

## Adaptive Randomization Using PROC MCMC

Qianyi Huang, McDougall Scientific Ltd.; John Amrhein, McDougall Scientific Ltd.

### ABSTRACT

Based on work by Thall et al. (2012), we implemented a Bayesian method for randomizing patients in a Phase II trial. According to this method, each patient is rated for efficacy and toxicity levels at a specified time point after being randomized to a dose. Evidence is accumulated that identifies which dose(s) of a cancer treatment provide the most desirable profile, per a utility score matrix of efficacy and toxicity combinations rated by experts (0-100). Experts also define the region of “Good” utility scores and criteria of dose inclusion based on toxicity and efficacy performance. Simulation work is done mainly using PROC MCMC in which priors and likelihood function for joint outcomes of efficacy and toxicity are defined to generate posteriors. Resulting joint probabilities for doses that meet the inclusion criteria are used to calculate the mean utility and probability of having “Good” utility scores. Adaptive randomization probabilities are proportional to the probabilities of having “Good” utility scores. A final decision of the optimal dose is made at the end of the trial.

### INTRODUCTION

On February 5<sup>th</sup>, 2010, the FDA (U.S Food and Drug Administration) issued a Guidance for the use of Bayesian Statistics in Medical Device Clinical Trials to provide guidance on statistical aspects of the design and analysis of clinical trials for medical devices that use Bayesian statistical methods. The learning process of the Bayesian framework can be used to modify adaptive characteristics of clinical trials. Accumulated information can also reduce the Sponsor's costs by stopping trials early, especially when futility is assessed. This paper discusses a simulation that powered adaptive allocation (randomization) for a recent oncology trial.

### ADAPTIVE RANDOMIZATION METHOD AND SIMULATION PROCESS

Thall et al (2012) proposed a phase I/II Bayesian adaptive dose randomization method for utility based bivariate ordinal responses. An adaptive allocation (randomization) probability for each dose is determined based on its posterior bivariate distribution as long as it satisfies several criteria. As more patients participate in the trial, more information is accumulated and the chance of a patient receiving a “better” dose improves. By the end of the trial or earlier, either one dose with the largest mean posterior utility score is chosen or no dose is chosen due to futility.

This method combines the purposes of phase I (evaluation of safety or toxicity) and phase II (evaluation of efficacy) by introducing utility scores (0 - 100) to categorized bivariate toxicity and efficacy outcomes. Experts assign utility scores  $U(y_1, y_2)$  (e.g. see Table 1) based on their evaluation of these categories. Toxicity and efficacy levels are ordinal outcomes: 1 means least toxic or least effective;  $m_1$  represents most effective and  $m_2$  stands for most toxic. Intuitively, Level 1 toxicity and level  $m_1$  efficacy will be the ideal outcome which in this example is given a 100 utility score.

Utility Scores	Efficacy level			
Toxicity Level	1	2	...	$m_1$
1	20	30	...	100
2	15	25	...	90
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$m_2$	0	5	...	40

**Table 1. Utility Scores for bivariate toxicity and efficacy outcomes**

Furthermore we can introduce randomness by assigning probabilities to each of the cells making it a bivariate discrete distribution for ordinal outcomes. The interpretation of the probability for each cell can

be the probability of obtaining the (toxicity, efficacy) ordinal bivariate outcome. Note that bivariate distributions vary across different doses.

Thall et al (2012) suggested using a saturated marginal model which is:

$$\pi_{k,y,x} = \Pr(Y_k = y|x, \theta_{k,y,x}) \quad (1)$$

Where  $k = 1, 2$  (1 for efficacy; 2 for toxicity);  $y = 1, 2, \dots, m_k$  represents the level of efficacy or toxicity;  $x = 1, 2, \dots, J$  represents different doses;  $\theta_{k,y,x}$  represents the parameters associated with each efficacy or toxicity level for each dose.

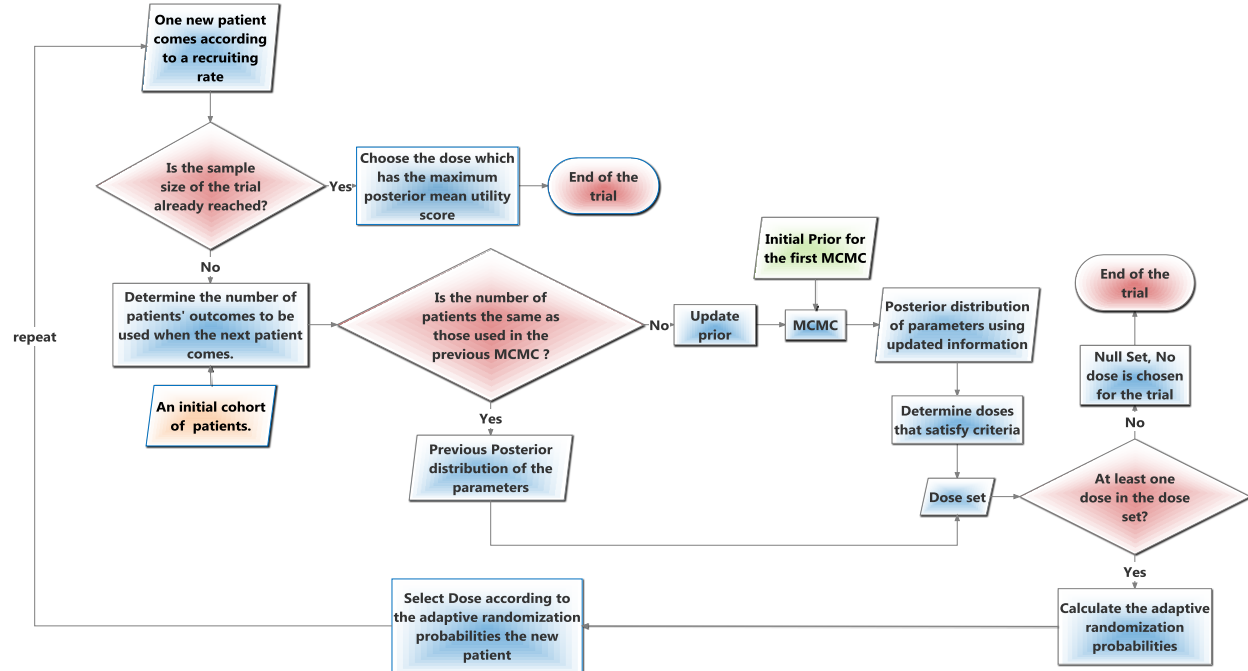
A bivariate Gaussian Copula is then used to combine efficacy and toxicity marginal distributions as a bivariate distribution by incorporating a correlation parameter  $\rho$ . The final likelihood based on the bivariate distribution for  $n$  patients can be denoted as:

$$L(Y|x, \theta) = \prod_{i=1}^n \prod_{y_1=1}^{m_1} \prod_{y_2=1}^{m_2} \{\pi(y_1, y_2|x_{(i)}, \theta)\}^{I(Y_{(i),1}=y_1, Y_{(i),2}=y_2)} \quad (2)$$

Where  $\theta$  represents all parameters used to construct the bivariate distribution and  $I(\cdot)$  is an indicator function with value 1 when conditions in the brackets are satisfied and 0 otherwise. The likelihood function in equation (2) is the product of all bivariate probabilities ( $\pi(y_1, y_2|x_{(i)}, \theta)$ ) associated with outcomes from  $n$  patients.

To reduce the possibility of giving a large randomization probability to a high toxicity and low efficacy dose, there are several criteria that determine which doses are to be included in the next allocation round. Intuitively, the probability of generating a high toxicity level should not be too large and the mean of the utility score for each dose, which is a good measure of efficacy that also accounts for toxicity, should not be too low.

Experts then define a region of good ordinal bivariate outcomes; i.e. the cells for which associated utility scores are greater than the expert-elicited threshold ( $U^*$ ). Adaptive randomization probabilities for all possible doses are then calculated based on their probabilities of generating good outcomes.



**Figure 1. Adaptive Randomization process**

The adaptive randomization process in practice is presented in Figure 1. The process starts with an initial cohort of patients for which doses are assigned with equal probability. Since patient recruitment and randomization is an ongoing process, when the next patient enters the study, not all patients in the cohort

may have completed the trial and/or provided the key efficacy and toxicity ratings required for the randomization process. Therefore, we include a process that determines the number of patient outcomes that can be used for the dose allocation of the new patients. If the number increases compared to that of the previous round, then new completers' outcomes will be used together with those of the old completers as the input dataset for the MCMC procedure. The initial prior distributions of parameters for the first round of MCMC are determined by elicited marginal probabilities by experts. Starting from the second round of MCMC, hyper parameters for the priors are determined using the information from the previous posterior distributions. The dose set is determined by several criteria (which will be discussed in detail later) using the posteriors of marginal and bivariate distributions. If none of the doses satisfy the criteria, then the trial will stop with no best dose chosen. Otherwise, adaptive randomization probabilities are assigned to possible doses for the new patient. If the dose set is not empty when the pre-determined sample size is reached, then the dose with the largest posterior mean utility score is chosen to be the best dose.

We developed macros for each of the independent parts shown in Figure 1; for example, creation of marginal and binomial models, and determination of the initial hyper-parameters and building the filtering criteria to assign randomization probabilities for the next patient. Then, all the independent parts were linked to complete the simulation process.

## IMPLEMENTATION OF THE PROCESS – A SIMULATION IN SAS

### MCMC PROCEDURE

The first task is to build up the probability model that can be used in PROC MCMC. Since the required likelihood function was not among the existing likelihood functions in the MCMC procedure, the bivariate distribution needed to be built from scratch; i.e. first building the marginal distributions and then linking the marginal distributions to bivariate distribution with a correlation parameter.

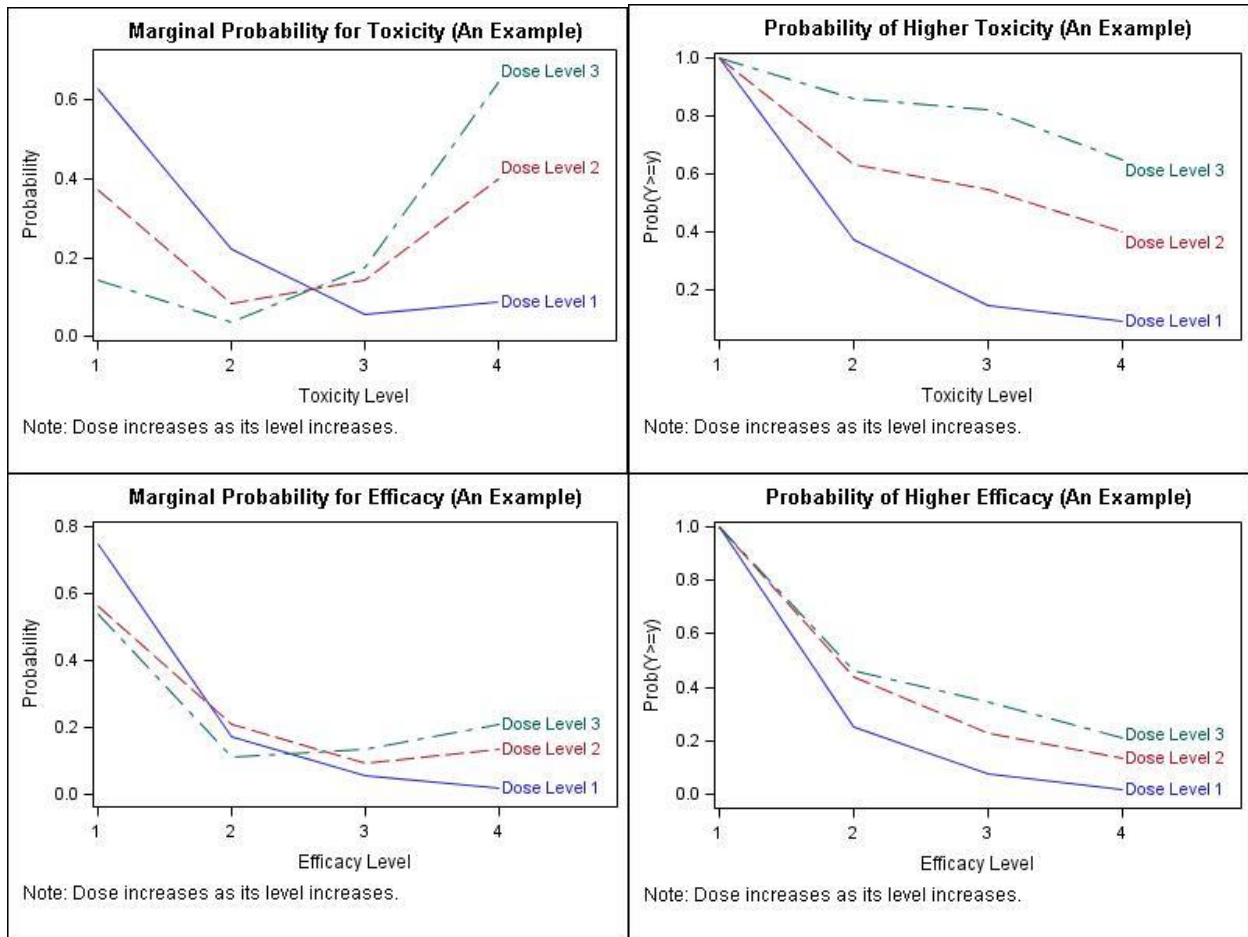
For the simulations, there were three dose levels and four levels for both efficacy and toxicity, i.e.  $J = 3, m_1 = 4, m_2 = 4$ .

$$\left\{ \begin{array}{l} \Pr(Y_k = 1|x, \theta_{k,1,x}) = 1 - \lambda_{k,1,x} \\ \Pr(Y_k = 2|x, \theta_{k,2,x}) = \lambda_{k,1,x} - \lambda_{k,1,x} \cdot \lambda_{k,2,x} \\ \Pr(Y_k = 3|x, \theta_{k,3,x}) = \lambda_{k,1,x} \cdot \lambda_{k,2,x} - \lambda_{k,1,x} \cdot \lambda_{k,2,x} \cdot \lambda_{k,3,x} \\ \Pr(Y_k = 4|x, \theta_{k,4,x}) = \lambda_{k,1,x} \cdot \lambda_{k,2,x} \cdot \lambda_{k,3,x} \end{array} \right\} \quad (3)$$

Where  $\lambda_{k,y,x} = \text{logit}(\theta_{k,y,x}) = \frac{e^{\theta_{k,y,x}}}{1 + e^{\theta_{k,y,x}}}$  and  $\theta_{k,y,x} = \mu_{k,y} + \sum_{z=2}^x \gamma_{k,y,z}$  and  $\gamma_{k,y,z} \geq 0$  for all  $x = 2, 3; y = 1, 2, 3, 4; k = 1, 2$  and for  $x = 1, \theta_{k,y,1} = \mu_{k,y}$ .

According to Thall et al., the parametrization given in equation (3) guarantees a monotone increasing relationship between dose levels (we assume higher level represents higher dose) and the probability of generating higher toxicity and efficacy levels, i.e. higher  $\Pr(Y_k \geq y|x, \theta_{k,y,x})$ . The marginal distributions for toxicity and efficacy can be determined with  $J \times (m_1 + m_2 - 2) = 18$  parameters (total number of  $\mu_{k,y}$  and  $\gamma_{k,y,z}$ ) in this case.

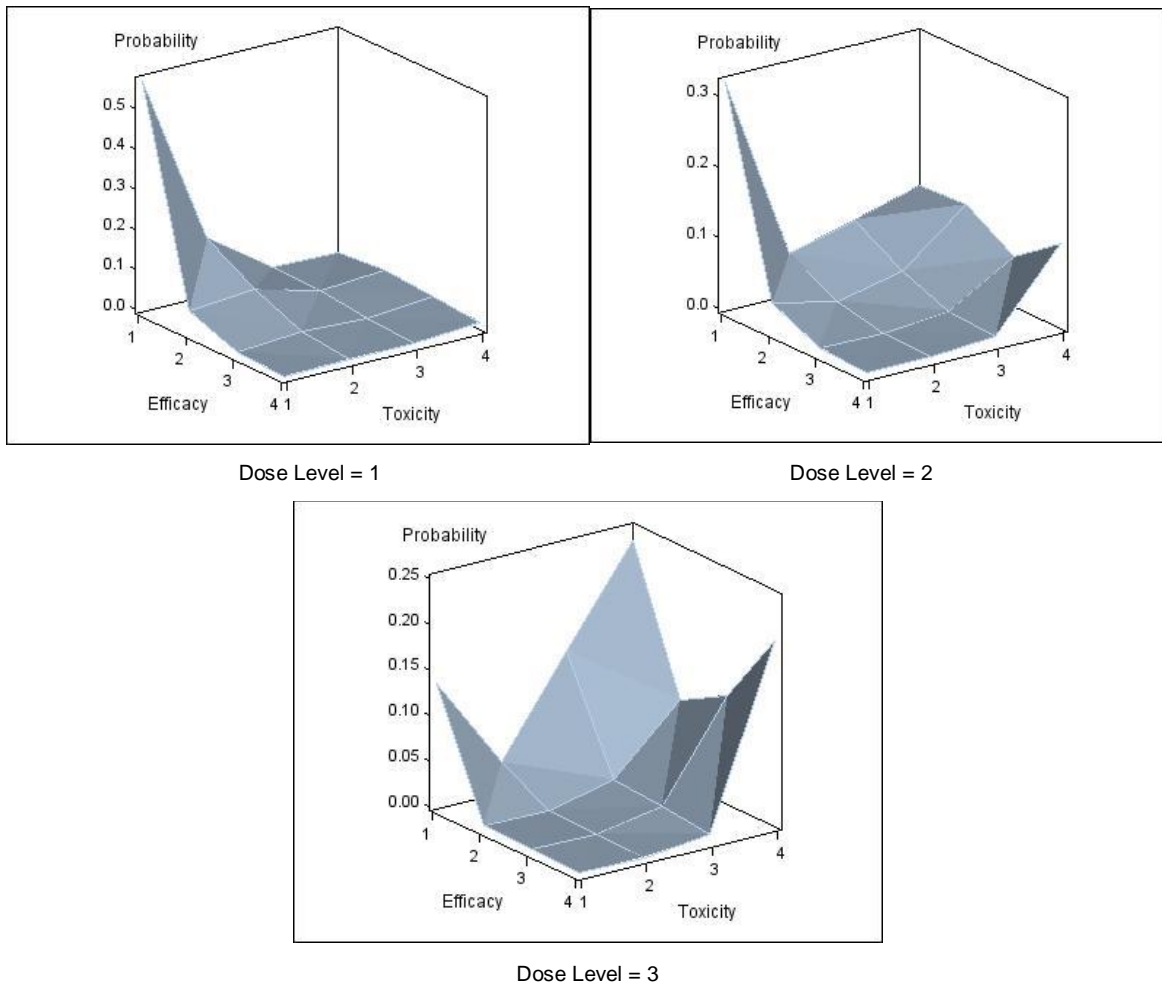
An example in Figure 2 shows the monotone relationship between dose level and  $\Pr(Y_k \geq y|x, \theta_{k,y,x})$  for which we generated  $\mu_{k,y}$  from a standard normal distribution and  $\gamma_{k,y,z}$  from a normal distribution truncated below 0 with mean = 0.1 and standard deviation = 1. In fact, mathematically any values of  $\mu_{k,y}$  and  $\gamma_{k,y,z} \geq 0$  can guarantee the monotone relationship condition. In the MCMC procedure we also use normal distributions for all  $\mu_{k,y}$  and the normal distributions truncated below 0 for all  $\gamma_{k,y,z}$  as priors but with mean and standard deviation determined by the posterior distributions from the previous round of MCMC.



**Figure 2. An example of Marginal models for Toxicity and Efficacy**

If the total number of dose, toxicity and efficacy levels is known, then by the same method of parametrization, marginal probabilities can be determined. This process can be written as a macro function with the number of dose levels, toxicity and efficacy levels as inputs and the marginal probabilities as the outputs i.e. %CSAT (nD = 3, nT = 4, nE = 4), where nD = Total Dose Level; nT = Total Toxicity Level; nE = Total Efficacy Level.

The second step is to create a bivariate distribution from the marginal distribution using a Gaussian Copula by introducing a correlation parameter  $\rho$ . A Copula is a link function defining the dependence structure of multivariate variables and, in this case, is bivariate. A simple standard bivariate normal distribution with correlation  $\rho$  is used. This can also be done using a macro %Cjoint (nD=3, nT=4, nE=4, rho) to create bivariate probabilities at each dose level.



**Figure 3. An example of bivariate distributions for toxicity and efficacy**

Plots in Figure 3 are 3D plots with the X axis as efficacy level, the Y axis as toxicity level and the Z axis as the joint probability. Marginal probabilities are the same as the example given in Figure 2. The correlation coefficient = 0.7. It also shows that as the dose level increases the probability of producing high toxicity and efficacy outcomes increases.

MCMC Procedure Structure in SAS is:

```
PROC MCMC
  DATA = input
  OUTPOST = posterior
  /* MCMC options */
  NBI = 1000
  NMC = 30000
  SEED = **
  THIN = 5
  INF = 1E20;

  /*Initialization of Parameters for Marginal Probability Models*/
  beginnodata;
    %Csat(nD=3, nT=4, nE=4);
  endnodata;

  /*Blocking Parameters in the MCMC Process*/
```

```

parms muT1 0;
parms muT2 0;
parms muT3 0;
parms muE1 0;
parms muE2 0;
parms muE3 0;
parms gammaT1 gammaT2 0;
parms gammaT3 gammaT4 0;
parms gammaT5 gammaT6 0;
parms gammaE1 gammaE2 0;
parms gammaE3 gammaE4 0;
parms gammaE5 gammaE6 0;
parms rho 0;

/* Prior Distributions */
beginprior;
  prior rho ~ uniform(-1,1);

  prior muT1 ~ normal(mean = &muT1, sd = &sdmT1);
  prior muT2 ~ normal(mean = &muT2, sd = &sdmT2);
  prior muT3 ~ normal(mean = &muT3, sd = &sdmT3);

  prior muE1 ~ normal(mean = &muE1, sd = &sdmE1);
  prior muE2 ~ normal(mean = &muE2, sd = &sdmE2);
  prior muE3 ~ normal(mean = &muE3, sd = &sdmE3);

  /*****
  /* Truncated Normal Density */
  *****/
  prior gammaT1 ~ normal(mean = &mgT1, sd = &sdgT1, lower=0);
  prior gammaT2 ~ normal(mean = &mgT2, sd = &sdgT2, lower=0);
  prior gammaT3 ~ normal(mean = &mgT3, sd = &sdgT3, lower=0);
  prior gammaT4 ~ normal(mean = &mgT4, sd = &sdgT4, lower=0);
  prior gammaT5 ~ normal(mean = &mgT5, sd = &sdgT5, lower=0);
  prior gammaT6 ~ normal(mean = &mgT6, sd = &sdgT6, lower=0);

  prior gammaE1 ~ normal(mean = &mgE1, sd = &sdgE1, lower=0);
  prior gammaE2 ~ normal(mean = &mgE2, sd = &sdgE2, lower=0);
  prior gammaE3 ~ normal(mean = &mgE3, sd = &sdgE3, lower=0);
  prior gammaE4 ~ normal(mean = &mgE4, sd = &sdgE4, lower=0);
  prior gammaE5 ~ normal(mean = &mgE5, sd = &sdgE5, lower=0);
  prior gammaE6 ~ normal(mean = &mgE6, sd = &sdgE6, lower=0);
endprior;

/*Create the Bivariate distribution (Likelihood)*/
beginndata;
  %Cjoint(nD=3, nT=4, nE=4, rho=rho);
endndata;

/* Log likelihood */
/* Note: Array probJ[*] is created by %Cjoint*/
llike = log(probJ[dosel, eff, tox]);
model dgeneral(llike);
run;

```

Since the simulation work is automatic, it is not possible to check every diagnostics and whether each posterior distribution is mixing well. For this reason, we set the NMC= option, the number of iterations in

the main simulation loop, to 30000 and the initial burn-in samples to 1000. In addition, to reduce the auto-correlation of the samples, we selected every 5th posterior sample from the posterior distribution to produce the final posterior distribution.

By default, the INF= option, the numerical definition of infinity, is 1E15. If the value is 1E16 or larger, then the procedure will treat the log likelihood as a missing value. There might be situations in which a large variation in the parameters results in very small or large bivariate probabilities, so we set the INF to 1E20 in order to avoid possible missing log likelihood values.

It is not very efficient to set all variables into one PARAMS statement because it requires PROC MCMC to update the multivariate distribution of parameters all at once. Thus, we blocked the parameters by separating parameters by dose to achieve some efficiency.

PRIOR statements specify prior distributions of the parameters. In our case, the correlation parameter  $\rho$  was reasonably assumed to be uniformly distributed between (-1, 1). Mean parameters  $\mu_{k,y}$  were assumed to be normally distributed with mean and variance based on their posterior distribution using the previous information. Parameters  $\gamma_{k,y,z}$  were assumed to be normal distribution truncated below 0 based on previous posterior distribution as well. The distributions were imposed with non-negative constraints to guarantee the monotonicity of the probability of getting higher levels of efficacy/toxicity when dose level increases.

It is possible to make variations on the prior distributions of the parameters as long as it can satisfy the constraints of the marginal distributions. Further, more non-standard log-likelihoods can be programmed using GENERAL/DGENERAL, and here we chose DGENERAL. The likelihood is given as `probJ[dose1, eff, tox]`, which is the probability of a efficacy and toxicity outcome given a dose level. The argument of the GENERAL (Continuous) / DGENERAL (Discrete) functions contains the expression of the log likelihood for the data as long as the expression is programmable using SAS functions and expressions.

There is a way of shortening the running time of MCMC by using BEGINNODATA; ENDNODATA; BEGINPRIOR; ENDPRIOR which reduced the observation-level computations. Statements inside can be a transformation of the parameters for which the same computation repeats for every observation.

## HYPER-PARAMETERS

Since Bayesian analysis always takes prior information into account, choosing proper priors is important. When not enough information is given, a non-informative prior for the parameters is usually a better choice. Morita et al. (2008) proposed a formal way of determining the non-informative prior distribution using Effective Sample Size (ESS). In our simulation we implemented a simplified version described in Thall's paper. Initial Hyper-parameter determination is a one-time process based on experts' elicited marginal distributions. Starting from the second run of MCMC, hyper parameters are determined by the sample mean and standard deviation of the previous posterior distributions. In the sample SAS code, these are denoted as macro variables, for example, "&muT1" stands for previous posterior mean and "&sdmT1" stands for previous posterior standard deviation of  $\mu_{2,1}$ .

## DETERMINE THE ADAPTIVE RANDOMIZATION PROBABILITY

In order to determine the adaptive randomization probability for each dose we need to consider both toxicity and efficacy. Thall et al. (2012) defined three criteria to exclude doses that are too toxic or not effective.

1. The first one is the posterior probability of the possibility of having higher toxicity levels greater than an expert elicited probability should not exceed a threshold value. In this criterion, there are three pre-specified threshold values: the first one is from which levels ( $y^*$ ) upward are high toxicity levels and the second one is the expert elicited high toxicity probability ( $\pi^*$ ) and the cut-off probability ( $p_T^*$ ). Mathematically, this criterion can be written as:

$$\Pr(\Pr(Y_2 \geq y^* | x, \theta) \geq \pi^* | I_n) \leq p_T^* \quad (4)$$

where  $I_n$  represents the information based on  $n$  patients' outcomes.

2. Second is that the dose should have the chance of reaching the highest mean utility score. In this criterion, acceptable probability ( $p_E^*$ ) of reaching the highest utility score is also pre-determined by experts.

$$\Pr(u(x, \theta) = \max_{z \in X} \{u(z, \theta)\} | I_n) \geq p_E^* \quad (5)$$

where  $u(x, \theta) | I_n = \sum_{y_1=1}^{m_1} \sum_{y_2=1}^{m_2} U(y_1, y_2) \pi(y_1, y_2 | x, \theta, I_n)$ .

3. The third criterion is the dose should generate a posterior mean utility score  $\varphi$  (which is the average of the utility using the posterior bivariate distribution) not too far from the largest posterior mean utility score  $\varphi_{max}$ . In this criterion, acceptable distance  $\delta_n$  from the largest mean posterior utility score needs to be pre-determined and is decreasing as  $n$  increases.

$$|\varphi(x, I_n) - \varphi_{max}(x, I_n)| \leq \delta_n \quad (6)$$

where  $\varphi(x, I_n) = E_\theta[u(x, \theta) | I_n]$ .

The MCMC procedure discussed previously produces a sample of posterior distributions of the parameters named as `posterior`. The probability of generating high toxicity levels which is  $\Pr(Y_2 \geq \pi^* | x, I_n)$  and mean utility scores  $u(x, \theta) | I_n$  and posterior mean utility scores  $\varphi(x, I_n)$  can be determined using the samples of the posterior distributions. Probabilities of generating good utility outcomes for each selected dose are standardized and then are used as the adaptive randomization probabilities. Note that if a dose is dropped then the adaptive randomization probability will be zero.

## AN EXAMPLE

In this example, the real marginal distributions (see Table 2) and bivariate distribution was constructed with a correlation coefficient = 0.2.

True Marginal Probabilities	Efficacy level				Toxicity Level			
Dose Level	1	2	3	4	1	2	3	4
1	0.1	0.6	0.1	0.2	0.55	0.3	0.1	0.05
2	0.1	0.55	0.1	0.25	0.45	0.4	0.1	0.05
3	0.1	0.47	0.15	0.28	0.25	0.5	0.15	0.1

**Table 2. True Marginal Probabilities**

The elicited Marginal distributions to determine the initial hyper parameters are as follows:

Elicited Marginal Probabilities	Efficacy level				Toxicity Level			
Dose Level	1	2	3	4	1	2	3	4
1	0.4	0.3	0.2	0.1	0.4	0.3	0.2	0.1
2	0.2	0.25	0.25	0.3	0.2	0.25	0.25	0.3
3	0.1	0.2	0.3	0.4	0.1	0.2	0.3	0.4

**Table 3. Elicited Marginal Distributions**

We used the elicited utility scores associated with toxicity and efficacy outcomes in Thall's paper which can be seen in the following table, the cells that are highlighted in blue are defined as good outcomes (i.e.  $U^* = 25$ ).



Elicited Utility Scores	Efficacy level			
Toxicity Level	1	2	3	4
1	50	85	92	100
2	25	50	60	75
3	10	15	20	25
4	0	5	7	10

**Table 4. Elicited Utility Scores**

The process started from 3 patients treated with the three dose levels and we assumed patients join the study on average 30 days and use 10 days to finish all the evaluations. Adaptive randomization started from the 4<sup>th</sup> patient, the total patients in the trial was planned to be 20.  $U^* = 25$ ;  $y^* = 4$ ;  $\pi^* = 0.1$ ;  $p_T^* = 0.9$ ;  $p_E^* = 0.05$ . From 4<sup>th</sup> patient to 10<sup>th</sup> patient we set  $\delta_n = 20$ ; for 11<sup>th</sup> patient to 20<sup>th</sup> patient, we set  $\delta_n = 15$ .

Dose	True Mean Utility Score	Average Mean Utility Score (Standard Deviation) evaluated at the end of the trial	Percentage of selection (%)
1	64.7	64.4 (6.23)	86.7
2	62.5	54.0 (11.63)	13.3
3	52.8	33.9 (10.04)	0

**Table 5. Simulation Results**

We repeated the process 15 times and the average mean utility score at the end of the study presented in Table 5 is consistent with the true mean utility scores. 86.7% of the time, dose level 1 was chosen to be the best dose and 13.3% of the time dose level 2 was chosen at the end of the trial. Mean utility scores for dose level 2 and dose level 3 were lower than their true mean utility scores and they had large variations which are due to the small amount of patients being randomized to these two dose levels.

## CONCLUSION AND FURTHER DISCUSSION

The methods described above provide insight into the techniques required to implement a Bayesian based adaptive randomization procedure. It translates novel methods into an application which can be used to address a real world problem in the pharmaceutical industry using SAS. The simulation technique can be modified according to different marginal and bivariate models and different dose selection criteria with the flexibility of the MCMC procedure and macro functionality of SAS.

The adaptive randomization process needs experts' information and even a little change in the elicited threshold values can change the dose selection process. This requires coordination among clinical, medical and statistical experts at the trial design stage.

Since the MCMC procedure is embedded in the adaptive randomization simulation, there are several points to be noted:

1. Diagnostics were not checked and thus the posterior distribution might not be stable after the pre-specified burn-in samples. So in the simulation, it is better to set the number of iterations large enough and have relatively larger burn-in samples and thinning rate. In the actual setting, convergence of the distributions parameters and diagnostics should always be carefully checked.
2. It is hard to tune the option parameters for each MCMC round in the simulation practice. Again, in the actual trial, the tuning process should be included.
3. As patient numbers increase, MCMC will require more running time. Also, it is possible that large variation in parameters from the proposal distribution (which is multivariate normal distribution by default in SAS) can result in very small bivariate probabilities (close to 0) thus generating imprecise estimation and affecting the corresponding posterior distributions. A more efficient sampling method is needed. In Thall et al (2012) paper, Gibbs sampling was used. This can be a further research

question to implement Gibbs sampling or more an efficient sampling method in the adaptive randomization process.

## REFERENCES

- [1] Thall, Peter F. and Nguyen, Hoang Q. 2012. "Adaptive Randomization to improve utility-based dose-finding with bivariate ordinal outcomes" *Journal of Biopharmaceutical Statistics*, 22:785-801.
- [2] Morita, Satoshi; Thall, Peter F.; Müller, Peter. "Prior Effective Sample Size in Conditionally Independent Hierarchical Models" *Bayesian Anal.* 7 (2012), no. 3, 591--614.
- [3] Berry, Scott M., et al. *Bayesian adaptive methods for clinical trials*. CRC press, 2010.
- [4] Carpenter, Art. 2004. *Carpenter's Complete Guide to the SAS® Macro Language. Second Edition*. Cary, NC: SAS Institute Inc.

## ACKNOWLEDGMENTS

We would like to thank Statistics Team in McDougall Scientific Ltd. for their support and advices.

## RECOMMENDED READING

- SAS/STAT® 9.2 User's Guide, *The MCMC Procedure*
- *Guidance for Industry and FDA Staff; Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials; February 5, 2010: Center for Devices and Radiological Health (CDRH) Division of Biostatistics Office of Surveillance and Biometrics and Center for Biologics Evaluation and Research (CBER)*

## APPENDIX

```
/******  
/* Macro for Creating Saturated Marginal Model */  
/******  
  
%macro cSat(nD=, /* number of Dose levels */  
            nT=, /* number of Toxicity levels */  
            nE= /* number of Efficacy levels */);  
  
    /*mu*/  
    array muT[%eval(&nT-1)];  
    array muE[%eval(&nE-1)];  
  
    /*gamma*/  
    array gammaT[%eval(&nD-1), %eval(&nT-1)];  
    array gammaE[%eval(&nD-1), %eval(&nE-1)];  
  
    /*Cumulative summation*/  
    array sumgT[%eval(&nD), %eval(&nT-1)];  
    array sumgE[%eval(&nD), %eval(&nE-1)];  
  
    /*Cumulative product*/  
    array prodT[%eval(&nD), %eval(&nT-1)];  
    array prodE[%eval(&nD), %eval(&nE-1)];  
  
    /*logistic(lambda) = theta*/
```

```

array lamT[&nD, %eval(&nT-1)];
array lamE[&nD, %eval(&nE-1)];

/*pi*/
array piT[&nD, &nT];
array piE[&nD, &nE];

/*pibar*/
array pibarT[&nD, &nT];
array pibarE[&nD, &nE];

array cumuT[&nD, &nT];
array cumuE[&nD, &nE];

array Copula[&nD, &nE, &nT];
/* Joint probability */
array probJ[&nD, &nE, &nT];

/*This is doing cumulative sum;*/
do y=1 to %eval(&nT-1);
    sumgT[1,y] = 0;
    do d = 2 to &nD;
        sumgT[d,y] = sumgT[d-1,y] + gammaT[d-1,y];
    end;
end;

do y = 1 to %eval(&nE-1);
    sumgE[1,y] = 0;
    do d = 2 to &nD;
        sumgE[d,y] = sumgE[d-1,y] + gammaE[d-1,y];
    end;
end;

/*Define lambda*/
do d = 1 to &nD;
    do y = 1 to %eval(&nT-1);
        lamT[d, y] = logistic( muT[y] + sumgT[d,y] );
    end;
end;

do d = 1 to &nD;
    do y = 1 to %eval(&nE-1);
        lamE[d, y] = logistic( muE[y] + sumgE[d,y] );
    end;
end;

/*Define cumulative product*/;
do d = 1 to &nD;
    prodT[d,1] = lamT[d,1];
    prodE[d,1] = lamE[d,1];
    do y = 2 to %eval(&nT-1);
        prodT[d,y] = prodT[d,y-1] * lamT[d,y];
    end;
    do y = 2 to %eval(&nE-1);
        prodE[d,y] = prodE[d,y-1] * lamE[d,y];
    end;
end;

```

```

end;

/*Define pi*/
do d = 1 to &nD;
    piT[d,1] = 1-lamT[d,1];
    piE[d,1] = 1-lamE[d,1];
    piT[d,&nT] = prodT[d,%eval(&nT-1)];
    piE[d,&nE] = prodE[d,%eval(&nE-1)];

    do y = 2 to %eval(&nT-1);
        piT[d, y] = (1-lamT[d, y]) * prodT[d, y-1];
    end;

    do y = 2 to %eval(&nE-1);
        piE[d, y] = (1-lamE[d, y]) * prodE[d, y-1];
    end;
end;

/*Define pibar*/
do d = 1 to &nD;
    pibarT[d,1] = 0.9999999999;
    pibarE[d,1] = 0.9999999999;

    do y = 2 to &nT;
        pibarT[d, y] = pibarT[d, y-1] -piT[d, y-1];
        /* avoid the possibility of error message from probbnrm function */
        if abs(pibarT[d, y]) < 10e-15 then pibarT[d, y] = 10e-15;
    end;

    do y = 2 to &nE;
        pibarE[d, y] = pibarE[d, y-1] -piE[d, y-1];
        if abs(pibarE[d, y]) < 10e-15 then pibarE[d, y] = 10e-15;
    end;

end;

/*Define cumulative probabilities*/
do d = 1 to &nD;
    cumuT[d,&nT] = 0.9999999999;
    cumuE[d,&nE] = 0.9999999999;

    do y = 1 to %eval(&nT-1);
        cumuT[d,y] = 1-pibarT[d,y+1];
        if abs(cumuT[d,y]) < 10e-15 then cumuT[d,y] = 10e-15;
    end;

    do y = 1 to %eval(&nE-1);
        cumuE[d,y] = 1-pibarE[d,y+1];
        if abs(cumuE[d,y]) < 10e-15 then cumuE[d,y] = 10e-15;
    end;

end;

%mend;

/*****/

```

```

/* Macro for Creating Bivariate Distribution */
/*****/
%macro Cjoint(nD=, /* number of Dose levels */
              nT=, /* number of Toxicity levels */
              nE=, /* number of Efficacy levels */
              rho= /* Correlation coefficient b/w Toxicity and Efficacy */);

do d = 1 to &nD;
  do e=1 to &nE;
    do t=1 to &nT;
      Copula[d,e,t]=probbnrm(probit(cumuT[d,t]),probit(cumuE[d,e]),&rho);
    end;
  end;

  probJ[d,1,1] = Copula[d,1,1];

  do i = 2 to &nT;
    probJ[d, 1, i] = Copula[d,1,i] - Copula[d,1,i-1];
  end;

  do i = 2 to &nE;
    probJ[d, i, 1] = Copula[d,i,1] - Copula[d,i-1,1];
  end;

  do e = 2 to &nE;
    do t = 2 to &nT;
      probJ[d,e,t] = Copula[d,e,t] + Copula[d,e-1, t-1] - Copula[d,e, t-1] -
Copula[d,e-1, t];
    end;
  end;

end;
%mend;

```

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Qianyi Huang  
 McDougall Scientific Ltd.  
[qhuang@mcdougallscientific.com](mailto:qhuang@mcdougallscientific.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.