

The International Journal of Biostatistics

Volume 6, Issue 2

2010

Article 10

CAUSAL INFERENCE

Optimal Dynamic Regimes: Presenting a Case for Predictive Inference

Elja Arjas*

Olli Saarela†

*University of Helsinki and National Institute for Health and Welfare, elja.arjas@helsinki.fi

†National Institute for Health and Welfare, olli.saarela@thl.fi

Optimal Dynamic Regimes: Presenting a Case for Predictive Inference

Elja Arjas and Olli Saarela

Abstract

Dynamic treatment regime is a decision rule in which the choice of the treatment of an individual at any given time can depend on the known past history of that individual, including baseline covariates, earlier treatments, and their measured responses. In this paper we argue that finding an optimal regime can, at least in moderately simple cases, be accomplished by a straightforward application of nonparametric Bayesian modeling and predictive inference. As an illustration we consider an inference problem in a subset of the Multicenter AIDS Cohort Study (MACS) data set, studying the effect of AZT initiation on future CD4-cell counts during a 12-month follow-up.

KEYWORDS: Bayesian nonparametric regression, causal inference, dynamic programming, monotonicity, optimal dynamic regimes

1 Introduction

A dynamic (sequential) treatment regime is a decision rule which determines the treatment decisions to be taken over time, and it can in general depend on both the patient's earlier known history and the anticipated future consequences of the decision. The definition of optimality of such a rule can involve formal consideration of loss or utility functions; here, however, we take optimality to simply mean the treatment regime which produces the highest expected value for a response variable of interest at the end of the study period.

Important contributions to the study of optimal treatment regimes have been provided, in particular, by Robins (1986, 1994, 2004) and Murphy (2003). The respective approaches of the two authors are summarized and compared by Moodie et al. (2007). Murphy (2003) proposed methodology for estimating optimal dynamic regimes from observational data. As discussed in this paper, Arjas (2003) suggested to solve the inferential problem using simulation from the predictive distributions of potential outcomes of the response variable, given alternative treatment regimes (see also Arjas and Parner, 2004). In contrast, Murphy parameterizes the problem directly in terms of differences of mean response values of the different treatment regimes to the mean of the optimal regime, calling these functions *regrets*. Basically, any unbiased estimation of 'causal effects' from observational data requires an assumption of *no unmeasured confounders*, meaning that the treatment decisions made at each point in time are based only on observed data that are available also in the statistical analysis. Rosthøj et al. (2006) present an application of Murphy's method in a situation where the treatment variable (dose of anticoagulation medication) is continuous, concentrating on parametric regret functions. In addition to the regret functions, the estimation procedure requires the determination of probability distributions for the treatment actions.

Determination of optimal decisions in more conventional likelihood-based analysis requires the modeling of the longitudinal distribution of the data, combined with backward induction (also known as *dynamic programming*) in time. This approach requires working out all the future optimal decisions, backwards from the last observation point to the first. One way to achieve this, at least in the case of discrete decision alternatives, is to draw samples from the predictive distributions estimated from the observed data, for each of the alternative treatment options. The expectations are then evaluated using Monte Carlo integration.

Carlin et al. (1998) describe this approach in the context of optimal stopping rules in clinical trials, where the decisions to be made are binary (stop the trial if a large enough positive or negative treatment effect has been ob-

served, or continue the trial to collect more data). If determination of one expectation using Monte Carlo integration requires a sample of size m , using nested simulation rounds through k backward steps would require a total of $m(m^k - 1)/(m - 1)$ simulations. For large values of k this leads to computational problems, and therefore Carlin et al. describe also a forward sampling algorithm applicable in certain special cases of likelihood functions. Here the nested integration can be replaced with an optimization task over $2k - 1$ decision rule boundaries. Also in the context of sequential clinical trials, Wathen and Thall (2008) revamp this algorithm by assuming more restrictive parametric forms for the boundaries. To account for uncertainty in the model definition, they incorporate a model selection based on approximate Bayes factors over a preselected set of candidate models.

Moodie et al. (2007) demonstrate optimal dynamic regime analysis based on modeling of regret or blip (Robins, 2004) functions. They argue that likelihood-based modeling of longitudinal data is sensitive to model misspecification. However, it is unclear whether misspecification of the regret functions is any less serious of a problem. In this paper we aim to wipe the dust off the combination of likelihood and backward induction in optimal dynamic regime analysis and claim that it is still a viable alternative for such purpose. Probability-based data analysis carries some obvious advantages over alternative methods, including the ability to handle missing data and providing a natural quantification of uncertainties at different levels of model specification. In order to avoid problems associated with too rigid model assumptions we apply Bayesian nonparametric monotonic regression, where a monotonicity assumption is used in place of more restrictive parametric or structural assumptions. In contrast to Wathen and Thall (2008), in our approach model selection is built into a part of the probability model and statistical inference is based on model averaged results over random realizations drawn from the space of possible models. We argue that, at least in situations where the decision space is small, the optimal regimes can be worked out using standard MCMC techniques, with the optimal decisions obtained as a byproduct of a single sequence of MCMC samples. We illustrate this approach by using the same example as in Moodie et al. (2007), based on data from Multicenter AIDS Cohort (MACS) study (Kaslow et al., 1987).

The plan of the paper is as follows: In the next section we briefly reintroduce the data and the notation of Moodie et al. (2007), and then present our probability model for the data. In Section 3 we discuss probability-based inference of the optimal treatment regime and the computational issues involved. We also pay attention to the problems caused by missing data and censoring

to causal inference in such observational settings. The paper concludes with a discussion in Section 4.

2 Statistical model

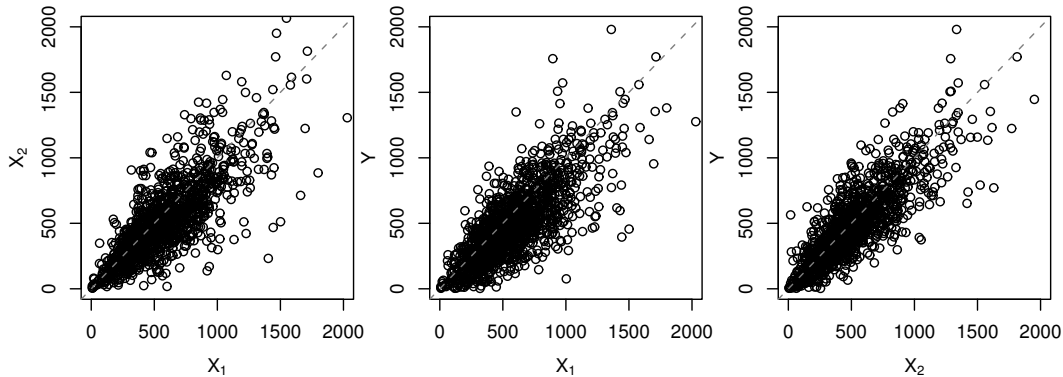
Multicenter AIDS Cohort study (MACS) was a longitudinal observational study starting from 1984 and following up HIV-positive homosexual and bisexual men recruited from four U.S. cities. As in Moodie et al. (2007), we restrict our attention to a small subset of this dataset, consisting of the first three consecutive follow-up visits (separated by approximately six months each) after antiviral drug azidothymidine (AZT) became available. We included HIV-positive individuals who did not have (self-reported) AIDS defining illness at the first eligible visit. The interest is in studying the effect of AZT treatment initiated before the 6 or 12 month follow-up visits to the response variable, which is the count of CD4-positive helper cells measured at the 12 month visit. First suppressing an index i referring to an individual, let

- X_1 = CD4-cell count at the beginning of the study,
- $A_1 = 1$ if AZT treatment initiated at $0 < t < 6$, and $A_1 = 0$ otherwise;
- X_2 = CD4-cell count at $t = 6$ months;
- $A_2 = 1$ if AZT treatment initiated at $6 < t < 12$ (only possible if $A_1 = 0$), and $A_2 = 0$ otherwise;
- Y = CD4-cell count at $t = 12$ months.

Further, let $\mathcal{A}_1 = \{i : A_{1i} = 1\}$, $\mathcal{A}_0 = \{i : A_{1i} = 0\}$, $\mathcal{A}_{01} = \{i : A_{1i} = 0, A_{2i} = 1\}$, and $\mathcal{A}_{00} = \{i : A_{1i} = 0, A_{2i} = 0\}$ denote the sets of individuals in the data observed in the different treatment alternatives. The observed counts in the dataset were $|\mathcal{A}_1| = 132$, $|\mathcal{A}_0| = 1295$, $|\mathcal{A}_{01}| = 131$ and $|\mathcal{A}_{00}| = 1164$. Figure 1 displays the scatterplots of the three CD4-cell measurements, showing that there are strong serial correlations between the consecutive cell counts, and also that there is a visible decreasing trend in the individual CD4-levels as a function of time.

We assume that treatment assignment after the baseline visit can depend on the measured covariate value X_1 , but that it otherwise satisfies the condition of no unmeasured confounders. The MACS data include many possibly relevant measured covariates, both questionnaire based and laboratory results. Here, however, we take this condition to mean simply that, given the CD4-count X_{1i} of individual i at time $t = 0$, there are no additional known individual characteristics of i , represented in the model by (latent) random variables,

Figure 1: Scatterplots for three CD4-level measurements



that would predict the assignment of i to either $A_{1i} = 0$ or $A_{1i} = 1$, nor does such a prediction of assignment depend on the model parameters (considered in detail below) which we use for describing the development of the CD4-count levels Y_i and X_{2i} . For more formal definition of no unmeasured confounders, see e.g. Arjas and Parner (2004). In similar fashion, when considering the assignment of treatment alternatives A_{2i} after time $t = 6$, we assume that they can depend on the, then known, covariate readings X_{2i} and X_{1i} , but that they otherwise satisfy a similar no unmeasured confounders postulate.

The statistical model is now specified in such a way that it is conditional on the observed values of X_{1i} , $i = 1, \dots, n$, in the data. Moreover, as our statistical inferences will be based on forming the likelihood of the data, the above no unmeasured confounders implies that the likelihood contributions arising from treatment assignments can be treated as constant proportionality factors, and therefore have no influence on the inferences based on such likelihoods. Thus, under the no unmeasured confounders postulate we can specify our statistical model in such a way that it is at $t = 6$ conditioned directly on the observed values of both X_{1i} and A_{1i} . The model is specified as a modular structure, where each module corresponds to a possible transition from one ‘state’ to the next ‘state’. In the following, we use the notation $[.]$ to refer to both probability density and probability mass functions, depending on whether the random variate is continuous or discrete.

Module 1: Model structure for treatment branch $A_1 = 1$:

- Consider first a nonparametric Bayesian model for ‘the systematic part’ of predicting response Y_i of individual i , given covariate value X_{1i} at baseline $t = 0$. Denote this systematic part by $f_1(X_{1i})$. Since the func-

tion f_1 , which is assumed to be common to all individuals, is unknown, we view it to be random and will estimate it from the combined data. In order to facilitate this, we make the following assumption: f_1 is (that is, all realizations of f_1 are) *monotonically increasing*. Considering that both Y_i and X_{1i} are CD4-counts measured from the same individual i , this assumption would seem to be well justified.

- **Modeling individual variation:** It is obvious from Figure 1 that the variability of the CD4-levels over time depends on the absolute CD4-level. To take this into account, conditionally on f_1 and X_{1i} , and independently across individuals, we draw a random variable, say Z_{1i} , from a gamma distribution: $[Z_{1i} \mid A_{1i} = 1, X_{1i}, f_1, \alpha_1] = \text{Gamma}(\alpha_1 f_1(X_{1i}), \alpha_1)$. Here α_1 is a hyperparameter, which can be either specified separately or estimated from the data; here we choose the latter option and therefore specify a prior $[\alpha_1]$ for it. This parameter controls the variance of the gamma distribution. Note that $E(Z_{1i} \mid A_{1i} = 1, X_{1i}, f_1, \alpha_1) = f_1(X_{1i})$.
- **Modeling the response Y_i :** Conditionally on $(A_{1i} = 1, X_{1i}, f_1, Z_{1i})$, suppose that $[Y_i \mid A_{1i} = 1, X_{1i}, f_1, Z_{1i}] = \text{Poisson}(Z_{1i})$. Note that then $E(Y_i \mid A_{1i} = 1, X_{1i}, f_1, \alpha_1) = f_1(X_{1i})$.

In summary, we have postulated the following hierarchical model structure for the individuals $i \in \mathcal{A}_1$ in branch $A_1 = 1$, conditionally on the treatment assignments and the corresponding X_{1i} ,

$$\begin{aligned} & [\{Y_i, Z_{1i} : i \in \mathcal{A}_1\}, f_1, \alpha_1 \mid \{A_{1i} = 1, X_{1i} : i \in \mathcal{A}_1\}] \\ &= \left(\prod_{i \in \mathcal{A}_1} [Y_i \mid Z_{1i}][Z_{1i} \mid A_{1i} = 1, X_{1i}, f_1, \alpha_1] \right) [f_1][\alpha_1]. \end{aligned} \quad (1)$$

Module 2: The model for X_{2i} in the treatment branch $A_1 = 0$ is specified exactly as the model for Y_i above, with a different nonparametrically specified and monotonically increasing function f_0 and the related parameterization, resulting in a joint distribution of the form

$$\begin{aligned} & [\{X_{2i}, Z_{0i} : i \in \mathcal{A}_0\}, f_0, \alpha_0 \mid \{A_{1i} = 0, X_{1i} : i \in \mathcal{A}_0\}] \\ &= \left(\prod_{i \in \mathcal{A}_0} [X_{2i} \mid Z_{0i}][Z_{0i} \mid A_{1i} = 0, X_{1i}, f_0, \alpha_0] \right) [f_0][\alpha_0]. \end{aligned} \quad (2)$$

The next step in the model construction is to consider the progression in branch $A_1 = 0$ from time $t = 6$ onwards, up to Y measured at $t = 12$,

separately for the alternatives $A_2 = 1$ and $A_2 = 0$. The main difference to the above two constructions is that now there are two covariate measurements, X_1 and X_2 , in the history that should be accounted for when predicting Y . Before continuing the model specification, we return briefly to the issue of treatment assignment, now at time $6 < t < 12$. As stated earlier, we assume that A_2 can depend on, then known, covariate readings X_1 and X_2 (as well as, in obvious way, on A_1), but that they otherwise satisfy the no unmeasured confounders postulate. As a consequence, the likelihood contribution arising from observing the value of A_2 for an individual in the data can be treated as a constant, and therefore need not be considered when drawing statistical inferences, unless it is specifically of interest (in Section 3.2 we will note one possible situation where this might be the case).

Module 3: Considering first the case $A_1 = 0$ and $A_2 = 1$, we introduce a nonparametrically defined function $f_{01}(X_1, X_2)$, and assume that it is monotonically increasing in X_1 and $X_2 - X_1$. The motivation behind this choice (instead of using the two measured levels directly as covariates) was to separate the baseline level and the longitudinal trend, thus resulting in two less correlated variables, both of which are seemingly compatible with the monotonicity postulate. We postulate a hierarchical model structure

$$\begin{aligned} & [\{Y_i, Z_{01i} : i \in \mathcal{A}_{01}\}, f_{01}, \alpha_{01} \mid \{A_{1i} = 0, A_{2i} = 1, X_{1i}, X_{2i} : i \in \mathcal{A}_{01}\}] \\ & = \left(\prod_{i \in \mathcal{A}_{01}} [Y_i \mid Z_{01i}][Z_{01i} \mid A_{1i} = 0, A_{2i} = 1, X_{1i}, X_{2i}, f_{01}, \alpha_{01}] \right) [f_{01}][\alpha_{01}], \quad (3) \end{aligned}$$

where

$$[Z_{01i} \mid A_{1i} = 0, A_{2i} = 1, X_{1i}, X_{2i}, f_{01}, \alpha_{01}] = \text{Gamma}(\alpha_{01} f_{01}(X_{1i}, X_{2i}), \alpha_{01})$$

and $[Y_i \mid Z_{01i}] = \text{Poisson}(Z_{01i})$. Note that

$$E(Y_i \mid A_{1i} = 0, A_{2i} = 1, X_{1i}, X_{2i}, f_{01}, \alpha_{01}) = f_{01}(X_{1i}, X_{2i}).$$

Module 4: Finally, considering the combination $A_1 = 0$ and $A_2 = 0$, we introduce a random function $f_{00}(X_1, X_2)$, similarly monotonically increasing in X_1 and $X_2 - X_1$, and postulate the joint distribution

$$\begin{aligned} & [\{Y_i, Z_{00i} : i \in \mathcal{A}_{00}\}, f_{00}, \alpha_{00} \mid \{A_{1i} = 0, A_{2i} = 0, X_{1i}, X_{2i} : i \in \mathcal{A}_{00}\}] \\ & = \left(\prod_{i \in \mathcal{A}_{00}} [Y_i \mid Z_{00i}][Z_{00i} \mid A_{1i} = 0, A_{2i} = 0, X_{1i}, X_{2i}, f_{00}, \alpha_{00}] \right) [f_{00}][\alpha_{00}], \quad (4) \end{aligned}$$

where the distributions for Z_{00i} and Y_i are defined analogously to above.

3 Statistical inference

3.1 Determining the optimal treatment assignment

Given the observed data \mathcal{D} , the model parameters $f_0, f_1, f_{00}, f_{01}, \alpha_0, \alpha_1, \alpha_{00}$, and α_{01} can now be estimated. The nonparametric random functions are defined and estimated following Saarela and Arjas (2009). Our approach to the inferential problem is Bayesian, with the consequence that we obtain the joint posterior distribution of these parameters, and can then consider the corresponding posterior predictive distribution of how a generic individual drawn from the same population as the actual individuals in the data (that is, we are using the same model for both) would respond to different treatment alternatives. Denote the ‘potentially observed’ characteristics of such an individual by X_1^*, X_2^* and Y^* . Note that the choice of the values of A_1^* and A_2^* can be viewed as representing ‘do’-conditioning (Pearl, 2000). (That we can switch from observed treatment values in the data to ‘optional’ or ‘forced’ in the predictions is again consequence of the no unmeasured confounders postulate.) It is now straightforward to compute, for example, the posterior predictive expectation of the response Y^* at time $t = 12$, as a function of the CD4-count X_1^* when $A_1^* = 1$, that is, if the AZT treatment is initiated after the baseline visit. We get

$$\begin{aligned} E(Y^* \mid A_1^* = 1, X_1^*, \mathcal{D}) \\ &= E(E(Y^* \mid A_1^* = 1, X_1^*, f_1, \alpha_1, \mathcal{D}) \mid A_1^* = 1, X_1^*, \mathcal{D}) \\ &= E(f_1(X_1^*) \mid A_1^* = 1, X_1^*, \mathcal{D}), \end{aligned}$$

where the outer expectation in the second line is with respect to the unknown parameters f_1 and α_1 . Similarly,

$$\begin{aligned} E(X_2^* \mid A_1^* = 0, X_1^*, \mathcal{D}) &= E(f_0(X_1^*) \mid A_1^* = 0, X_1^*, \mathcal{D}), \\ E(Y^* \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) \\ &= E(f_{01}(X_1^*, X_2^*) \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) \end{aligned}$$

and

$$\begin{aligned} E(Y^* \mid A_1^* = 0, A_2^* = 0, X_1^*, X_2^*, \mathcal{D}) \\ &= E(f_{00}(X_1^*, X_2^*) \mid A_1^* = 0, A_2^* = 0, X_1^*, X_2^*, \mathcal{D}). \end{aligned}$$

The computation of these expectations is carried out in practice most easily alongside the MCMC sampling which is used for estimating the model parameters, applying data augmentation. A more detailed description of the computational algorithm is given in the Appendix.

To solve the main problem, that is, finding the optimal treatment regime, we need to consider how the choice between the three alternatives ($A_1^* = 1$), ($A_1^* = 0, A_2^* = 1$) and ($A_1^* = 0, A_2^* = 0$) should be made in a situation in which the measurement X_1^* is available at $t = 0$, and, if $A_1^* = 1$ is chosen, also X_2^* is measured at $t = 6$ in order to facilitate the choice between $A_2^* = 1$ and $A_2^* = 0$. Following the principles of dynamic programming, we consider first how this latter choice after $t = 6$ should be made, conditioning on the observed past history ($A_1^* = 0, X_1^*, X_2^*$) of a generic individual as described above, and then moving backwards in time in order to make a choice between branch ($A_1^* = 0$) and branch ($A_1^* = 1$).

After $t = 6$, the key comparison to be made is between the posterior expectations $E(Y^* | A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D})$ and $E(Y^* | A_1^* = 0, A_2^* = 0, X_1^*, X_2^*, \mathcal{D})$, both of which were obtained above. We now define the optimal treatment choice after $t = 6$ to be the one under which, for given values of X_1^* and X_2^* , the expected value of Y^* is larger, denoted by indicator

$$A_2^{*opt}(X_1^*, X_2^*) = 1_{\{E(Y^* | A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) > E(Y^* | A_1^* = 0, A_2^* = 0, X_1^*, X_2^*, \mathcal{D})\}}.$$

After $t = 0$, we then make a comparison between $E(Y^* | A_1^* = 1, X_1^*, \mathcal{D})$ and

$$E(E(Y^* | A_1^* = 0, A_2^{*opt}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) | A_1^* = 0, X_1^*, \mathcal{D}),$$

where the outer expectation is with respect to the posterior predictive distribution of X_2^* , and let

$$A_1^{*opt}(X_1^*) = 1_{\{E(Y^* | A_1^* = 1, X_1^*, \mathcal{D}) > E(E(Y^* | A_1^* = 0, A_2^{*opt}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) | A_1^* = 0, X_1^*, \mathcal{D})\}}.$$

The optimal dynamic treatment regime A^{opt} will thereby be defined as follows: For $0 < t < 6$ choose $A_1^{*opt}(X_1^*)$ and for $6 < t < 12$, if $A_1^{*opt}(X_1^*) = 0$, choose $A_2^{*opt}(X_1^*, X_2^*)$.

It is of some interest to compare results obtained when following the three deterministic regimes ($A_1^* = 1$), ($A_1^* = 0, A_2^* = 1$) and ($A_1^* = 0, A_2^* = 0$) to each other, and all these to A^{opt} . The comparison becomes most clear and intuitive if we present it in terms of the posterior predictive distributions of the responses Y^* , and since these predictions will depend on the value of covariate X_1^* determined at the baseline, we should study the predictive distributions as functions of X_1^* . In each case, producing the predictive distributions can be realized with relative ease by data augmentation within the MCMC sampling scheme, by augmenting the hierarchical models (1), (2), (3) and (4) with the potential outcomes, with the inferential model, and then considering marginals of Y^* . In Figure 2 we show the predictive distributions of Y^* for the three

different regimes given observed decile points of X_1 in the data. The three distributions are often overlapping, and even if some systematic differences between the groups could be discerned, they appear to be quite small.

Model fit can be evaluated by comparing the predictive distributions, drawn given the individual covariate levels, to the actual observed outcome values. Figure 3 shows the medians of the posterior distributions of residuals from the four modules. For modules 3 and 4 these are plotted with respect to both X_1 and X_2 . Only the observations which contribute to the likelihood in the specific module are shown. The residual plots seem to indicate reasonably good model fit.

Producing predictive distributions corresponding to the optimal regime A^{opt} (or $A_1^{\text{opt}}(X_1^*)$) is slightly more involved since then the choice of what particular module, of the four possibilities, is used in the simulation will depend on the current values of X_1^* and X_2^* , the former being chosen in advance for each prediction, and the latter being obtained in the course of the simulation if $A_1^{\text{opt}}(X_1^*) = 0$ (see Appendix). We evaluated the optimal regime for each observed X_1 -value present in the data; in Figure 4 the results are presented as differences between the decision to initiate the treatment after the baseline visit, and the decision not to initiate, given that the later decision after $t = 6$ is optimal. In panel a) the medians and credible intervals obtained from the posterior predictive distributions are presented, as functions of X_1 , for the response itself, while panel b) shows these statistics for the corresponding means, represented in the model by the random functions f . The credible intervals indicate no real differences between the groups based on these data.

When two CD4-measurements have been recorded, finding the optimal decision will be based on both of these. Figure 5 a) shows the predictive distributions of the 6 months change in the CD4-level as a function of the baseline value. A decreasing trend is clearly visible here. Panel b) shows the A_2^{opt} values as a function of the baseline CD4-level and of the 6 months change. There is no clear pattern separating the alternative decisions, indicating again that the differences between the groups in these data are small.

3.2 Considerations due to missing data and censoring

The probability models defined in Section 2 were written for fully observed measurement data and follow-up histories. In reality the data described were subject to various kinds of incomplete observation. Least problematic of these are ordinary missing data in the CD4 laboratory measurements. There are only a small number of these in the data, and due to the full probability model they can be handled easily by using Bayesian data augmentation.

Figure 2: Predictive distributions for three alternative treatment regimes

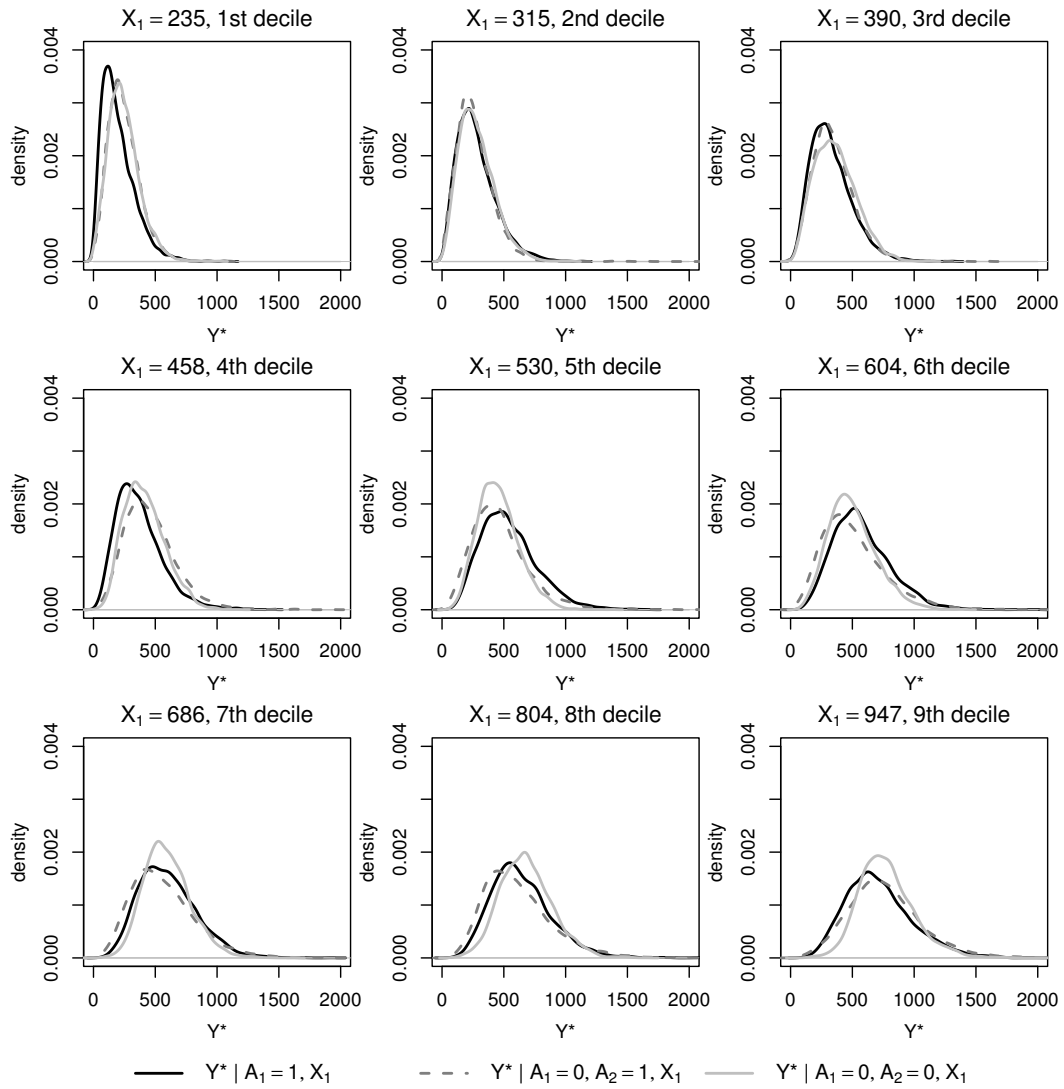


Figure 3: Posterior median residuals

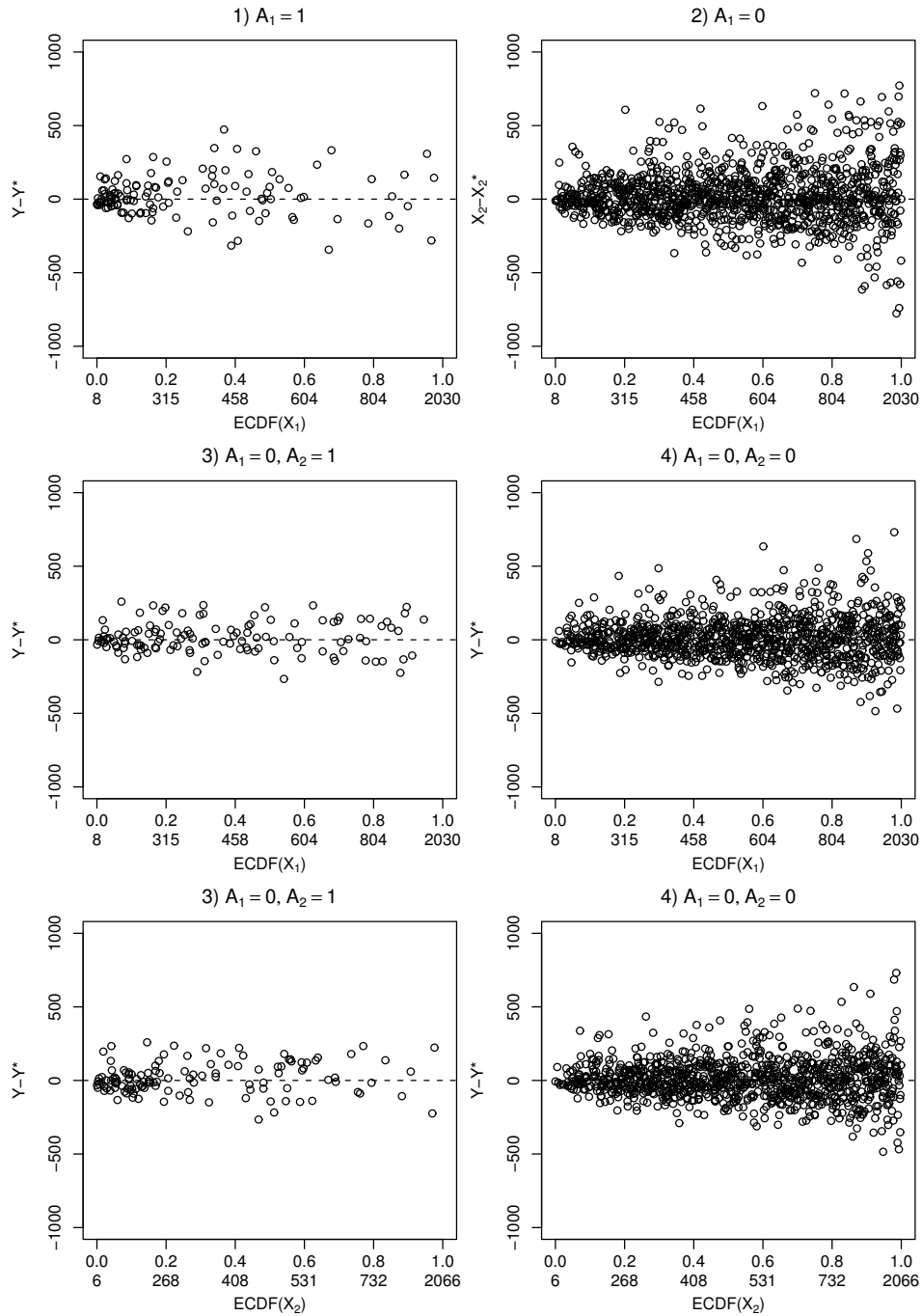


Figure 4: Comparison between two alternative treatment decisions after the baseline visit

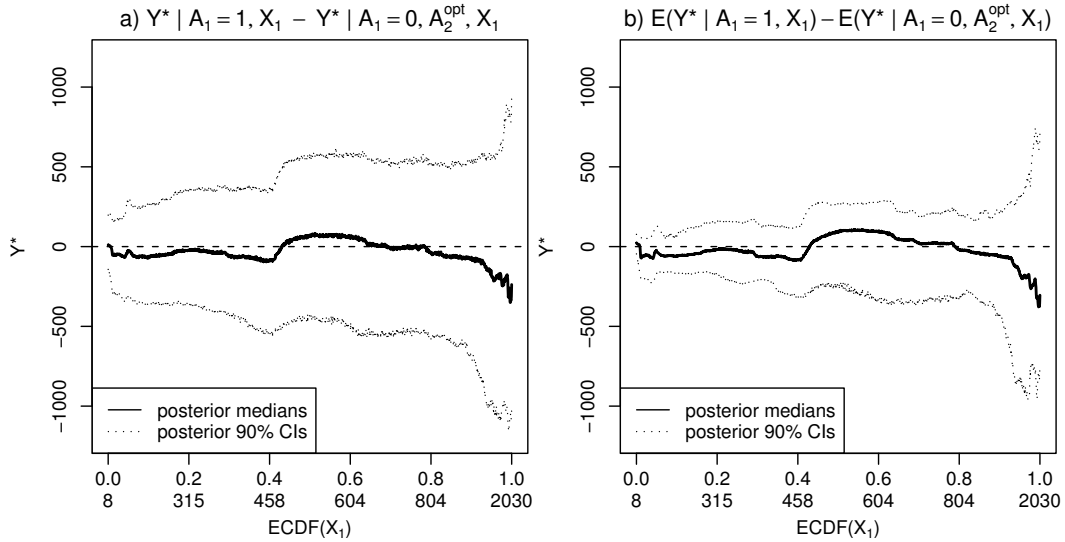
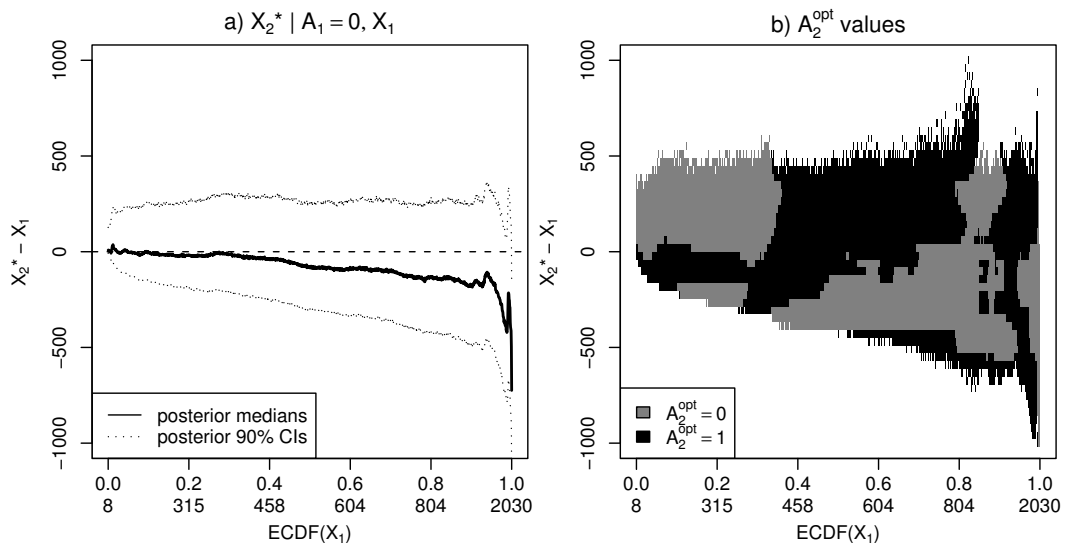


Figure 5: Treatment decision based on two covariate measurements



Incompletely observed treatment histories due to missed follow-up visits are a more problematic issue. Consider first the individuals who dropped out from the follow-up after the first or second visit. (It is possible that the same individuals have reappeared for later MACS visits, but since we consider only follow-up of 12 months we do not utilize this information here.) The data included 200 individuals who were censored from the follow-up after the baseline visit and another 245 who were censored after the intermediate visit. These we omitted from the analysis. Effects of the drop-out for the statistical inference depend on the (typically unknown) censoring mechanism (Little, 1995). For example, here *missing at random* would refer to the assumption that the censoring after the baseline visit would depend only on the observed baseline CD4 measurement X_1 , while the censoring after the intermediate visit would depend only on observed values of (X_1, X_2, A_1) . This assumption is in many ways analogous to the no unmeasured confounders postulate, although the latter may be easier to understand, since it means that the clinician making the treatment decisions and the statistician analyzing the data would be working on the basis of the same observed variables. In contrast, missing at random means assuming that confounding variables really do not exist. If the assumption is valid, censored individuals are similar to individuals who remained in the study with similar covariate histories and the inference is unbiased. A nonignorable censoring mechanism would mean that the drop-out depends on the unobserved response values themselves, or on some latent characteristics related to them. Correcting the resulting bias with likelihood-based inference would require also modeling of the censoring process (including its dependence on the outcome). However, in practice there may be little or no information available for such modeling.

If only the intermediate visit of the three is missed, and since the drug usage items in the MACS questionnaire concern medications taken after the previous follow-up visit, it is possible to infer the corresponding treatment branch and the corresponding likelihood contribution for those who missed the intermediate visit and in the 12 months visit did not report having received AZT treatment (there were 119 such individuals in our dataset). Assuming missing at random, we can include them in the branch $(A_1 = 0, A_2 = 0)$, with missing X_2 measurements to be handled as ordinary missing data, allowing also dependence on the observed values of Y . However, those who in their 12 months visit reported having taken AZT may have initiated the treatment either before or after 6 months. To utilize these individuals in the inference, we would also have to define a probability model for A_1 in order to augment the missing treatment combination. Since there were only 25 such individuals in the data, we did not pursue this here. However, in case of more complicated

designs with many different treatment alternatives possibly only partially observed, and in order to utilize as much of the data as possible, it would be necessary to model also the probabilities of different treatment decisions and their dependence on the covariates.

Rubin (2006) discusses causal inference in situations where follow-up histories are subject to censoring due to death and the response variable reading is thus unobserved for the censored individuals. Here the membership in a ‘principal stratum’ is an unobservable individual characteristic which is a source of nonignorable censoring. For example, the MACS data could include individuals who would have survived with AZT treatment, but who did not receive the treatment and were then censored due to death before 12 months (Rubin’s principal stratum ‘LD’). If treated, an individual belonging to this principal stratum survives, but may end up with a lower CD4-count than those who would survive through the follow-up period irrespective of the treatment decision (principal stratum ‘LL’), thus confounding causal inference. In the MACS study, a more common situation than censoring due to death may well be that an individual simply feels too sick to go to a follow-up visit. This may be even more problematic than censoring due to death since here knowing the dates and the causes of death of the individuals in the study would not help the modeling task. For example, if we knew who has died of AIDS during the 12 month period, we could try to include these individuals in the analysis using zero as the outcome CD4-level. This would not help in cases where the cause of censoring is also unknown.

If the missing at random assumption is unverifiable in practice, how should one then interpret the results obtained from such an observational setting? At least, the results should always be presented conditionally on the baseline measurement of CD4-level (as we have done throughout). As discussed by Rubin (2006), such a covariate may be predictive of both the final and intermediate outcomes, including possible censoring, and thus may shed some light on the unobservable principal strata. For example, in the MACS study we could speculate that those with normal levels of CD4-cells (e.g. > 500) would be primarily individuals who would survive through the follow-up period irrespective of the treatment decision, and thus the results for this group would be less affected by a possible bias due to censoring. Correspondingly, the results may be more uncertain for those whose disease has already progressed to a more advanced stage at the baseline.

4 Discussion

In this paper we have presented a brief illustration of how methods belonging to Bayesian nonparametric inference and dynamic programming can be used in establishing an optimal treatment regime. Although our example was simplified, essentially the same methods can be applied directly to more complicated designs as long as the treatment decisions are discrete and there are sufficiently many observations in each branch of the tree of alternative treatment histories. In practice the modeling would likely involve more covariates than what we considered here. Since completely nonparametric inference of high dimensional problems is usually not feasible, dimension reducing modeling assumptions will be required. For higher dimensional problems we propose Bayesian hierarchical parameterization and “packaging” of covariates (see e.g. Arjas and Liu, 1996), where parametric assumptions are relaxed in the most critical parts of the model. Further work is required with models which could accommodate continuous treatment decisions.

A main challenge is to collect enough covariate information, in order to convince oneself (and others) about the adequacy of the no unmeasured confounders postulate in the considered context, and then model such processes in a way that would lead to plausible predictive distributions. However, this approach alone does not provide a complete solution in problems involving incomplete observation, which cannot be fitted into the standard probability-based missing data framework. As was evident in the present example, the no unmeasured confounders postulate, even if true, will not ensure unbiased inference if the longitudinal design is subject to self-selection due to censoring caused by a missed follow-up visit. At least, attempts should be made to gather information on the causes of such censoring (Little, 1995). Also, the assumed conditions for non-confounded statistical inference should be made explicit in an intuitively understandable way. Here the concept of principal stratification (Frangakis and Rubin, 2002; Rubin, 2006) might be helpful.

In reporting the results from an empirical study, there is often considerable emphasis on statistical tests, or estimates, relating to the parameters of simple (often simplistic) statistical models. Given that sufficient amounts of data are available, nonparametric Bayesian modeling, combined with MCMC methodology, offers an attractive and flexible alternative for statistical inference. This task is to some extent facilitated by the use of constrained (for example, monotone or U-shaped) multivariate nonparametric Bayesian models. Our preference would be to report the results from an empirical study in the form of predictive distributions of the response of interest, with each such prediction corresponding to a specific (sequence of) intervention(s) or

choice(s) of control variables. In studies involving real data the computational challenge can become formidable, and even exceed what is feasible in practice. Nevertheless, we would view the relative simplicity of the present conceptual, entirely probabilistic, framework to be a valuable asset in such an enterprise.

Appendix: Computational details

We outline an algorithm for the sequential inference problem described in Section 3, which requires only a single MCMC sample of size m from the joint distribution of all model parameters and potential outcomes. This differs from the sampling scheme described by Carlin et al. (1998), which utilized nested rounds of simulations in calculating the integrals involved. Let

$$\begin{aligned}
 & E(Y^* \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) & (5) \\
 & = E(E(Y^* \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, f_{01}, \mathcal{D}) \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) \\
 & = \int E(Y^* \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, f_{01}, \mathcal{D}) \\
 & \quad \times p(f_{01} \mid A_1^* = 0, A_2^* = 1, X_1^*, X_2^*, \mathcal{D}) \, df_{01} \\
 & = \int f_{01}(X_1^*, X_2^*) p(f_{01} \mid A_1^* = 0, A_2^* = 1, \mathcal{D}) \, df_{01},
 \end{aligned}$$

and equivalently,

$$\begin{aligned}
 & E(Y^* \mid A_1^* = 0, A_2^* = 0, X_1^*, X_2^*, \mathcal{D}) & (6) \\
 & = \int f_{00}(X_1^*, X_2^*) p(f_{00} \mid A_1^* = 0, A_2^* = 0, \mathcal{D}) \, df_{00}.
 \end{aligned}$$

Note here that the posterior distributions of the random functions f_{01} and f_{00} do not depend on the pair (X_1^*, X_2^*) , since these potential outcomes are not ‘data’ and thus do not affect the estimates of the unknown parameters in the model. With fixed (X_1^*, X_2^*) , the above integrals can be evaluated using Monte Carlo integration, by sampling from the posterior distributions of f_{01} and f_{00} .

To determine $A_1^{\text{opt}}(X_1^*)$, we have to evaluate the expectation

$$\begin{aligned}
 & E(E(Y^* \mid A_1^* = 0, A_2^{\text{opt}}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) \mid A_1^* = 0, X_1^*, \mathcal{D}) \\
 &= \int E(Y^* \mid A_1^* = 0, A_2^{\text{opt}}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) p(X_2^* \mid A_1^* = 0, X_1^*, \mathcal{D}) dX_2^* \\
 &= \int \int E(Y^* \mid A_1^* = 0, A_2^{\text{opt}}(X_1^*, X_2^*), X_1^*, X_2^*, f_{0, A_2^{\text{opt}}}, \mathcal{D}) \\
 &\quad \times p(f_{0, A_2^{\text{opt}}} \mid A_1^* = 0, A_2^{\text{opt}}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) \\
 &\quad \times p(X_2^* \mid A_1^* = 0, X_1^*, \mathcal{D}) df_{0, A_2^{\text{opt}}} dX_2^* \\
 &= \int \int \int \int \int f_{0, A_2^{\text{opt}}}(X_1^*, X_2^*) p(f_{0, A_2^{\text{opt}}} \mid A_1^* = 0, A_2^{\text{opt}}(X_1^*, X_2^*), X_1^*, X_2^*, \mathcal{D}) \\
 &\quad \times p(X_2^* \mid Z_0^*) p(Z_0^* \mid A_1^* = 0, X_1^*, f_0, \alpha_0) \\
 &\quad \times p(f_0, \alpha_0 \mid A_1^* = 0, \mathcal{D}) dZ_0^* df_0 d\alpha_0 df_{0, A_2^{\text{opt}}} dX_2^*.
 \end{aligned} \tag{7}$$

Again, some obvious conditional independencies have been applied in the last expression. The split of the joint posterior distribution into the conditional distributions shown above suggests how these will be used in the MCMC simulation. The algorithm to evaluate (7) now goes as follows. For $k = 1, \dots, m$, given a fixed value of X_1^* , do:

1. Draw parameter values (f_0^k, α_0^k) from the posterior $p(f_0, \alpha_0 \mid A_1^* = 0, \mathcal{D})$.
2. Draw the latent variable Z_0^{*k} from $p(Z_0^* \mid A_1^* = 0, X_1^*, f_0^k, \alpha_0^k)$.
3. Draw the corresponding X_2^{*k} value from $p(X_2^* \mid Z_0^{*k})$.
4. Draw function realizations f_{01}^k and f_{00}^k from $p(f_{01} \mid A_1^* = 0, A_2^* = 1, \mathcal{D})$ and $p(f_{00} \mid A_1^* = 0, A_2^* = 0, \mathcal{D})$, respectively.
5. Optionally, if the predictive distributions of the response variable are of interest, draw a value of Y^k given both $(A_1^* = 0, A_2^* = 1, X_1^*, X_2^{*k}, f_{01}^k)$ and $(A_1^* = 0, A_2^* = 0, X_1^*, X_2^{*k}, f_{00}^k)$, similarly as in steps 2.-3.

Simultaneously, an MCMC sample of parameter values and potential outcomes is also produced from model (1). Now using the obtained samples (f_{01}^k, f_{00}^k) , $k = 1, \dots, m$, of the function realizations, the expectations (5) and (6) are evaluated at each point (X_1^*, X_2^{*k}) , $k = 1, \dots, m$, to determine the values of the indicators $A_2^{\text{opt}}(X_1^*, X_2^{*k})$. Now, with the values of $A_2^{\text{opt}}(X_1^*, X_2^{*k})$ known, expectation (7) can also be evaluated using Monte Carlo integration, using, depending on the value of A_2^{opt} at the point (X_1^*, X_2^{*k}) , the realization f_{01}^k or f_{00}^k in place of $f_{0, A_2^{\text{opt}}}(X_1^*, X_2^{*k})$. Note again that the posterior distributions of

the random functions do not depend on the simulated potential outcomes and thus the evaluation of the function level at any given point (X_1^*, X_2^{*k}) is purely deterministic, given the current realization of the random function. Now (7) can be compared to $E(Y^* | A_1^* = 1, X_1^*, \mathcal{D})$, evaluated using the sample from model (1), to determine the indicator $A_1^{\text{opt}}(X_1^*)$.

The MCMC sampler described here was ran for 25000 iterations after a 15000 round burn-in, saving every 5th state of the chain. Gamma(0.1, 0.1) prior was used for the α parameters. One iteration took 2.7 seconds on a standard desktop computer. To put this into context, each of these iterations involved 100 proposals to update each the four random functions f . The computational time needed in our approach would generally depend roughly linearly on the numbers of knots (likelihood modules) in the tree representing the different treatment alternatives.

References

- Arjas, E., C. Jennison, A. P. Dawid, D. R. Cox, S. Senn, R. G. Cowell, V. Didelez, R. D. Gill, J. B. Kadane, J. M. Robins, and S. A. Murphy (2003). Optimal dynamic treatment regimes - Discussion on the paper by Murphy. *Journal of the Royal Statistical Society, Series B* 65, 355–366.
- Arjas, E. and L. Liu (1996). Non-parametric Bayesian approach to hazard regression: a case-study with a large number of missing covariate values. *Statistics in Medicine* 15, 1757–1770.
- Arjas, E. and J. Parner (2004). Causal reasoning from longitudinal data. *Scandinavian Journal of Statistics* 31, 171–187.
- Carlin, B. P., J. B. Kadane, and A. E. Gelfand (1998). Approaches for optimal sequential decision analysis in clinical trials. *Biometrics* 54, 964–975.
- Frangakis, C. E. and D. B. Rubin (2002). Principal stratification in causal inference. *Biometrics* 58, 21–29.
- Kaslow, R. A., D. G. Ostrow, R. Detels, J. P. Phair, B. F. Polk, and C. R. Rinaldo, Jr (1987). The Multicenter AIDS Cohort study: rationale, organization, and selected characteristics of the participants. *American Journal of Epidemiology* 126, 310–318.
- Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 90, 1112–1121.

- Moodie, E. E. M., T. S. Richardson, and D. A. Stephens (2007). Demystifying optimal dynamic treatment regimes. *Biometrics* 63, 447–455.
- Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society, Series B* 65, 331–355.
- Pearl, J. (2000). *Causality: models, reasoning and inference*. Cambridge: Cambridge University Press.
- Robins, J. M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period - application to control of the healthy worker survivor effect. *Mathematical Modelling* 7, 1393–1512.
- Robins, J. M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics* 23, 2379–2412.
- Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In D. Y. Lin and P. Heagerty (Eds.), *Proceedings of the second Seattle symposium in biostatistics*, New York, pp. 189–326. Springer.
- Rosthøj, S., C. Fullwood, R. Henderson, and S. Stewart (2006). Estimation of optimal dynamic anticoagulation regimes from observational data: a regret-based approach. *Statistics in Medicine* 25, 4197–4215.
- Rubin, D. B. (2006). Causal inference through potential outcomes and principal stratification: application to studies with “censoring” due to death. *Statistical Science* 21, 299–309.
- Saarela, O. and E. Arjas (2009). Postulating monotonicity in nonparametric Bayesian regression. Submitted.
- Wathen, J. K. and P. F. Thall (2008). Bayesian adaptive model selection for optimizing group sequential clinical trials. *Statistics in Medicine* 27, 5586–5604.