Continual Reassessment Method with Bayesian Variable Selection in Phase I Clinical Trails

Zhenyu Zhao *

Abstract

This study presents a variable selection procedure in Bayesian adaptive design of Phase I clinical trials. For the Phase I dose-finding problem, besides the dose level, the dose-limiting toxicity (DLT) may depend on other covariates. The variable selection procedure presented is designed to involve the proper covariates in the model and use the selected model to predict the maximum tolerated dose (MTD). This variable selection procedure can be applied in a large range of Bayesian adaptive designs.

Key Words: Adaptive Design, Bayesian, Model Selection, Variable Selection, MCMC

1. Introduction

In Phase I clinical trials, a major goal is to find a safe dose at the target toxicity level (TTL). The model based designs assume there is a monotonic function describing the relation between dose level and the probability of DLT. Under the Bayesian framework, after each cohort of patients are treated, the parameters in the model will be estimated adaptively based on the posterior distribution. Then the updated functional curve will be used to update the current estimate of MTD and suggest the dose level for the next cohort.

The continual reassessment method (CRM) by O'Quigley et al. (1990) is the first Bayesian model-based phase I design appearing in the literature (see Berry et al. 2010). However, when an improper model is assumed, CRM may leads to overdosing. Due to safety considerations, modifications are added on the CRM to make the process safer, though at the cost of algorithm efficiency. For more details of the modifications, see Korn et al. (1994), Faries (1994), Goodman et al. (1995), Piantadosi et al. (1998), Heyd and Carlin (1999) and Berry et al. (2010).

The feature that the CRM relies on the binary outcomes indicated that this method does not use the information efficiently. This leads to another direction of extension. Cheung & Chappell (2000) introduced the modified CRM method based on the time-to-event (TITE) outcomes. This TITE-CRM not only uses the binary outcome recording whether the patient meets a DLT, but also includes the time to this toxicity event. This method is based on censored time and can accelerate the dose-finding process. For more discussion about the TITE-CRM, see also Thall et al. (2005) and Cheung et al. (2006).

In addition, another extension is to include covariates in the function characterising the relation between the probability of DLT and the dose level. For example, Bailey et al. (2009) described a case study using the model with covariates to study the combination therapy problem. The covariates introduced are indicator variables of the dose level of the second drug. Also see Cheung et al. (2006) for studies on the TITE-CRM with covariates.

Some works have been done on the model selection approach in the Bayesian adaptive design. Wathen & Thall (2008) studied Bayesian model selection method for optimizing group sequential Phase III clinical trials. Yin & Yuan (2009) introduced Bayesian model averaging CRM in Phase I clinical trials, which worked on multiple parallel CRM models, each with a different skeleton of prespecified toxicity probabilities.

*Northwestern University, 2006 Sheridan Rd, Evanston, IL 60208

This study proposed a variable selection method for Bayesian adaptive methods with covariates in Phase I clinical trials. This procedure will be illustrated on the CRM framework, but it is straightforward to be applied on a large range of modified versions of CRM, including TITE-CRM. This article will be arranged as follows: Section 2 first reviews the CRM procedure with covariates and then describes the CRM with Bayesian variable selection (BVS-CRM). Section 3 illustrates the BVS-CRM algorithm in a simulation study. Section 4 discusses limitations and extensions.

2. Methods

2.1 Contimual Reassessment Method with Covariates

Suppose the probability of the DLT is a monotonic function of the dose level (given other parameters are fixed), then after the TTL is defined, there is a unique dose level corresponding to the TTL, which is defined as the MTD. The primary goal of the CRM is to identify the MTD. A parametric model will be developed to characterise the relationship between the probability of the DLT and the dose level. Generally there are $20 - 80$ patients participating in the Phase I clinical trails. The patients will be divided into several cohorts, each with equal or unequal number of patients. After setting the initial dose, the first cohort will be treated at this dose level. After observing the binary outcome which indicates whether the DLT occurs for each patient, the parameters in the model will be updated based on the posterior distribution. Then the dose level for the next cohort will be assigned based on the updated model. After each cohort is treated, the parameter values are updated and the updated model will give the final estimate for the MTD and dose suggestions for the next phase clinical trial.

For modeling the relationship between the probability of the DLT and the dose level, Berry et al. (2010) stated three popular models: hyperbolic tangent model, logistic model, and power model. In practice, the model will be selected based on historical data or the expert opinions. This study will focus on the logistic model, but the idea can be extended to other models as well. Note that the characteristics of the patients and the other drugs taken at the same time may also have effect on the toxicity behavior of the drug investigated. Covariates can be introduced in the model to represent these effects. For example, Bailey et al. (2009) use the covariates to indicate the dose level of the second drug under investiagation.

Suppose there are N_t patients in the tth cohort and T cohorts in total, then the total patient number is $N = \sum_{t=1}^{T} N_t$. Let $i \in \{1, 2, ..., N\}$ denote the index of the patient, and the $i_{\tilde{t}}$th ($i_{\tilde{t}} \in \{1, 2, ..., N_{\tilde{t}}\}$) patient in the $\tilde{t}$th cohort will have the index number of $i = \sum_{t \leq \tilde{t}-1} N_t + i_{\tilde{t}}$. In this way, each patient will have a unique identification number. Let D_0 denote the dose escalation grid set. The algorithm of the CRM with covariates can be summarized as follows:

- Step 1: Charaterize the relationship between dose and toxicity by a parametric model, for example (Bailey et al. 2009):

$$logit(p(d, \alpha, \beta, \gamma, x)) = log(\alpha) + \beta log(\frac{d}{d^*}) + \sum_{j=1}^{J} \gamma_j x_j,$$

  where α and β are two unknown parameters, d^* is a fixed reference dose, $\{x_j\}_{j=1}^{J}$ denote covariates and $\{\gamma_j\}_{j=1}^{J}$ are corresponding covariate coefficients.

- Step 2: Assume a vague or fully non-informative prior for the parameters: $p(\alpha, \beta, \gamma)$

- Step 3: Treat the tth cohort of patients at the current dose level. The initial value of t is 1 and the initial dose for the first cohort of patients is set at a low and safe level.

- Step 4: Observe the toxicity outcome. For the patients in the current cohort, the range of the index number will be $\sum_{t' \leq t-1} N_{t'} + 1 \leq i \leq \sum_{t' \leq t} N_{t'}$. If a DLT is observed for the ith patient, then record $Y_i = 1$, otherwise let $Y_i = 0$. Also record the covariate values $\{X_{ij}\}_{j=1}^{J}$ for the ith patient.

- Step 5: Compute the posterior distribution of (α, β, γ). This calculation will use all data available up to the current step. The total number of patients who have been treated is $\tilde{N} = \sum_{t' \leq t} N_{t'}$ The likelihood after treating n patients is given by

$$L(\alpha, \beta, \gamma; d, X, Y) = \prod_{i=1}^{\tilde{N}} p(d_i, \alpha, \beta, \gamma, X_{i\cdot})^{Y_i} [1 - p(d_i, \alpha, \beta, \gamma, X_{i\cdot})]^{1-Y_i},$$

  where d_i is the dose level for patient i, and $X_{i\cdot} = (X_{i1}, X_{i2}, ..., X_{iJ})$ is the covariate for patient i. Then the posterior density is known up to a normalizing constant:

$$p(\alpha, \beta, \gamma | X, Y) \propto L(\alpha, \beta, \gamma; d, X, Y) p(\alpha, \beta, \gamma).$$

- Step 6: Update the values of (α, β, γ) by the posterior mean $(\bar{\alpha}, \bar{\beta}, \bar{\gamma})$. This step can be done by the MCMC method.

- Step 7: Let p^* denote the desired toxicity level and then select the dose $d^* = argmin_{d \in D} |p(d, \bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{X}) - p^*|$ as the dose level for the next cohort, where $\bar{X}$ denotes the sample average covariate values.

- Step 8: Repeat Steps $3 - 7$ until a sufficiently precise estimate of the parameters is achieved or the maximum sample size is reached.

One drawback of the original CRM approach is that it may lead to overdosing in the first few steps when the model is not estimated accurately. For instance, if no one meets the DLT at the initial level, then the updated model tends to give a high dose level for the next cohort. Certain protections can be made to make this process safer. For example, we can increase the dose by at most one prespecified level at one time. Another modification is to select p^* in Step 7 properly. At the original form, p^* is a fixed value which is the same as the TTL, so each time the model will set the dose closest to the current estimate of MTD for the next cohort. One modification is to start with a low p^* at the beginning, and then increase p^* as the process goes. For example, if the TTL is set as 25%, then we can start with a $p^* = 10\%$ and gradually increase it at each step; in the end, at the final few steps, when we believe the model to be accurate, we set p^* to be 25%. See Berry et al. (2010) for more discussions on modifications to the CRM. However, these safety considerations will not affect the variable selection procedure described below, hence without loss of generality, we will continue to use the original form of the CRM in the following sections.

2.2 Contimual Reassessment Method with Bayesian Variable Selection

There are two main motivations to introduce the variable selection into the CRM. First, this procedure will help us better interprete the data and the model. At the beginning, there might be many covariates proposed. The variable selection procedure will rule out the nuisance covariates and only keep the important covariates in the model. So we can know which covariates have remarkable impact on the toxicity behavior while the other

are neglectable. Second, the variable selection procedure can alleviate the effect of the so-called "curse of dimensionality". When the number of parameters in the model is reduced by the variable selection procedure, a more accurate estimation of the paramters can be expected, which will lead to a more accurate estimate of the MTD.

Regarding the variable selection methods, there are plenty available. See O'Hara and Sillanpää(2009) and Dellaportas et. al. (2002) for an overview of different Bayesian model selection techniques. Here we adopt the reversible jump MCMC method, introduced by Green (1995). The reasons are that (1) it is straigtforward to be applied to different situations; (2) it can be well integrated with the CRM procedure.

To introduce the idea of the reversible jump MCMC, consider the logistic model with covariates described in the section above. Suppose there are J covariates proposed, then there will be 2^J possible models, each including certain covariates while excluding the other. Let I_m denote the model indicator vector: if the jth covariate x_j ($1 \leq j \leq J$) is included in the model m, then the jth component of the indicator vector will be 1, i.e. $I_m^j = 1$; otherwise, set $I_m^j = 0$. Let J_m denote the covariate index set for model m, defined as $J_m = \{j : I_m^j = 1, j = 1, 2, ..., J\}$. At each step, a model and its associated parameters $(m', \alpha', \beta', \{\gamma_j' : j \in J_m\})$ will be proposed according to a specified proposal distribution. Then an acceptence probability is calculated to decide whether or not to accept the move from the current state to the proposed state. The outcome of the reversible jump MCMC algorithm is a posterior sample of different model and the parameters associated. The model with the highest frequency in the sample will be selected. The feature of the reversible jump MCMC is that it enables the Markov chain jump between models with parameter spaces of different dimensions.

The BVS-CRM procedure can be summarized as follows:

- Step 1: Consider the model class (includes 2^J models):

$$logit(p(d, \alpha, \beta, \gamma, x)) = log(\alpha) + \beta log(\frac{d}{d^*}) + \sum_{j \in J_m} \gamma_j x_j$$

- Step 2: Set priors for $(m, \alpha, \beta, \gamma)$, where $\gamma = \{\gamma_j : j \in J_m\}$.

- Step 3: Treat 1 cohort of patients at the level closest to the current estimate of the MTD. The initial value is set at a low dose level. Denote the current dose level as d_t.

- Step 4: For each patient in the current cohort, record the covariate value and outcome: (X_i, Y_i). Also record the dose level used for each patient $D_i = d_t$.

- Step 5: Variable Selection Step. Run the reversible jump MCMC algorithm.

  - Set initial values for the model and parameters. Suppose the current state is $(m, \alpha, \beta, \gamma)$.

  - Propose a model m' and the jump probability from the current model m to model m' is $j(m, m')$. Generate (α', β') from proposal distributions which has density $q(\alpha', \beta'|m', m, \alpha, \beta)$. Generate u from a specified proposal distribution: $q(u|\gamma, m, m')$.

  - Set $(\gamma', u') = g_{m,m'}(\gamma, u)$ where $\gamma' = \{\gamma_j' : j \in J_{m'}\}$ and $g_{m,m'}$ is a invertible function, which has the property $g_{m,m'} = g_{m',m}^{-1}$.

– Accept the proposed move to $(\alpha', \beta', m', \gamma')$ with probability

$$min(1, \frac{L(\alpha', \beta', \gamma'; D, X, Y)p(\alpha', \beta', \gamma'|m')p(m')j(m', m)}{L(\alpha, \beta, \gamma; D, X, Y)p(\alpha, \beta, \gamma|m)p(m)j(m, m')} \times \frac{q(\alpha, \beta|m, m', \alpha', \beta')q(u'|\gamma', m', m)}{q(\alpha', \beta'|m', m, \alpha, \beta)q(u|\gamma, m, m')} \times |\frac{\partial g_{m,m'}(\gamma, u)}{\partial(\gamma, u)}|)$$

– Repeat this procedure until the desirable sample size is obtained or the specified accuracy is reached.

- Step 6: Select the model with the highest frequency in the sample and denote it as m^*. Select the subsample of the MCMC sample which includes m^* as model component. Update the values of (α, β, γ) by the posterior mean of the subsample, denoted as $(\bar{\alpha}, \bar{\beta}, \bar{\gamma})$. To get better accuracy, another MCMC can be conducted to calculate the posterior mean of (α, β, γ) conditional on the model m^*.

- Step 7: Let p^* denote the desired toxicity level and then select the dose $d^* = argmin_d|p(d, \bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{X}) - p^*|$ as the dose level for the next cohort, where $\bar{X}$ denotes the sample average covariate values.

- Step 8: Repeat Steps $3 - 7$ until a sufficiently precise estimate of the parameters is achieved or the maximum sample size is reached.

Green (1995) suggested to use a local proposal distribution for the model and parameter values. That is, at each step, the proposed model either includes one additional covariate or reduces one current covariate compared with the current model. If the proposed model has one more covariate, then a covariate coefficient will be proposed for this covariate according to a proposal distribution; the remaining parameters of the proposed model will take the same value as those in the current model. Otherwise, if the proposed model reduces one covariate compared with the current model, then all the parameters in the proposed model will keep the same values as those in the current model.

3. Simulation Study

In this study, let us consider 10 cohorts with 10 patients per cohort. The dose escalation grids are defined as $D_0 = (0.05, 0.1, 0.2, 0.3, 0.5, 0.7)$. Suppose there are four covariates interested: $x = (x_1, x_2, x_3, x_4)$. Set the reference dose level as $d^* = 0.3$. Assume the true model only includes two covariates (x_1, x_2), which is set as

$$logit(p(d, \alpha_0, \beta_0, \gamma_0, x)) = log(\alpha) + \beta log(\frac{d}{d^*}) + \gamma_{01}x_1 + \gamma_{02}x_2,$$

where $\alpha_0 = 0.3$, $\beta_0 = 1$, and $\gamma_0 = (-1, 1.2)$. Each covariate x is generated according to a Bernoulli distribution with parameter 0.5. Under this setting, if we set the target toxicity level (TTL) as $p_0 = 0.2$, then the maximum tolerated dose (MTD) will be $d_0 = 0.226$ and 0.2 is the closest dose level available in the escalation grids.

The simulation procedure of this BVS-CRM is proceeded as follows:

- Step 1: Consider the model class (includes 16 models):

$$logit(p(d, \alpha, \beta, \gamma, x)) = log(\alpha) + \beta log(\frac{d}{d^*}) + \sum_{j \in J_m} \gamma_j x_j$$

- Step 2: Priors:

$$p(log(\alpha)) \sim N(0,1)$$
$$p(\beta) \sim N(0,1)$$
$$p(\gamma_j | I_m[j] = 1) \sim N(0,1)$$

and

$$p(m) \sim Uniform$$

- Step 3: Treat 1 cohort of patients at the level closest to the current estimate of the MTD. The initial value is taken as the lowest dose level. Denote the current dose level as d_t.

- Step 4: For each patient in the current cohort, generate a covariate vector x and substitude x and d_t into the true model to calculate the probability of DLT: $p(d_t, \alpha_0, \beta_0, \gamma_0, x)$. Then generate a binary outcome y from a Bernoulli distribution with probability $p(d_t, \alpha_0, \beta_0, \gamma_0, x)$. Record the value of (x, y) for each patient.

- Step 5: Variable Selection Step. Run the reversible jump MCMC algorithm.

  - Starting values: $(\alpha, \beta, \gamma) = (0.1, 0.1, 0, 0, 0, 0)$ and $m = (0, 0, 0, 0)$.

  - Proposal distribution:

$$q(\alpha') \sim U[0,5]$$
$$q(\beta') \sim U[0,5]$$

    $j(m, m')$: Simulate j^* from $U\{1, 2, 3, 4\}$ and then let $m'_{j^*} = 1 - m_{j^*}$.

    Let $\gamma'_{-j^*} = \gamma_{-j^*}$, which means set the covariate coefficients in the proposed model the same value as the current model, except the j^*th component.

    If $m'_{j^*} = 1$, which means the proposed model includes one more covariate, then define $q(u') = 1$ and simulate u from $q(u)$ as γ'_{j^*}. Let $N(0,1)$ be the proposal distribution for u, so $q(u)$ is the standard normal density.

    Otherwise set $m'_{j^*} = 0$ and $\gamma'_{j^*} = 0$, then define $q(u) = 1$ and calculate $q(u')$ where $u' = \gamma_{j^*}$ and q is then density of the standard normal.

  - Accept the proposed move to $(\alpha', \beta', m', \gamma')$ with probability

$$min(1, \frac{L(\alpha', \beta', \gamma'; D, X, Y) p(\alpha', \beta', \gamma' | m') p(m') j(m', m)}{L(\alpha, \beta, \gamma; D, X, Y) p(\alpha, \beta, \gamma | m) p(m) j(m, m')}$$
$$\times \frac{q(\alpha, \beta | m, m', \alpha', \beta') q(u' | \gamma', m', m)}{q(\alpha', \beta' | m', m, \alpha, \beta) q(u | \gamma, m, m')} \times |\frac{\partial g_{m,m'}(\gamma, u)}{\partial(\gamma, u)}|),$$

    which can be simplified as

$$min(1, \frac{L(\alpha', \beta', \gamma'; D, X, Y) p(\alpha', \beta', \gamma' | m') q(\alpha) q(\beta) q(u')}{L(\alpha, \beta, \gamma; D, X, Y) p(\alpha, \beta, \gamma | m) q(\alpha') q(\beta') q(u)})$$

  - For each of the first 9 cohorts, repeat this procedure $5,500$ times and drop the initial 500 samples to get $5,000$ valid samples in the end. For the last (10th) cohort, repeat this procedure $33,000$ times and drop the initial $3,000$ samples to get $30,000$ valid samples.

- Step 6: Select the model with the highest frequency in the sample and denote it as m^*. Select the subsample of the MCMC sample which includes m^* as model component. Update the values of (α, β, γ) by the posterior mean of the subsample, denoted as $(\bar{\alpha}, \bar{\beta}, \bar{\gamma})$.
- Step 7: Let $p^* = 0.2$ and then select the dose $d^* = argmin_d|p(d, \bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{X}) - p^*|$ as the dose level for the next cohort, where $\bar{X}$ denotes the sample average covariate values.
- Step 8: Repeat Steps $3 - 7$ until a sufficiently precise estimate of the parameters is achieved or the maximum sample size is reached.

The simulation results for the models with high posterior probabilities are given in Table 1. The standard error is calculated by dividing the 30000 MCMC samples at the last step into 10 batches and computing the standard error of the model posterior probability accross batches. Figure 1 and Figure 2 compare the two optimal models with the model with no covariate and full covariates. Figure 3 shows the dose levels assigned for patients in the cohorts, which are also the estimated MTD in each step. Note that the the final MTD suggested by the optimal model using the full set of data is also 0.2. Parameter estimates of the optimal model are:

$$(\alpha, \beta, \gamma) = (0.520, 1.023, -0.707, 0.820, 0, 0)$$

Table 1: Top 5 models selected

Model	(x_1, x_2)	(x_1, x_2, x_3)	(x_1)	(x_2)	(x_1, x_2, x_4)
Posterior Probability	0.156	0.103	0.100	0.082	0.076
Standard Error	0.020	0.018	0.026	0.026	0.012

As shown in table 1, the optimal model selected is exact the true model. Also, the optimal model suggests the right MTD in the end, and is the closest to the true functional curve in Figure 1 and 2. However, we can also note that the posterior probability of selecting the true model is not very high. There are two possible reasons: (1) The data size is not large, considering there are 16 possible models in total. In addition, the information contained in the binary outcome is limited. (2) The optimal model has no large difference with other models with high posterior probabilities given the data set. As we can see, the optimal model and the second optimal model perform similarly in Figure 1 and 2. The limited information problem also lead to rough estimates of the parameters, and as a result, all estimated models are away from the true models. However, though the estimates are rough, the optimal model (or the true model) still performs the best among all models, and gives the correct MTD in the end. In this study, a flat prior is used for models. The purpose is to examine the procedure without preference on models. In practice, a informative prior can be assigned to state preference.

4. Disccusion and Extensions

Considering the small sample size of the dose-finding procedure and the limited information contained in the binary outcome, the number of covariates proposed can not be too many in the BVS-CRM. To get over this difficulty, more information is required. One possible solution is to apply this covariate selection approach on TITE-CRM. If the time for

the occurence of DLT is well modeled, then this model will offer extra information, which means it is able to include more covariates in the Bayesian variable selection procedure, as well as produce parameter estimates with higher accuracy.

Another extension of this approach is to assign dose to individuals, rather than the cohort. The benefit of this modification is to protect patients who are easy to encounter the DLT from overdosing. This also indicates we could suggest different MTD's for different scenarios or patient groups in the end of the dose-finding study.

REFERENCES

[1] Berry, S.M., Carlin, B.P., Lee, J.J., & Muller, P. (2010). Bayesian adaptive methods for clinical trials. Boca Raton, FL: CRC Press.

[2] Bailey, S., Neuenschwander B., Laird G. & Branson M. (2009) A Bayesian Case Study in oncology phase I combination dose-uinding using logistic regression with covariates. Journal of Biopharmaceutical Statistics, 19:3, 469-485

[3] Cheung, Y.K. and Chappell, R. (2000). Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics, 56, 1177-1182.

[4] Cheung, Y.K., Inoue, L.Y.T., Wathen, J.K., and Thall, P.F. (2006). Continuous Bayesian adaptive randomization based on event times with covariates. Statistics in Medicine, 25, 55-70.

[5] Faries, D. (1994). Practical modications of the continual reassessment method for phase I cancer clinical trials. J. Biopharmaceutical Statistics, 4, 147-164.

[6] Goodman, S.N., Zahurak, M.L., and Piantadosi, S. (1995). Some practical improvements in the continual reassessment method for phase I studies. Statistics in Medicine, 14, 1149-1161.

[7] Green, P. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 4, 711-32

[8] Heyd, J.M. and Carlin, B.P. (1999). Adaptive design improvements in the continual reassessment method for phase I studies. Statistics in Medicine, 18, 1307-1321.

[9] Korn, E.L., Midthune, D., Chen, T.T., Rubinstein, L.V., Christian, M.C., and Simon, R.M. (1994). A comparison of two phase I trial designs. Statistics in Medicine, 13, 1799-1806.

[10] O'Hara, R.B. and Sillanpää, M. J. (2009) A Review of Bayesian Variable Selection Methods: What, How and Which. Bayesian Analysis. 4, 1,85-118

[11] O'Quigley, J., Pepe, M., and Fisher, L. (1990). Continual reassessment method: a practical design for phase I clinical trials in cancer. Biometrics, 46, 33-48.

[12] Piantadosi, S., Fisher, J.D., and Grossman, S. (1998). Practical implementation of a modified continual reassessment method for dose-nding trials. Cancer Chemother. Pharmacol., 41, 429-436.

[13] Thall, P.F., Wooten, L., and Tannir, N. (2005). Monitoring event times in early phase clinical trials: some practical issues. Clinical Trials, 2, 467-478.

[14] Wathen, J.K. and Thall P.F.(2008) Bayesian adaptive model selection for optimizing group sequential clinical trials. Statist. Med. 27:55865604

[15] Guosheng Yin and Ying Yuan (2009): Bayesian Model Averaging Continual Reassessment Method in Phase I Clinical Trials, Journal of the American Statistical Association, 104:487, 954-968

[16] Dellaportas P., Forster, J.J. and Ntzoufras, I. (2002). On Bayesian model and variable selection using MCMC. Statistics and Computing 12:27-36

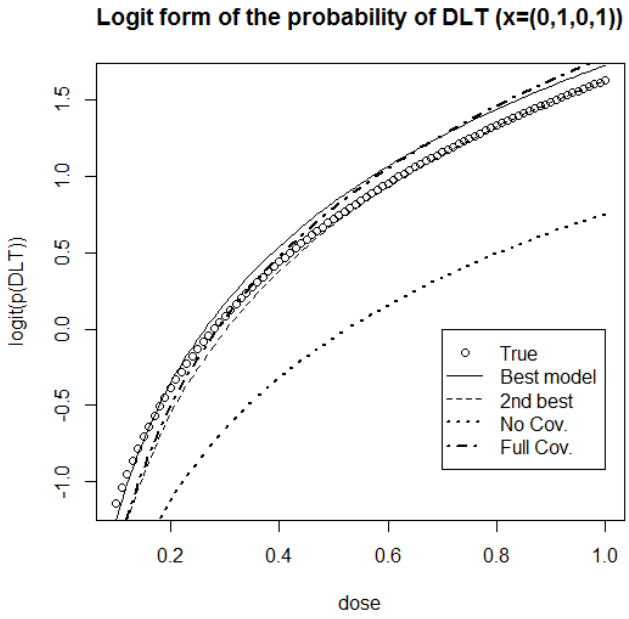

Figure 1: Model comparison given covariate $x = (0, 1, 0, 1)$

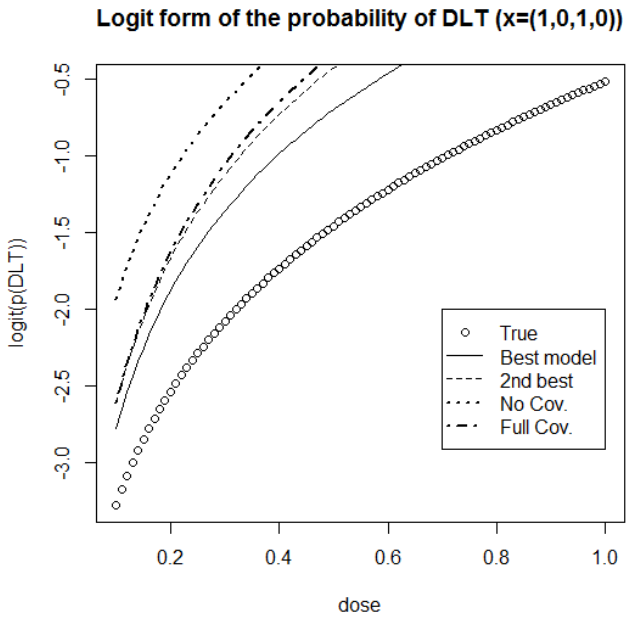

Figure 2: Model comparison given covariate $x = (1, 0, 1, 0)$

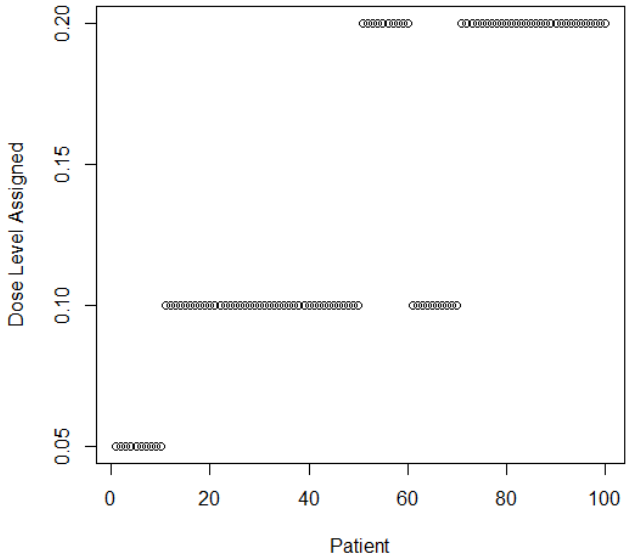

Figure 3: Dose level assigned to patients