i)\} \right]^2 + \lambda_c^\gamma P_c(\gamma) + \lambda_r^\gamma P_r(\gamma),$$

(6)

where ϕ is an arbitrary function of the pre-treatment covariate X which may or may not approximate h_0 well. In the observational study setting we can model the propensity score (e.g., via logistic regression) and replace $\pi(X_i)$ with $\hat{\pi}(X_i)$ predicted from the model. As in Section 3.1, rather than minimizing (6) directly, identifiability constraints prompt us to consider minimizing a loss function similar to that described in Appendix B.

Omitting the penalty terms, the estimating equations associated with the loss function given by (6) have the same form as those solved in estimating the parameters in the A -learning framework. This is illustrated in Appendix C. In the “classical” A -learning setting, the pre-treatment covariates consist of only scalar values and the contrast parameters of interest can be consistently estimated, provided that the contrast is correctly specified and that the propensity scores are either known or can be estimated consistently. This means that the contrast estimates are robust to misspecification of $h_0(X)$. Hence the modeler can essentially ignore this component of the model for large samples.

Our setting is more complex owing to the functional nature of the pre-treatment covariate. Bias in the estimated contrast may be expected as a result of (1) using a finite set of basis functions to represent the surface G and (2) using a penalized fitting procedure. However, we note that it may be reasonable to expect that the estimates derived for θ_0 and G from the A -learning approach enjoy properties similar to those contrast parameters derived from A -learning in the classical setting provided that a large enough set of basis functions are considered in the representation of G and that the penalties reflect true underlying properties of the surface. If these conditions hold, then applying a flexible A -learning approach in our setting will yield high quality contrast estimates and correspondingly high quality ITRs. Ciarleglio *et al.* (2015) presented evidence from Monte Carlo experiments that this is indeed the case in settings where the contrast can be modeled as a linear function of pre-treatment covariates, including both scalars and functions. This suggests that combining A -learning with a flexible contrast model such as the one that we propose here may provide additional protection against model misspecification resulting in high quality ITRs based on functional pre-treatment covariates.

Although it is not necessary to specify ϕ , it is common to employ some parametric form. Since the form of ϕ used should not have an impact on the contrast estimates for a large sample size, it is advantageous to consider a simple working model. Obviously, if we let $\phi(X) = \theta_0 + \int_0^1 F\{X(s), s\} ds$ then we are using the same model discussed in Section 3.1. Two simpler models that may be considered are $\phi(X) = \theta_0$ or $\phi(X) = \theta_0 + \int_0^1 \beta(s) X(s) ds$ and either of these potentially require far fewer parameters to estimate than the more complex working model. Note that additional penalty terms may be required in (6) depending on the model chosen for ϕ . In Ciarleglio *et al.* (2015), constant and linear forms for ϕ were investigated and it was determined that the richer working model, even if mis-specified, resulted in contrast estimates that yielded better performing treatment decision rules.

3.3. Note on Adding Covariates

Although we have presented the Q -fGAM and A -fGAM approaches with only one functional pre-treatment covariate, both methods allow for the inclusion of multiple functional and scalar covariates. The extension is straightforward, however, including these additional covariates can substantially increase the number of parameters that need to be estimated.

Not all relationships between the pre-treatment covariates, treatment, and the response need to be left unspecified. In fact, the user can posit models in which some covariates exhibit linear effects on the response while others exhibit more general additive effects. See McLean *et al.* (2014) for further discussion on including multiple functional and scalar covariates in fGAMs as well as discussion on useful transformations of functional covariates.

4. Numerical Investigations

Here we investigate the finite sample properties of the proposed flexible functional Q and A -learning approaches in various settings. In these Monte Carlo experiments we are interested in comparing the performance of the two approaches with respect to the expected value of the response under estimated optimal treatment and the probability of selecting the correct treatment. We also compare the performance of the new functional approaches to an approach that uses Q -learning but that models the Q -function as a GAM depending on the scalar mean value of the functional pre-treatment covariate.

4.1. Simulation Setup

Treatment and Pre-Treatment Covariate—For each data set we generate functional pre-treatment covariates $X(s) = 1.5 \sum_{j=1}^{15} \beta_j b_j(s)$ with $\beta_j \sim N(0.25, 0.25^2)$ and b_j are order-4 B -spline basis functions with equally spaced knots on $[0, 1]$ for $j = 1, \dots, 15$. The curves are sampled at 100 equally spaced locations on $[0, 1]$. Figure 1 displays ten realizations of X . We also generate a treatment indicator, $A \in \{0, 1\}$, under two settings: (1) in which treatment received is random such that $\pi(X) = 0.5$ and (2) in which treatment received depends on X such that $\pi(X) = \frac{1}{1+e^{-\zeta}}$ where $\zeta = -1.75 + 4 \int_0^1 X(s) ds$.

Response—We generate the responses from the models given by

$Y = \theta_0 + \int_0^1 F_\delta\{X(s), s\}dt + \theta_1 A + A \int_0^1 G_j\{X(s), s\}ds + \varepsilon$, where $\delta \in \{-10, -1, 1\}$ and $j \in \{1, 2, 3\}$. We consider all F_δ and G_j pairs with

- $F_\delta\{X(s), s\} = \alpha e^{-\delta^2(s-0.5)^2} \sin\{2^\delta \pi X(s)\}$; $(\alpha, \delta) = (250, -10), (70, -1), (60, 1)$,
- $G_1\{X(s), s\} = 14 \sin(\pi s) X(s) - 3.5$,
- $G_2\{X(s), s\} = 175[(s-0.5)\{X(s)-0.5\}]^2 - 1.5$,
- $G_3\{X(s), s\} = 13e^{-4\{X(s)-1\}^2 - 11.1(s-0.5)^2} - 1.7$,

and $\varepsilon \sim N(0, \sigma^2)$, with σ selected such that $\text{var}(Y)/\text{var}(\varepsilon) = 2$. Values of θ_0 and θ_1 are provided in Table 1. Plots of the F_δ and G_j surfaces are displayed in Figure 3. We note that, for each model considered, we have $\text{var}[\int_0^1 F_\delta\{X(s), s\}]/\text{var}[\int_0^1 G_j\{X(s), s\}] \in (1, 1.6)$ (i.e., in all settings, the variation in the interaction effect is comparable to the variation in the main effect).

4.2. Methods Compared

We employ three different methods for estimating ITRs in each of the settings described above:

- **Q-fGAM**: Q -learning using the model given in (1) and representation given by (3); we let $B_x = B_s = C_x = C_s = 5$ with all sets of basis functions comprised of order-4 B -splines; row and column penalties as described in Appendix A,
- **A-fGAM**: A -learning corresponding to the loss given in (6); we let $C_x = C_s = 5$ with all sets of basis functions comprised of order-4 B -splines; we let $\phi(X) = \theta_0 + \int_0^1 \beta(s) X(s) ds$ and use the following order-4 B -spline representation $\beta(s) = \sum_{j=1}^{10} \alpha_j a_j(s)$ with roughness penalty given by $\lambda_\beta \int_0^1 [\frac{\partial^2}{\partial s^2} \beta(s)]^2 ds$,
- **Q-GAM**: We first reduce X to a scalar quantity, $\bar{X} = \int_0^1 X(s) ds$; fit the GAM corresponding to $E(Y | \bar{X}, A) = \theta_0 + f(\bar{X}) + \theta_1 A + A g(\bar{X})$ where f and g are unspecified functions each represented by linear combinations of ten order-4 B -splines.

We choose to include the Q -GAM approach for comparison since the reduction of functional data to “expertderived” scalar values is frequently carried out in practice (e.g., Tenke *et al.* (2011); Bruder *et al.* (2008)). For settings in which treatment is randomly assigned, we set $\pi(X_i) = 0.5$ for all i when using the A -fGAM approach. For settings in which treatment assignment depends on X , we predict the propensity score, $\hat{\pi}(X_i)$, for each observation using functional logistic regression via the penalized functional regression approach (Goldsmith *et al.*, 2011) and treat the predicted propensity scores as known. We employ a correctly specified propensity score model: $\text{logit}\{\pi(X)\} = \nu + \int_0^1 \omega(s) X(s) ds$.

For all methods we use GCV to select tuning parameters with $\kappa = 1$ (see Appendix A). We use the `pfr` function from the R package `refund` (Crainiceanu *et al.*, 2014) to fit the Q -

fGAM and A-fGAM ITRs and use the `gam` function from the R package `mgcv` (Wood, 2016) to fit the Q-GAM ITRs.

We consider three total sample sizes in each setting: $n=100$, 200, or 400. For each setting and sample size combination we simulate 250 training sets and one test set with 10,000 observations for evaluating performance of the estimated ITRs.

4.3. Evaluation Metrics

We evaluate the performance of the proposed approaches using two measures computed after applying the estimated ITRs to the independent validation data. Let $V(d) = E\{Y^*(d)\}$, the mean response under treatment regime d in the population and let $R(\hat{d}) = E\{V(\hat{d})\}/V(d^{opt})$ where d^{opt} is the true optimal decision and expectation in the numerator is taken with respect to the distribution of the parameters estimated in $\hat{d}$. $R(\hat{d})$ provides a measure of efficiency that $\hat{d}$ achieves in performance relative to the true optimal regime. Clearly, if $R(\hat{d})$ is close to 1 then $\hat{d}$ is considered to perform well. We approximate $V(d^{opt})$ using the independent validation data in each setting. Measures of efficiency for the rules $d(X) = 0$ and $d(X) = 1$ (i.e., give everyone treatment 0 or 1 respectively) are given by $R(0)$ and $R(1)$ respectively and are shown in Table 1 for each simulation setting.

We also assess the accuracy of each method for estimating ITRs by computing the percent of correct decisions (PCD) made. This is given by $PCD(\hat{d}) = 1 - E(|\hat{d} - d^{opt}|)$, where again, expectation is taken with respect to the distribution of the parameters estimated in $\hat{d}$. We provide the Monte Carlo mean (standard deviation) for each method/setting combination.

4.4. Summary of Simulation Results

Figure 3 shows box plots of the $R(\hat{d})$ values for each the 250 simulated data sets under each of the nine different response models when treatment is assigned randomly with $\pi(X_i) = 0.5$. Figure 4 is similar but for settings in which treatment assignment depends on the functional pre-treatment covariate. The values on the left side of each panel show the mean (sd) of the $PCD(\hat{d})$ values.

We focus on results from Figure 3 since the findings in Figure 4 are very similar. Across all settings and sample sizes, the Q-fGAM and A-fGAM approaches show comparable efficiency as well as comparable mean PCD values. We note that when F_{-10} serves as the baseline function (top three panels), the working model, ϕ , employed in the A-fGAM approach is approximately correctly specified. Thus we thought it reasonable to expect that ITRs estimated via the A-fGAM approach would perform better than those estimated via the Q-fGAM approach in these settings. However we see that there is no clear difference in the performance of the two methods with respect to efficiency or PCD .

When the response is generated with either G_2 or G_3 (i.e., surfaces that are non-linear in both X and s), there is clear evidence that the A-fGAM and Q-fGAM approaches yield superior ITRs compared to those derived from the Q-GAM approach (“QS” in box plots) especially for larger sample sizes. When the response is generated with G_1 (i.e., surface that is linear in X and non-linear in s), the Q-GAM approach produces surprisingly decent ITRs

compared to those derived from the Q -fGAM and A -fGAM approaches particularly for smaller sample sizes.

5. Application to Psychiatric RCT Data

To illustrate our proposed approaches for developing treatment decision rules we consider data from an ongoing multisite clinical trial investigating heterogeneous treatment response in MDD patients. Upon study enrollment, investigators collect data from each subject on a multitude of pre-treatment variables including measures from various imaging modalities. After all pre-treatment variable information is collected, each subject is randomized to either sertraline, a selective serotonin reuptake inhibitor (SSRI), or placebo. Subjects are then monitored via assessments at 1, 2, 3, 4, 6, and 8 weeks after initiation of treatment. Several possible outcome measures are collected at each of these time points, but here we focus on the Hamilton Rating Scale for Depression (HRSD) score. The score is based on responses to 17 items which together provide information on severity of depression. Higher values indicate more severe depression. In our application we use the change in HRSD score from baseline (before treatment is administered) to the eighth week of the study as our outcome of interest (week 8 - baseline). Thus lower values indicate a greater reduction in depression symptom severity.

We consider several treatment decision rules that depend on a pre-treatment functional covariate derived from resting electroencephalography (EEG) measures. EEG is a relatively inexpensive and noninvasive tool for measuring certain aspects of brain function and it is of considerable interest to be able to use data derived from EEG experiments to develop quality antidepressant treatment decision rules. In the study, subjects wear a cap embedded with 72-electrodes which provide measures of gross neuronal activity at each electrode's location. Subjects sit with their eyes open and an activity time series is recorded at each electrode over a two-minute period. Each time series is converted to a current source density (CSD) power spectrum value ($\mu V/m^2$) curve which provides information about rhythms in brain activity at a given electrode. Prior studies have suggested that posterior theta and alpha rhythm of the EEG may be useful in differentiating patients who do or do not respond to SSRIs or other treatments for MDD (Bruder *et al.*, 2008; Tenke *et al.*, 2011). Thus we focus on using the CSD curve from the Oz electrode, located in the posterior region of the electrode cap, over a frequency range of 3 to 16 Hz (theta and alpha bands) as the pre-treatment functional covariate. Figure 5 shows the curves for 20 subjects in the sample that we consider.

We estimated treatment decision rules using data from 121 subjects with complete information on baseline HRSD score, week 8 HRSD score, treatment assignment, and Oz CSD power curve. In these 121 subjects, 56 were randomized to sertraline. The mean change in HRSD over 8 weeks is -5.94 for those on placebo and -8.13 for those on sertraline.

We consider five approaches for assigning treatment: (1) use the estimated rule from Q -fGAM; (2) use the estimated rule from A -fGAM; (3) use the estimated rule from Q -GAM; (4) assign everyone placebo; (5) assign everyone sertraline. For (1) – (3) we estimated the ITRs using the same procedures described in Section 4.2. In the estimation procedures, we used the opposite of the change in HRSD score as the response when estimating the ITRs

with each approach so that larger values are in fact desirable. We also scaled the domain of the CSD curves to be $[0, 1]$. For each approach used we computed the K -fold cross-validated alternative inverse probability weighted estimator (AIPWE) given by

$$AIPWE_{CV}(\hat{d}) = \frac{1}{K} \sum_{k=1}^K \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \left\{ \frac{C_{\hat{d}_k, i}^{(k)} Y_i^{(k)}}{\pi(X_i^{(k)})} - \frac{C_{\hat{d}_k, i}^{(k)} - \pi(X_i^{(k)})}{\pi(X_i^{(k)})} \hat{m}_Q^{(k)}(X_i^{(k)}) \right\} \right], \text{ where}$$

$C_{\hat{d}_k, i}^{(k)} = A_i^{(k)} \hat{d}_k(X_i^{(k)}) + (1 - A_i^{(k)}) \{1 - \hat{d}_k(X_i^{(k)})\}$ indicates whether subject i from test split k actually received the treatment recommended by $\hat{d}_k$, the ITR estimated using training split k , $\hat{m}_Q^{(k)}(X_i^{(k)}) = \hat{Q}^{(k)}(X_i^{(k)}, 1) \hat{d}_k(X_i^{(k)}) + \hat{Q}^{(k)}(X_i^{(k)}, 0) \{1 - \hat{d}_k(X_i^{(k)})\}$ with $\hat{Q}^{(k)}$ given by the model fit using the Q -fGAM method on training split k and $\pi(X_i^{(k)})$ is taken to be 0.5. We set $K = 5$ and repeated the CV procedure 100 times for each of the five approaches for assigning treatment.

In the table on the left side of Figure 6 we report $\overline{AIPWE}_{CV}$, the mean (sd) $AIPWE_{CV}$ value for five competing regimes over the 100 splits. (Note that the response values here are the actual change scores rather than the opposites.) The Q -fGAM and A -fGAM approaches appear to give slightly better ITRs on average although the additional reduction in HRSD score resulting from using the ITRs derived from these methods may not be clinically meaningful compared to the simple rule of giving everyone sertraline. One can assess the need for the more complicated models presented here by plotting the estimated surface $\hat{G}$ as well as inspecting its partial second derivatives with respect to X for fixed values of s . These plots are shown in Figure 7. Note that a flexible model for the response may be appropriate here since there is evidence of non-linearity along both the X and s dimensions in the surface plot. We also observe non-zero second partial derivatives with respect to X for several values of s .

6. Discussion

Development of ITRs that are tailored for individuals is of critical importance. This is especially true for diseases like major depressive disorder for which there is substantial heterogeneity in response for most existing treatments. Furthermore, in order to improve treatment recommendations, the methods for estimating ITRs need to allow for the inclusion of functional data such as those derived from various imaging modalities. We have proposed two novel approaches for estimating treatment decision rules that depend on functional pre-treatment covariates. Both approaches allow for flexible estimation of the possibly complex relationship between functional covariates, treatment, and a scalar response. This flexibility provides protection against model mis-specification. This is important since estimates derived from mis-specified response or contrast models can negatively affect the performance of the ITRs that depend on those estimates.

Our proposed Q -fGAM approach posits a model for the response which may require the estimation of many parameters. On the other hand, the A -fGAM approach focuses on modeling the contrast and can require estimation of far fewer parameters, depending on the working model for ϕ , thus making it a viable option when small sample size is an issue. Our

simulations suggest that the Q -fGAM and A -fGAM approaches typically result in similarly performing ITRs; lending more support for the use of the A -fGAM method.

One practical limitation of our proposed methods is that the number of parameters that need to be estimated increases quickly with (1) the number of functional covariates that are included in the model and (2) the number of basis functions needed to adequately represent F_j and G_j corresponding to functional covariates $X_j, j = 1, \dots, J$. Hence these methods are best suited for settings in which there are few informative pre-treatment functional covariates (i.e., J is small) and the sample size is large enough to allow for use of adequately rich sets of basis functions to represent each F_j and G_j surface. This highlights the fact that the important pre-treatment covariates must be known in advance based on clinical knowledge. Unfortunately, such knowledge may be unavailable. This is particularly problematic in research on treatment selection for MDD. Therefore incorporating a variable selection procedure with our flexible ITR estimation approaches is of considerable interest.

Zhang *et al.* (2012a) recently proposed a framework in which an ITR is estimated via a two-stage approach in settings with scalar pre-treatment covariates. The first stage consists of deriving a high quality estimate of the contrast for each subject and the second stage uses those contrast values to create weights and corresponding binary responses which can be input into a classification algorithm that can determine the OITR by minimizing the weighted misclassification error. Zhang *et al.* (2012a) claim that the separation of the two stages results in increased flexibility and allows for the use of a wide range of classification algorithms. Such an approach can be extended to settings in which functional pre-treatment covariates are available. Our proposed approaches can be used in the first stage of such a procedure to provide quality contrast estimates and therefore weights and binary responses for the second stage. Methods such as functional logistic regression can be employed for the subsequent minimization of the weighted misclassification error. We intend to investigate such two stage approaches in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by MH099003-01 from the National Institutes of Health. This work has utilized computing resources at the High Performance Computing Facility of the Center for Health Informatics and Bioinformatics at NYU Langone Medical Center.

References

- Bruder G, Sedoruk J, Stewart J, McGrath P, Quitkin F, Tenke C. Electroencephalographic alpha measures predict therapeutic response to selective serotonin reuptake inhibitor antidepressant: Pre and post-treatment findings. *Biological Psychiatry*. 2008; 63:1171– 1177. [PubMed: 18061147]
- Ciarleglio A, Petkova E, Ogden R, Tarpey T. Treatment decisions based on scalar and functional baseline covariates. *Biometrics*. 2015; 71:884– 894. [PubMed: 26111145]
- Ciarleglio A, Petkova E, Ogden R, Tarpey T. Variable selection for treatment decision rules with scalar and functional predictors. 2016 (under review).

- Crainiceanu, C., Reiss, P., Goldsmith, J., Huang, L., Huo, L., Scheipl, F., Swihart, B., Greven, S., Harezlak, J., Kundu, M., Zhao, Y., McLean, M., Xiao, L. refund: Regression with Functional Data (R Package version 0.1–13). 2014.
- Goldsmith J, Bobb J, Crainiceanu C, Reich D. Penalized functional regression. *Journal of Computational and Graphical Statistics*. 2011; 20:830–851. [PubMed: 22368438]
- Laber E, Zhao Y. Tree-based methods for individualized treatment regimes. *Biometrika*. 2015; 102:501–514. [PubMed: 26893526]
- Lu W, Zhang H, Zeng D. Variable selection for optimal treatment decision. *Statistical Methods in Medical Research*. 2013; 22:493–504. [PubMed: 22116341]
- McKeague I, Qian M. Estimation of treatment policies based on functional predictors. *Statistica Sinica*. 2014; 24:1461–1485. [PubMed: 25165416]
- McLean M, Hooker G, Staicu A, Scheipl F, Rupert D. Functional generalized additive models. *Journal of Computational and Graphical Statistics*. 2014; 23:249–269. [PubMed: 24729671]
- Moodie E, Dean N, Sun Y. Q-learning: Flexible learning about useful utilities. *Statistics in Biosciences*. 2014; 6:223–243.
- Qian M, Murphy S. Performance guarantees for individualized treatment rules. *Annals of Statistics*. 2011; 39:1180–1210. [PubMed: 21666835]
- Ramsay, JO., Silverman, BW. *Functional Data Analysis*, Second Edition. Springer; New York: 2005.
- Robins, J. Optimal structured nested models for optimal sequential decisions. In: Lin, D., Heagerty, P., editors. *Proceedings of the Second Seattle Symposium on Biostatistics*. Springer; New York: 2004. p. 189–326.
- Rubin D. Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*. 1978; 6:34–58.
- Tenke C, Kayser J, Manna C, Fekri S, Kroppmann C, Schaller J, Alschuler D, Stewart J, McGrath P, Bruder G. Current source density measures of electroencephalographic alpha predict antidepressant treatment response. *Biological Psychiatry*. 2011; 70:388–394. [PubMed: 21507383]
- Tian L, Alizadeh A, Gentles A, Tibshirani R. A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*. 2014; 109:1517–1532. [PubMed: 25729117]
- Watkins C, Dylan P. Q-learning. *Machine Learning*. 1992; 8:279–292.
- Wood, S. R Package version 1.8–11. 2016. mgcv: Mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation.
- Zhang B, Tsiatis A, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. *Stat*. 2012a; 1:103–114. [PubMed: 23645940]
- Zhang B, Tsiatis A, Laber E, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012b; 68:1010–1018. [PubMed: 22550953]
- Zhao Y, Zeng D, Rush A, Kosorok M. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*. 2012; 107:1106–1118. [PubMed: 23630406]
- Zhou X, Mayer-Hamblett N, Khan U, Kosorok M. Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association*. 2015; doi: 10.1080/01621459.2015.1093947

Figure 1.
10 realizations of X for in simulation experiments.

Figure 2.
 F_{δ} and G_j functions used in simulation settings.

Figure 3.

Box plots of $R(\hat{d})$ efficiency values (on horizontal axis) for all settings with treatment randomly assigned. Colored values on the left side of each panel are mean (sd) PCD values for the corresponding box plot on the right. Labels on the right side of each panel show the method (Q = Q-fGAM, A = A-fGAM, QS = Q-GAM) and sample size. Row headings indicate which F_δ under consideration and column headings indicate which G_j under consideration.

Figure 4.

Box plots of $R(\hat{d})$ efficiency values (on horizontal axis) for all settings with treatment randomly assignment depending on X . Colored values on the left side of each panel are mean (sd) PCD values for the corresponding box plot on the right. Labels on the right side of each panel show the method (Q = Q-fGAM, A = A-fGAM, QS = Q-GAM) and sample size. Row headings indicate which F_δ under consideration and column headings indicate which G_j under consideration.

Figure 5. Twenty CSD curves from the Oz electrode under the eyes-open condition for subjects on placebo (red) and on sertraline (blue).

Regime	$\overline{AIPWE}_{CV}$ (sd)
(1) Q-fGAM	-8.34 (0.50)
(2) A-fGAM	-8.43 (0.49)
(3) Q-GAM	-8.00 (0.49)
(4) Placebo	-5.83 (0.19)
(5) Sertraline	-8.07 (0.19)

Figure 6.

Left: Mean (sd) of the 100 $AIPWE_{CV}$ values for each regime. **Right:** Histograms of the 100 $AIPWE_{CV}$ values for the A-fGAM and Sertraline approaches.

Figure 7. **Left:** Estimated $\hat{G}$ surface from A -fGAM ITR using all 121 observations. **Right:** Plots of second partial derivatives of $\hat{G}$ with respect to X for $s = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$.

Table 1

Table of simulation parameters. The efficiency measures $R(0)$ and $R(1)$, defined in Section 4.3, are approximated using 10^5 simulated observations.

δ	G_1				G_2				G_3			
	θ_0	θ_1	$R(0)$	$R(1)$	θ_0	θ_1	$R(0)$	$R(1)$	θ_0	θ_1	$R(0)$	$R(1)$
	θ_0	θ_1	$R(0)$	$R(1)$	θ_0	θ_1	$R(0)$	$R(1)$	θ_0	θ_1	$R(0)$	$R(1)$
-10	-3	-0.50	0.82	0.73	-3	-0.50	0.77	0.84	-3	-0.50	0.79	0.80
-1	-3	-0.50	0.91	0.87	-3	-0.50	0.89	0.92	-3	-0.50	0.90	0.90
1	2	-0.50	0.91	0.86	2	-0.50	0.88	0.92	2	-0.50	0.89	0.90