A Simulation Study of a Bayesian Hierarchical Changepoint Model with Covariates

Wonsuk Yoo[(1)], **Elizabeth H. Slate**[(2)]

[(1)] Department of Mathematical Sciences
New Jersey Institute of Technology
Newark, NJ 07102

[(2)] Department of Biostatistics, Bioinformatics and Epidemiology
Medical University of South Carolina
Charleston, SC 29425

CAMS Report 0405-19, Spring 2005

Center for Applied Mathematics and Statistics

NJIT

A Simulation Study of a Bayesian Hierarchical Changepoint Model with Covariates

Wonsuk Yoo[1] and Elizabeth H. Slate[2]

1. Department of Mathematical Sciences, New Jersey Institute of Technology

2. Department of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina *

Abstract

This paper presents a simulation study to investigate the behavior of estimates of covariate effects in a Bayesian hierarchical changepoint model. The model is a segmented linear regression model with one changepoint within a fully Bayesian framework. We introduce covariate effects into the model in three ways: as an affect on the interaction term, an affect on the slope after the changepoint, or as an affect on the timing of the changepoint. We estimate all parameters using Markov chain Monte Carlo and compute bias, relative bias and average mean square error of the posterior means of the parameters capturing the covariate effcts. We also investigate the effects of model misspecification on the characteristics of the effect of using an inappropriate fitted model.

1 Introduction

Prostate Cancer (PCa) is the most common type of cancer (other than skin cancer), and the second leading cause of cancer death in US men. The American Cancer Society (ACS) estimates that during 2004 about 230,110 new cases of prostate cancer will be diagnosed and about 29900 men will die of prostate cancer in the United States. Prostate cancer accounts for about 11% of male cancer-related deaths. Thus, 1 man in 32 will die of this disease. Nevertheless, early detection of this cancer, while it is still confined to the prostate, increases the chances of full recovery and hence can reduce mortality attributable to this disease. Biomarkers, biological quantities that can be indicative of underlying disease status, have potential to improve early detection. Currently, prostate specific antigen (PSA) is the most useful biomarker for prostate cancer. PSA is a protein produced by the cells of the prostate gland and is measured through a simple blood test. When the prostate gland enlarges, PSA levels in the blood tend to rise. An elevated PSA level may lead more possibility of presence of prostate cancer. A longitudinal series of PSA measurements taken periodically from a subject may lead to the development of diagnostic criteria with much higher sensitivities and specificities. A few papers have modelled longitudinal PSA readings to better understand the natural history of the PSA trajectories in those who do and do not develop prostate cancer and shown that PSA tends to increase only very slowly among healthy men and increases much more rapidly among men with prostate cancer (Carter *et al.* 1992, Slate *et al.* 1999, Whittemore *et al.* 1995). This suggests that by statistically modeling these trends, longitudinal series of PSA readings taken over time for individuals may be used to improve the value of PSA for prostate cancer detection. With these views, this paper aims to provide a most appropriate model which can figure out the relationship between prostate cancer and longitudinal series of PSA measurements. Thus we want to have early detection of prostate cancer using an efficient biomarker of prostate cancer, PSA.

*This work was funded by grants from DEPT OF DEFENSE - Federal DAMD 17-02-1-0138.

We consider a fully Bayesian hierarchical changepoint model similar to that proposed by Slate and Clark (1999), which is a Segmented Random Effects Model with one change-point representing an onset time of a disease (initiation of the prostate cancer) within full Bayesian framework. There have been several researches that interested in relationship between PSA level and its diagnostic role, "what probability is there that a subject can have the prostate cancer". Similarly, our interest focuses on estimating the subject-specific parameters such as subjects' changepoint. One of typical practical problems in statistical application may be related with consideration of any related covariates into the interested model. It can be called generalization of the model. A risk factor is anything that increases the chance of developing a disease such as cancer. It can be any kind of external or internal conditions including food habit, life-style, or health condition which can lead higher chance to develop a specific disease. American Cancer Society (ACS) considers several risk factors for prostate cancer such as age, race, diet, family history, physical inactivity and others. There have been many researches to figure out risk factors that increase the risk of developing prostate cancer mainly with smoking habit, alcohol consumption and body mass index (Hiatt *et al.* (1994), Platz *et al.* (1999), Demark-Wahnefried *et al.* (2000), Josept *et al.* (2002)). It would be substantially important to public health to have any evidence that even a moderate level of smoking or alcohol consumption can increase relative risk(RR) since there might be high incidence between prostate cancer and potential risk factors such as cigarette smoking, use of alcohol, body mass index(BMI). The problem is related with how we incorporate risk factors into our model since there has been no previous research on addition of covariate into a fully Bayesian hierarchical changepoint model. Thus, it is reasonable to consider three ways to incorporate risk factors into our model as similar as that proposed by Cronin (1994). A covariate might affect PSA readings at the overall level, through the age at which cancer initiates (changepoint), and by the growth rate following the changepoint. For example, assume that smoking habit might be a risk factor of prostate cancer. it is believed to increase PSA level in men with and without prostate cancer while smoking can be believed to affect PSA when a changepoint occurs. In this study, we investigate the covariate effects in these three ways.

2 A Bayesian Hierarchical Changepoint Model with Covariates

The time-dependent measurements which can be one of typical type of Change-point problem arises in various statistical application fields such as medical, pharmaceutical, biological and even economic and social researches. For example, if a researcher want to investigate the effectiveness of new drug in pharmaceutical research, a change-point model can be one of promising candidates in the experimental design for an interest model since she can get information of the onsets of effectiveness through comparison of subjects' changepoints. A Bayesian model can be good alternative to figure out this problem since the method permits the "borrowing of strength" which can bring population information into the model and incorporates expert knowledge. The development of computer-intensive method can answer questions in complex hierarchical, not-closed formed situations where explicit solutions are not available or can not be possible. With such development of computing technology, Bayesian changepoint models have been effective tools for researches that focuses on estimating the subject-specific parameters such as subjects' changepoint. Even though the model is somewhat complicated with hierarchical structure for prior information, we can fit the changepoint easily using computer-intensive techniques such as Markov chain Monte Carlo (MCMC) techniques, an iterative Monte Carlo algorithm.

Since the Gibbs sampler, a special type of Markov chain Monte Carlo method, was introduced by Geman and Geman (1984) and had been developed by Gelfand and Smith (1990) and Gelfand *et al.* (1990), problems related with complicated and not closed-formed integration such as Bayesian

model, hierarchical model, and changepoint model can be easily solved for the Gibbs algorithm to be able to draw samples from the the marginal posterior distribution without exact calculation. Carlin *et al.* (1992) gave a general approach to hierarchical Bayes changepoint models, including the use of the Gibbs sampler, an iterative Monte Carlo method, to solve for all desired marginal posterior distributions of model parameters. They showed how the Markov chain changepoint problem from a Bayesian viewpoint can be examined. Cronin *et al.* (1994) used the Gibbs sampler to detect changepoints in a fully Bayesian hierarchical model. They were interested in sequentially estimating the marginal posterior distributions for the changepoint and the slope after the chagepoint and proposed the usage of posterior distributions of the changepoint as one diagnostic rule for detecting prostate disease status. Kiuchi *et al.* (1995) used hierarchical Bayes change point model to investigate disease progression of HIV infection through changepoint in the series of T4 counts data. Slate *et al.* (2000) also applied a segmented linear regression model in fully Bayesian framework. The Gibbs sampler was used to investigate the posterior probabilities of changepoints for prospective prostate cancer detection. They considered the log-transformed PSA level for dependent variables but response measurements can be any biological markers which can be indicative of underlying disease status and have potential to improve early detection.

2.1 No-covariate Model

First of all, we explain just the no-covariate Bayesian hierarchical changepoint model and provide three methods to consider covariates into the Bayesian model, so called the generalization of Bayesian hierarchical changepoint model. This Bayesian model consists of three-stage with prior and hyper-prior distributions. Prior and hyper-prior information come from previous literature (Cronin *et. al.* (1994) and Slate *et. al.* (2000)). Slate and Clark (1999) considered the log-transformed PSA level for dependent variables but response measurements $Y_{ij}{}'s$ can be any biological markers which can be indicative of underlying disease status and have potential to improve early detection. We also have log-transformed PSA responses, $\log(PSA+1)$. The Bayesian hierarchical changepoint model with no covariate is

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \tag{1}$$

where the indexes of i and j represent the subject and the number of readings in a subject respectively. The superscript "+" is an indicator function with $z^+ = z$ for $z > 0$ and zero otherwise. In the PSA context, x_{ij} may denote the age in the j^{th} visit of a subject i. The interpretation of all random effects for the trajectory for subject i is that a_{0i} is the intercept, a_{1i} is slope before the changepoint, the b_i is slope after a changepoint and t_i is the changepoint. Our model is three-step hierarchical changepoint model with prior and hyper-prior distributions given below:

$$\begin{pmatrix} a_{0i} \\ a_{1i} \end{pmatrix} | \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \sim \text{MVN}\left\{ \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}, \Omega_a \right\} \tag{2}$$

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \sim \text{MVN}\left\{ \begin{pmatrix} 1 \\ 0.02 \end{pmatrix}, \begin{pmatrix} 100 & 0 \\ 0 & 10000 \end{pmatrix} \right\}$$

$$\Omega_a \sim \text{Wishart}\left\{ \left[5 \begin{pmatrix} 0.1 & 0 \\ 0 & 0.0001 \end{pmatrix} \right]^{-1}, 5 \right\}$$

$$b_i | \beta, \tau_b \sim \text{N}(\beta, \tau_b)$$

$$\beta \sim \text{N}(0.15, 3600) \;\; \& \;\; \tau_b \sim \text{Gamma}(48, 0.0133)$$

$$t_i | \mu_t, \tau_t \sim \text{N}(\mu_t, \tau_t) \;\; \& \;\; \mu_t \sim \text{N}(80, 0.10) \;\; \& \;\; \tau_t \sim \text{Gamma}(47, 4700)$$

$$\varepsilon_{ij} | \tau_i \sim \text{N}(0, \tau_i) \;\; \& \;\; \tau_i \sim \text{Gamma}(5, 0.25)$$

All normal distributions are parameterized in terms of a mean and a precision; thus τ_i, τ_b, τ_t, and Ω_a, are all precisions. The prior information of the random effects for the trajectory for subject

i, $(a_{0i}, a_{1i})'$, is assumed to follow the multivariate normal distribution with mean vector $(\alpha_0, \alpha_1)'$ and precision matrix Ω_a. The model indicates that the hyper-prior for mean vector $(\alpha_0, \alpha_1)'$ follows also multivariate normal distribution with known mean and covariance matrix, and the precision matrix Ω_a is assumed to follow Wishart distribution. The Wishart distribution, a multivariate generalization of the gamma distribution, is the conjugate prior distribution for the inverse covariance matrix Ω_a which is symmetric and positive definite in a multivariate normal distribution. In other words, the conjugate prior for the covariance matrix Σ in multivariate normal distribution is the multivariate generalization of the inverse gamma distribution. In the random effects model with an unknown variance-covariance matrix Σ, it's typical to specify a Wishart distribution for Σ^{-1}. The Gamma(α, β) distributions has mean α/β where shape parameter $\alpha > 0$ and inverse scale $\beta > 0$. Therefore, all prior and hyper-prior distributions are conjugate in the full model. All prior information used in this research came from previous literatures including Carter *et al.* (1992), Whittemore *et al.* (1995) and Clark *et al.* (1996). This model assumes that all men eventually reach their changepoint if they live long enough. Since a large fraction of men does not encounter their changepoint during their life, the population mean of the changepoints, μ_t, tends to be quite large.

2.2 Addition of Covariates

This additive model is to consider risk factors as covariates into Bayesian hierarchical changepoint model (1). The risk factors can be smoking status, drinking usages, BMI, family history of specific disease in the biomedical research and can be categorized as continuous or discrete variables. The covariates must be permitted to influence the dependent observation's trajectory both before and after the changepoint, and also to influence the timing of the changepoint. There are three ways in adding covariates as follow. Let's consider a dummy variable c_i as a covariate of interest in subject i and its corresponding coefficient γ. The covariate c_i has binary values: 1 if the subject i is a man holding the risk factor and 0 if not holding the risk factor.

$$c_i = \begin{cases} 1 & \text{if the subject } i \text{ has the risk factor} \\ 0 & \text{else} \end{cases}$$

The **type I** is to extend the segmented linear regression model (1) by adding the covariate at the overall PSA level. The model (3) shows that c_i does not depend on the observation time j. Thus, it is believed that the covariate affect the readings (response measurement levels) prior to the changepoint without affecting the changepoint itself or the post-changepoint slope.

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + c_i\gamma + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \tag{3}$$

As an illustration, if family history is believed to increase observations levels in men with and without a cancer, we can consider a family history indicator as the fixed effect c_i in the model (3). We consider the corresponding coefficient γ as a population parameter. Since there is no population information for the parameter, we can assume a normal flat prior distribution for γ with zero mean and large variance. If a researcher believe that BMI level increases response measurement levels, we can consider the BMI levels as a continuous fixed effect c_i in the model. Furthermore, when we have more than two levels for discrete variable of interest, we can consider it an ordinal variable if we know the relationship between levels. But since we do not know the exact relationship between levels, we use dummy variables. If we have a variable with three levels, we use two dummy variables which are dichotomous:

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + c_{1i}\gamma_1 + c_{2i}\gamma_2 + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \tag{4}$$

The model (4) shows the addition to the model of a variable with three levels. Assume that a covariate smoking status has three levels such as none, former smokers and current smokers, and we

want to know the effects of smoking on PSA level. Since the smoking indicator has more than two levels, we should have multiple indicator fixed effect terms. Specially the smoking effect has three levels, as described above, so the term, c_{1i} should take the binary values 1 (for former smokers) and 0 (for everyone else), and the term c_{2i} should also take binary values of 1 for current smokers and 0 for everyone else.
The **type II** is when the covariate is believed to affect the change in slope after a changepoint occurs. This is represented in the model shown in (5). We would use model (5) to model observation reading if a risk factor is believed to affect the rate of tumor growth once a tumor is present. For this model, we also assume normal prior distribution for γ with known mean and variance.

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + ((b_i + c_i\gamma) - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \tag{5}$$

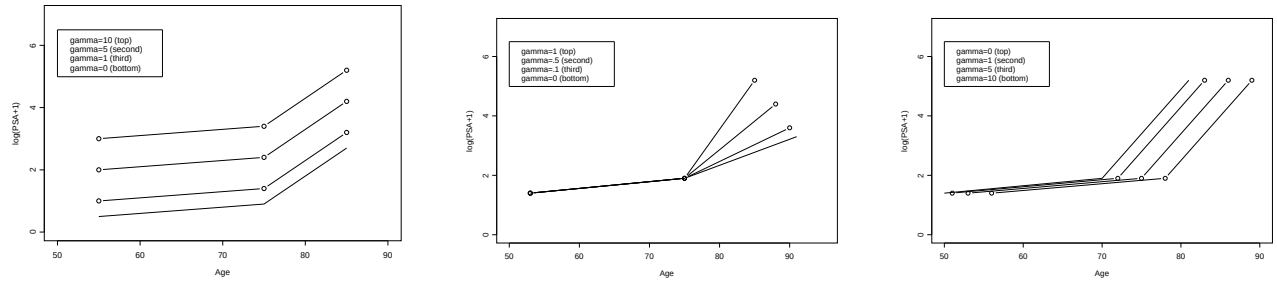
The model (6) can be applied if the covariate is believed to affect the time when a changepoint occurs. This **type III** model is appropriate if it is believed that smoking actually affects when a changepoint occurs. The model (5) is equivalent to shifting the mean of the prior distribution of t_i for the subset of the population with a value of $c_i \neq 0$.

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - (t_i + c_i\gamma))^+ + \varepsilon_{ij} \tag{6}$$

An additive model can contain any combination of covariates for all three addition types. Let's assume that a covariate c_1 of smoking status affects the slope after the changepoint, and that a covariate c_2 of body mass index is expected to affect the changepoint. Then we can consider an additive model:

$$Y_{ij} = a_{0i} + a_{1i}x_{ij} + ((b_i + c_{1i}\gamma_1) - a_{1i})(x_{ij} - (t_i + c_{2i}\gamma_2))^+ + \varepsilon_{ij} \tag{7}$$

We can propose any combination of covariates with the three different types of model described in (3), (5) and (6). Figure (1) shows how the γ's affect the full model when we give three different values for γ for each of the three types. Since type I is assumed for a covariate which only directly affects the response level, the first graph (a) shows that graphs depending on different γ values have different log-transformed PSA levels. As the level of γ increases, we see that the risk factor has a positive effect on the log-transformed PSA levels. The graph (b) for type II indicates that the different γ values lead to three different slopes after the changepoint. Under the type III assumption that a covariate only affects the changepoint, graph (c) shows variation of a changepoint. If a risk factor is significant to a changepoint, the changepoint moves horizontally.



(a) type I (b) type II (c) type III

Figure 1: How the covariate affects the full model depending on three values of γ (a) of type I, (b) of type II, (c) of type III .

2.3 Posterior Estimation for γ

Bayesian inference is used to determine and summarize charactersitics of marginal posterior distributions for the parameters of interest. To estimate parameters of the posterior distribution, these models may be fit using the Gibbs sampler, a special type of MCMC methods similar to those used in Slate and Clark (1999) and Cronin et al. (1994). we use WinBUGS to fit the model with covariates. We consider a model (3) with fully Bayesian framework (2) for Gibbs estimation of **type I**. Let us denote the parameter space θ for a subject i as

$$\theta_i = (\gamma, a_{0i}, a_{1i}, b_i, t_i, \tau_i, \alpha_0, \alpha_1, \Omega_a, \beta, \tau_b, \mu_t, \tau_t) \tag{8}$$

and also denote the full conditional distribution of γ as $p(\,\gamma \mid \cdots)$. The Gibbs sampler enables us to obtain samples from the complete conditional distribution $p(\,\gamma \mid \cdots)$

$$p(\,\gamma \mid \cdots) \;\propto\; p(\gamma) \cdot \prod_{i=1}^{I} \prod_{j=1}^{n_i} p(y_{ij} | a_{0i}, a_{1i}, b_i, t_i, \tau_i, \gamma) \tag{9}$$

where $p(\gamma)$ is the prior distribution of a covariate effect γ, I is the number of subjects, and n_i is the number of readings for subject i. The prior follows a normal distribution with zero mean and a known small precision τ_γ. Under the assumption that random effects and random error are independent,

$$y_{ij} | a_{0i}, a_{1i}, b_i, t_i \sim \mathrm{N}(\mu_{ij}^I, \tau_i I_{n_i}) \tag{10}$$

where $\mu_{ij}^I = a_{0i} + a_{1i}x_{ij} + c_i\gamma + (b_i - a_{1i})(x_{ij} - t_i)^+$. Then the full conditional distribution of γ follows the normal distribution as follows:

$$\begin{aligned} p(\,\gamma \mid \cdots) \;&\propto\; \exp\left\{ -\frac{\tau_i}{2} \sum_{i=1}^{I} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ij}^I)^2 \right\} \cdot \exp\left\{ -\frac{\tau_\gamma}{2}\gamma^2 \right\} \\ &\propto\; \exp\left\{ -\frac{\tau_\gamma + \tau_i n_i \sum_{i=1}^{I} c_i^2}{2} \left(\gamma - \frac{\tau_i \sum_{i=1} \sum_{j=1} \epsilon_{ij}}{\tau_\gamma + \tau_i n_i \sum_{i=1}^{I} c_i^2} \right)^2 \right\} \\ &\sim\; \mathrm{N}\left(\frac{\tau_i \sum_{i=1} \sum_{j=1} \epsilon_{ij}}{\tau_\gamma + \tau_i n_i \sum_{i=1}^{I} c_i^2}, \left(\tau_\gamma + \tau_i n_i \sum_{i=1}^{I} c_i^2 \right)^{-1} \right) \end{aligned} \tag{11}$$

where $\epsilon_{ij} = y_{ij} - (a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - t_i)^+)$ and $y_{ij} - \mu_{ij}^I = \epsilon_{ij} - c_i\gamma$. Similarly, we consider models (5) and (6) with fully Bayesian framework (2), and then can derive the full conditional distributions of γ for **type II** and **type III**. Let $\mu_{ij}^{II} = a_{0i} + a_{1i}x_{ij} + ((b_i + c_i\gamma) - a_{1i})(x_{ij} - t_i)^+$ and $\mu_{ij}^{III} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - (t_i + c_i\gamma))^+$. Then the full conditional distribution of γ for **type II** is

$$p(\gamma | \cdots) \propto \mathrm{N}\left(\frac{\tau_i \sum_i \sum_j \epsilon_{ij} c_i (x_{ij} - t_i)^+}{\tau_\gamma + \tau_i \sum_i \sum_j c_i^2 \left[(x_{ij} - t_i)^+ \right]^2}, \left(\tau_\gamma + \tau_i \sum_i \sum_j c_i^2 \left[(x_{ij} - t_i)^+ \right]^2 \right)^{-1} \right) \tag{12}$$

where $y_{ij} - \mu_{ij}^{II} = \epsilon_{ij} - \gamma c_i (x_{ij} - t_i)^+$, and the full conditional distribution for **type III** exists only when x_{ij} is greater than $t_i + c_i\gamma$,

$$p(\gamma | \cdots) \propto \mathrm{N}\left(\frac{\tau_i \sum_i \sum_j c_i \epsilon_{ij}^0}{\tau_\gamma + \tau_i n_i \sum_i c_i^2}, \left(\tau_\gamma + \tau_i n_i \sum_i c_i^2 \right)^{-1} \right) \tag{13}$$

where $\epsilon_{ij}^0 = y_{ij} - (a_{0i} + a_{1i} + (b_i - a_{1i})(x_{ij} - t_i) - c_i\gamma$.

One common practice to summarize marginal posterior densities is to obtain their basic statistics such as the mean and standard deviation, or to tabulate $100(1-\alpha)\%$ posterior credible intervals for interest parameters. Edwards, Lindman, and Savage (1963) call posterior intervals based on noninformative priors as "credible intervals". We can obtain the credible intervals analytically or by using a MCMC method. Consider a Bayesian posterior density with the form

$$p(\theta|\mathbf{y}) = \frac{1}{C}\ L(\mathbf{y}|\theta)\ P(\theta)$$

where $\mathbf{y}$ denotes observed data, and θ is the collection of all population and subject-specific parameters. $L(\mathbf{y}|\theta)$ is a likelihood function given the data $\mathbf{y}$, p(θ) is a prior, and C is a normalizing constant. Let $p(\theta|\mathbf{y})$ and $\Psi(\theta|\mathbf{y})$ denote the marginal posterior density function and the marginal posterior cumulative distribution function of θ. Then a $100(1-\alpha)\%$ **Bayesian credible interval** for θ is

$$\left(\theta^{\Psi(\theta^{(\alpha/2)}|y)}, \theta^{\Psi(\theta^{(1-\alpha/2)}|y)}\right)$$

We focus on estimating population parameters γ, effects of fixed covariates. We can obtain MCMC samples, $\{\theta^{(i),i=1,2,\ldots,n}\}$, from the marginal posterior distribution $p(\theta|\mathbf{y})$. The order statistic estimator of a $100(1-\alpha)\%$ Bayesian credible interval is

$$\left(\theta_{([(\alpha/2)n])}, \theta_{([(1-\alpha/2)n])}\right)$$

where $\theta_{([(\alpha/2)n])}$ and $\theta_{([(1-\alpha/2)n])}$ are the $[(\alpha/2)n]^{th}$ smallest and the $[(1-\alpha/2)n]^{th}$ smallest observations of $\theta^{(i)}$. $[(\alpha/2)n]$ and $[(1-\alpha/2)n]$ are the integer part of $(\alpha/2)n$ and $(1-\alpha/2)n$. If the marginal posterior densities $p(\theta|\mathbf{y})$ are not symmetric, the credible intervals can be biased. It is appropriate to apply a highest posterior density (HPD) interval when the marginal posterior distributions are skewed.

We want to investigate which covariate is significant from a generalized Bayesian hierarchical changepoint model. This is a univariate testing hypothesis within the full Bayesian framework. Consider a diachotomous covariate c_1 which has an unknown effect on prostate cancer through PSA. As was mentioned earlier, there are three possible ways to describe the effects of a covariate. We would like to know which type has most significant association with the covariate among three covariate addition types: the overall level, the age at which cancer initiates (changepoint), and the growth rate following the changepoint. We can perform hypothesis testing for a **type I** covariate c_1 if it is believed to affect the overall PSA level. Thus, we have a testable null hypothesis $H_0 : c_{1i} = 0$:

$$\begin{aligned} H_0:&\ \text{a reduced model } (c_{1i}{=}0)\\ &\ Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij}\\ H_1:&\ \text{a generalized model } (c_{1i} \neq 0)\\ &\ Y_{ij} = a_{0i} + a_{1i}x_{ij} + c_{1i}\gamma + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \end{aligned}$$

Similarly we can derive a hypothesis test for the **type II** covariate c_2:

$$\begin{aligned} H_0:&\ \text{a reduced model } (c_{2i}{=}0)\\ &\ Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij}\\ H_1:&\ \text{a generalized model } (c_{2i} \neq 0)\\ &\ Y_{ij} = a_{0i} + a_{1i}x_{ij} + ((b_i + c_{2i}\gamma) - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \end{aligned}$$

Moreover, if the covariate is believed to have an important effect on the age at which cancer initiates (changepoint), we can perform the following hypothesis testing procedure:

$$\begin{aligned} H_0: &\text{ a reduced model } (c_{3i}=0) \\ &Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij} \\ H_1: &\text{ a generalized model } (c_{3i} \neq 0) \\ &Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - (t_i + c_{3i}\gamma))^+ + \varepsilon_{ij} \end{aligned}$$

The decision is based on values of credible intervals. If the interval includes zero, we fail to reject the null hypothesis, and we can make the conclusion that the covariate has no significant effect on development of prostate cancer or changes in PSA level.

3 Simulation Study

We provide a simulation study to investigate the behavior of estimates of covariate effects in a Bayesian hierarchical changepoint model. We consider three ways to include covariates in the model (2): **type I** (3), **type II** (5), and **type III** (6) as described in previous section. We generate data sets from these three different types of models, and fit the data through those three models respectively. Then, we examine estimates $\hat{\gamma}$ of a covariate effect, and compare them with the true value of γ. The simulation results include mean values, bias, relative bias and average mean square error (MSE)

3.1 Data generation

We generate 135 data sets depending on five replications of three types of generalization, three types of case-control proportions, and three different true values of γ. Each data set consists of 100 subjects which have unequal log-transformed PSA readings. We consider three different types of case-control proportions: 30% cases and 70% controls, 50% cases and 50% controls, and 70% cases and 30% controls. The model assumes that everyone eventually has a changepoint if he lives long enough. Therefore, we define a case when a changepoint occurs earlier than the age at the last reading, and a control if it does not. For a more complete investigation, we also suppose three different true values for γ: 1, 5 and 10 for type I & III, and 0.1, 0.5 and 1.0 for type II. Therefore, each category has five data sets, and each model type has 45 data sets.

The model has a three-stage hierarchical structure. For the first stage, we use 100 subjects ($i = 1, \ldots, 100$) but the number of readings in a subject, j, is assumed to follow a Poisson distribution with a mean of 9. The lag interval between visits is 0.5 years (6 months). To generate a baseline age for each subject, we consider NPCT data which shows a mean of 72, and range of 56 through 87. Thus, 99% of subject ages have a range between 56.7 and 84.6, so we generate them as $72 + N(0, 5)$. The second stage describes the prior distribution of random effect parameters a_{0i}, a_{1i}, b_i and t_i, while population parameters $\alpha_0, \alpha_1, \Omega_a, \beta, \tau_b, \mu_t$, and τ_t are generated at the third stage under the same condition described in Table 1. It is assumed that we do not know the prior information of parameters of related covariates. In other words, we can not assume that there is a positive effect for a specific covariate. For example, when we investigate whether or not current smoking status has a significant effect on PSA level or disease onset time, we use a relatively *flat* prior population information. This can lead to a negative effect as well as a positive effect. Consequently, the prior information does not dominate the likelihood. The normal prior distribution can be a diffuse prior with a mean zero and a large variance 100. Since the prior minimally affects the joint likelihood function, it is a non-informative prior. The models in this simulation study for type I, II and III are: The generation model for types II and III is the same as for type I, except for having a different first

1. First stage : [I] $Y_{ij} = a_{0i} + a_{1i}x_{ij} + c_i\gamma + (b_i - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij}$
[II] $Y_{ij} = a_{0i} + a_{1i}x_{ij} + ((b_i + c_i\gamma) - a_{1i})(x_{ij} - t_i)^+ + \varepsilon_{ij}$
[III] $Y_{ij} = a_{0i} + a_{1i}x_{ij} + (b_i - a_{1i})(x_{ij} - (t_i + c_i\gamma))^+ + \varepsilon_{ij}$
where $\varepsilon_{ij}|\tau_i \sim \mathrm{N}(0, \tau_i)$

2. Second stage : prior information for random effects parameters
$\begin{pmatrix} a_{0i} \\ a_{1i} \end{pmatrix} | \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \sim \mathrm{MVN}\left\{ \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}, \Omega_a \right\}$
$b_i|\beta, \tau_b \sim \mathrm{N}(\beta, \tau_b)$ and $t_i|\mu_t, \tau_t \sim \mathrm{N}(\mu_t, \tau_t)$
$\tau_i \sim \mathrm{Gamma}(5, 0.25)$

3. Thirds stage : hyper-prior information for population parameters
$\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \sim \mathrm{MVN}\left\{ \begin{pmatrix} 1 \\ 0.02 \end{pmatrix}, \begin{pmatrix} 100 & 0 \\ 0 & 10000 \end{pmatrix} \right\}$
$\Omega_a \sim \textsc{Wishart}\left\{ \left[5 \begin{pmatrix} 0.1 & 0 \\ 0 & 0.0001 \end{pmatrix} \right]^{-1}, 5 \right\}$
$\beta \sim \mathrm{N}(0.15, 3600)$ and $\tau_b \sim \mathrm{Gamma}(48, 0.0133)$
$\mu_t \sim \mathrm{N}(80, 0.10)$ and $\tau_t \sim \mathrm{Gamma}(47, 4700)$
$\gamma \sim \mathrm{N}(0, 100)$

Table 1: The Model description in this simulation study for type I, II, and III

stage. The table 2 describes mean values of log-PSA, slope and changepoint for each type. The mean values of log-transformed PSA, slope after a changepoint and a changepoint are 2.48, 0.15 and 82.5 for type I, 4.31, 0.16 and 79.5 for type II, and 3.28, 0.17 and 77.1 for type III. It is of interest that the slope after a changepoint increases as the case-control ratio increases, and that a changepoint which indicates the onset of prostate cancer decreases as the ratio increases. The figure (2) shows

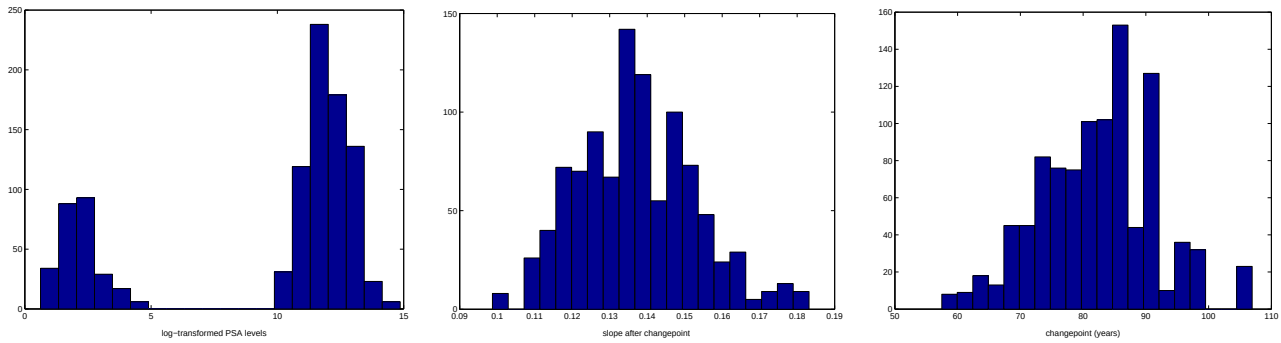
	type I	type II	type III
log-PSA	2.48	4.31	3.28
Baseline Age	72.2	73.1	71.3
slope	0.150	0.158	0.165
changepoint	82.5	79.5	77.1

Table 2: Mean values of log-PSA, slope and changepoint for three types

some information of one of our 135 generated data sets. The data is generated with gamma = 10, 30:70 for case-control ratio, and type I covariate method. Since we assume that the ratio between values "0" and "1" is 30% vs 70% to generate the coefficient of γ, c_i, the first graph (a) shows two groups of log-transformed PSA levels. The left side histogram shows log-transformed PSA levels of never-smoker group which is 30 % from generated subjects and the right one indicates those of ever-smoker group which occupies 70 % among 100 subjects. The graph (b) and (c) shows that the median of the slope after changepoint and the changepoint are around 0.15 and 80 respectively. We know that since only response measurement is affected by the covariate under type I, the slope after changepoint (b) and the changepoint (c) does not much varied.

3.2 Bayesian inference

We estimate the marginal posterior distribution of γ for each data set. When the prior and hyper-prior distributions of the generalized model are conjugate with each other, the Gibbs sampler can be easily applied to find the estimate of the marginal posterior distribution of γ. We give a non-



(a) (b) (c)

Figure 2: The distributions of (a) log-transformed PSA, (b) slope after changepoint, (c) changepoint

informative prior $N(0, 0.01)$ for γ, since we want a posterior distribution that is dominated by the likelihood function. The normal prior is a proper prior, which always leads a proper posterior distribution. We used the WinBUGS program to execute the Gibbs sampler. We denote the parameter vector θ as $\theta = (\alpha_0, \alpha_1, \Omega_a, \beta, \mu_t, \tau_b, \tau_t, a_{0i}, a_{1i}, b_i, t_i, \tau_i, \gamma)$. The algorithm samples from the full conditional distributions as follows:

1. Set starting values : $\alpha_0^{(0)}, \alpha_1^{(0)}, \beta^{(0)}, \ldots, t_i^{(0)}, \tau_i^{(0)}, \gamma^{(0)}$
2. Update full conditional density for α_0 : $\alpha_0^{(1)} \mid \alpha_1^{(0)}, \beta^{(0)}, \ \ldots, \ t_i^{(0)}, \tau_i^{(0)}, \gamma^{(0)}$
3. Update full conditional density for α_1 : $\alpha_1^{(1)} \mid \alpha_0^{(1)}, \beta^{(0)}, \ \ldots, \ t_i^{(0)}, \tau_i^{(0)}, \gamma^{(0)}$
4. ...
5. Update full conditional density for γ : $\gamma^{(1)} \mid \alpha_0^{(1)}, \alpha_1^{(1)}, \beta^{(1)}, \ \ldots, \ t_i^{(1)}, \tau_i^{(1)}$
6. Repeat steps 2 through 5 until reaching covergency

We ran the Gibbs sampler with initial values of $a_{0i} = 0.1, a_{1i} = 0.01, b_i = 0.5, t_i = 72, \beta = 1, \alpha_0 = 0.5, \alpha_1 = 0.02, \tau_b = 50, \tau_t = 50, \tau_i = 0.1, \gamma = 0.5$. We iterated five thousand times after five thousand burn-in iterations. The running time is approximately 4 min, 12 min and 27 min for type I, II and III respectively. Markov chain Monte Carlo (MCMC) algorithm produces samples from a distribution which is believed to be converged to the stationary distribution of an irreducible Markov chain. Thus it is natural to determine what point can be reasonably believed that the samples comes truly from the underlying stationary distribution. Since Markov property generally leads correlation among members of this sample, it slows the algorithm to sample from the entire stationary distribution. Cowles and Carlin (1996) compared several convergence diagnostic methods. The failure of convergency on a Markov chain induced by the MCMC algorithm causes biased and unreliable posterior estimates. So it is one of the most important steps in MCMC sampling to investigate convergence diagnostics.

3.3 Simulation results

The statistics of interests are both means of posterior means of an interest parameter of a covariate and means of posterior standard deviations of an interest parameter of a covariate. Suppose γ be the true value of the interest parameter of a covariate and $\hat{\gamma}_i$ be the mean of posterior means. Then, we can get information of the estimates through the mean square estimate, bias or relative bias. The mean square error combines both bias and variance.

$$\begin{aligned} MSE &= E_\gamma \left\{\hat{\gamma} - \gamma\right\}^2 \\ &= E_\gamma \left\{\hat{\gamma} - E(\hat{\gamma})\right\}^2 + E_\gamma \left\{E(\hat{\gamma}) - \gamma\right\}^2 \end{aligned}$$

Type I		Case-Control proportion					
		30 : 70		50 : 50		70 : 30	
		Mean	Std	Mean	Std	Mean	Std
Gamma = 1.0	Mean	0.6136	0.3527	0.5698	0.5078	0.5324	0.7277
	Bias	0.3864		0.4302		0.4676	
	Relative Bias	0.3864		0.4302		0.4676	
	MSE	0.1911		0.2037		0.2336	
Gamma = 5.0	Mean	3.7850	0.4974	3.2835	0.6405	3.1527	0.8841
	Bias	1.2150		1.7165		1.8473	
	Relative Bias	0.2430		0.3433		0.3695	
	MSE	1.5199		2.9813		3.4125	
Gamma = 10.0	Mean	7.7401	0.7922	7.4797	1.1129	6.4080	1.5095
	Bias	2.2599		2.5203		3.5920	
	Relative Bias	0.2260		0.2520		0.3592	
	MSE	6.5969		7.2071		14.806	

Table 3: Summary of Bayesian estimation of marginal posterior distribution of γ and mean square errors of type I.

$$= \quad \mathrm{Var}_\gamma[\hat{\gamma}] + [\mathrm{bias}_\gamma(\hat{\gamma})]^2$$

We focus on average mean square error and relative bias. The average mean square error is average of the mean square errors and

$$avgMSE(\gamma) = \frac{1}{N}\sum_{i=1}^{N} E_\gamma\left\{\hat{\gamma}_i - \gamma\right\}^2 \tag{14}$$

the relative bias is the proportion of difference between the true value and the estimate

$$\text{Relative Bias} = \frac{\gamma - \bar{\gamma}}{\gamma} \tag{15}$$

We provided three types of additive methods to consider a covariate into the full Bayesian framework. The followings are three tables of simulation result. Each type has nine categories which grouped by three different types of true gamma values and case-control ratios. Each categories repeated five times and calculate average of the estimates and average mean square error. Each estimate value for γ is calculated using Gibbs sampler and is investigated the critical region for it. Even though five estimates per each category are somewhat varied each other, all estimates have critical regions including their true gamma values. Since the estimates are repeated for only five times, the average of the means can be far from the true value. But if the enough sample sizes are allowed, the average of the estimates for each category go very close to the true value. Three tables 2, 3 and 4 are the simulation result. Each table shows that the average of the estimate for γ looks close to the true value which implies that this Bayesian additive model can be effective for Bayesian inference for investigating the effects of risk factors. For type I, we set initial values of γ with 10 for true gamma values of 1 and 5, and 1 for true value 10. The others are $b = 0.5$, $t = 72$, $a = \begin{pmatrix} 1 \\ 0.02 \end{pmatrix}$, $\tau = 0.1$ for random coefficients and $\beta_c = 1$, $\beta_\tau = 50$, $t_c = 75$, $t_\tau = 50$, $\alpha = \begin{pmatrix} 0.5 \\ 0.01 \end{pmatrix}$ and $\Omega = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ for fixed coefficients. We used first 10000 iterations with 5000 burn-in iterations for Bayesian estimates of γ for type I. Table 2 shows that parameter estimates for a fixed coefficient parameter, γ, of type I, are close to the true values. For type II, we use same initial values for all related parameters including γ and also used first 10000 iterations with 5000 burn-in iterations for Bayesian estimates of γ for type II. Table 3 shows that parameter estimates for a fixed coefficient parameter, γ, of type II, are

Type II		Case-Control proportion					
		30 : 70		50 : 50		70 : 30	
		Mean	Std	Mean	Std	Mean	Std
Gamma = 0.1	Mean	0.0508	0.0225	0.0644	0.0166	0.0247	0.0127
	Bias	0.0492		0.0356		0.0753	
	Relative Bias	0.492		0.356		0.753	
	MSE	0.0029		0.0032		0.0059	
Gamma = 0.5	Mean	0.4376	0.0245	0.3797	0.0203	0.3883	0.0318
	Bias	0.0624		0.1203		0.1117	
	Relative Bias	0.1248		0.2406		0.2234	
	MSE	0.0129		0.0562		0.0264	
Gamma = 1.0	Mean	0.9433	0.0207	0.9739	0.0418	0.8618	0.0538
	Bias	0.0567		0.0263		0.1382	
	Relative Bias	0.0567		0.0263		0.1382	
	MSE	0.0339		0.0089		0.0309	

Table 4: Summary of Bayesian estimation of marginal posterior distribution of γ and mean square errors of type II.

Type III		Case-Control proportion					
		30 : 70		50 : 50		70 : 30	
		Mean	Std	Mean	Std	Mean	Std
Gamma = 1.0	Mean	0.818	3.098	0.709	3.190	0.621	3.508
	Bias	0.192		0.291		0.379	
	Relative Bias	0.192		0.291		0.379	
	MSE	35.95		85.10		36.01	
Gamma = 5.0	Mean	3.958	2.914	3.944	2.949	3.839	2.231
	Bias	1.042		1.056		1.161	
	Relative Bias	0.208		0.211		0.232	
	MSE	16.75		34.76		73.88	
Gamma = 10.0	Mean	6.816	0.588	7.390	0.604	6.785	0.788
	Bias	3.184		2.610		3.215	
	Relative Bias	0.318		0.261		0.322	
	MSE	52.42		34.78		55.51	

Table 5: Summary of Bayesian estimation of marginal posterior distribution of γ and mean square errors of type III.

close to the true values. For type III, we run 50000 iterations with 20000 burn-in iterations since the model needs more iterations rather than the other two types.
This simulation study shows that parameter estimates for a fixed coefficient parameter, γ, are close to the true values. The bias, relative bias and MSE of the Bayesian inference relatively small, which indicates that this Bayesian additive model can be effective for Bayesian inference for investigating the effects of risk factors.

4 Investigation of Misspecification

The word "misspecification" has been widely used for any inappropriateness of the fitted model. If observed data are not conditionally normal but we use a normal model to fit the data, it can be called distributional misspecification or model misspecification. In this simulation, we have shown that these three types of generalizing Bayesian hierarchical changepoint model can be effective to have an early detection of a underlying disease when an important covariate is added into the model. In summary, we generate data sets depending on three types of covariate addition to model, fit the

data using a model same as data generated, investigate each critical regions, and finally provide average mean square error and relative bias to show that this Bayesian additive model can be effective to investigate the effects of risk factors. We have two aims here: One is to investigate the effects of model misspecification for our simulated three types of data sets, and the other is to show that the misspecification can be used as a method of model diagnosis by showing that misspecification effects can lead to wrong posterior parameter estimation of subjects-specific parameters and no population parameter effect of a covariate. We fit data set generated by type I model with type II model. Then we investigate posterior distribution of subject-specific coefficients of interception, slope after a changepoint and an added covariate parameter between model specification of type I and misspecification to model II. Similarly we do type II and type III data sets.

We consider just one covariate into the model (2), and the types I, II & III are the same as the models (3),(5) and (6). We generate three types of data sets corresponding three types of model types. Firstly we fit the data generated by type I using both a correct model, type I and a wrong model, type II model. We use same prior information and initial values as used in section 2, and run ten thousands iterations with five thousands burn-in to get the Gibbs samplers.

			Node γ				
Data from	Fitted by	true value	mean	std	2.5%	median	97.5%
type I	type I	1.0	1.04	0.57	0.408	0.723	2.072
	type II	1.0	-0.03	0.02	-0.067	-0.029	0.011
type II	type II	1.0	0.96	0.06	0.082	0.971	1.081
	type III	1.0	-0.38	1.28	-2.245	-0.419	1.849
type III	type II	1.0	0.04	0.02	-0.039	0.004	0.047
	type III	1.0	4.32	2.69	-0.079	4.047	10.44

Table 6: Summary table : Posterior estimation of mispecification for a coefficient γ

The data set has a value 1 for a true value of λ. The table 5 shows that the 95 % critical regions for γ include a value 1 when we use a type I model as a fitted model, which means that the fitted model is correct. Meanwhile, that of γ for a type II fitted model include zero which indicates that there exists no significant effect of a covariate. Similarly we investigate the effect of misspecification for the wrong fitted model of type III to a data set generated by type II model and for the wrong fitted model of type II to a data set generated by type III. For a data set generated from type II model, the critical regions for γ include a value 1 when we use a right model type II as a fitted model while that of γ for a wrong type III fitted mode include zero which indicates that there exists no significant effect of a covariate. The result with a data set generated from type III shows same result as like type II data set. That is, the credible interval of the correct model(type III) contains the true value 1 of the covariance coefficient γ. For a misspecification(type II model fitted), the 95 % credible interval includes zero, that shows no significance on the covariate effect.

5 Discussion

We only use one covariate for this simulation study. But this methods can be easily extended to a model with two or more covariates. The choice of the additive type might depend on characteristics of risk factors. Thus it is recommendable to consult with related previous studies. The covariates can be added into a model as a form of linear combination or non-linear combination. We can also consider possibilities to use non-normal priors or mixture normal prior for estimates of the covariates.

References

[1] American Cancer Society. "Cancer Facts and Figures 2004" and "Prostate Cancer", Cancer Reference Information, www.cancer.org, 2004.

[2] Burnham, K.P. and Anderson, D.R. (1998). Model Selection and Inference: A Practical Information-Theoretical Approach. Springer.

[3] Carlin, B.P., Gelfand, A.E. and Smith, A.F.M. (1992). Hierarchical Bayesian Analysis of Changepoint Problems. *Applied Statistics,* 41(2). 389-405.

[4] Carter, H.B., Morrell, C.H., Pearson, J.D., Brant, L.J., Plato, C.C., Metter, E.J., Chan, D.W., Fozard, J.L. and Walsh, P.C. (1992). Estimation of prostatic growth using serial prostate-specific antigen measurements in men with and without prostate disease. *Cancer Rsearch,* 52,

[5] Chen, M.H., Shao, Q., and Ibrahim, J.G. (2000). Monte Carlo Methods in Bayesian Computation. Springer-Verlag.

[6] Chen, M.H. and Shao, Q. (1999). Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *J. Comp. Graph. Statist.,* **8,** No. 1, 69-92.

[7] Clark, L.C., Combs, G.F., Turnbull, B.W., Slate, E.H., Alberts, D.S., Abele, D., Allison, R.J., Chalker, D.K., Gross, E.G., Hendrix, J.D., Herlong, J.H., Hixson, L.J., Kight, F., Krongrad, A., Lesher, J.L., Moore, J., Park, H.K., Rice, J.S., Rogers, A.I., Sanders, B.B., Schuman, B., Smith, C.L., Smith, E.H., Tayor, J.R., and Woodard, J.C. (1996). The nutritional prevention of cancer with selenium 1983-1993: A randomized clinical trial. *Journal of the American Medical Association,* 276, 1957-1963.

[8] Cowles, M.K. and Carlin, B.P. (1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comprehensive Review. *J. Amer. Statist. Assoc.,* **91,** No. 434, 883-904.

[9] Cronin, K.A. (1995). Detection of Changepoints in Longitudinal Data, Phd. Thesis. School of Operations Reasearch, Cornell University.

[10] Cronin, K.A., Slate, E.H., Turbull, B.W. and Wells, M.T.(1994). Using the Gibbs sampler to detect changepoints: application to PSA as a longitudinal marker for prostate cancer. *Comput. Sci. Statist.,* 26, eds. J. Sall and A. Lehman, Interface Foundation of North America, pp. 314-318.

[11] Gelfand A.E. and Smith, A.F. (1990) Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association,* **85,** No.410, 398-409.

[12] Gelfand A.E. and Smith, A.F. (1991) Gibbs Sampling for marginal posterior expectations. *Communication in Statistics,* **20(5 & 6),** 1747-1766.

[13] Gelman, A., Carlin, J.B. and Stern, H.S. (1997) *Bayesian Data Analysis.* Boca Raton: Chapman & Hall/CRC.

[14] George, E.I. and McCulloch, R.E. (1993). Variable selection via Gibbs sampling. *J. Am. Statis. Ass,* **88,** 881-889.

[15] Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996). Markov Chain Monte Carlo In Practice. Chpaman & Hall

[16] Hiatt, R.A., Armstrong, M.A., Klatsky, A.L. and Sidney, S. (1994). Alcohol consumption, smoking, and other risk factors and prostate cancer in a large health plan cohort in California (United States). *Cancer Cases and Control,* 5, 66-72.

[17] Kiuchi, A.S., Hartigan, J.A., Holford, T.R., Rubinstein, P. and Stevens, C.E. (1995). Change Points in the Series of T4 Counts Prior to AIDS. *Biometrics,* 51. 236-248.

[18] Lange, N., Carlin, B.P., Gelfand, A.E. (1992). Hierarchical Bayes Models for the Progression of HIV Infection Using Longitudinal CD4 T-Cell Numbers. *Journal of the American Statistical Association,* 87, 615-632. School of Operations Reasearch, Cornell University.

[19] Murtaugh, P.A. and Schaid, D.J. (1991). Application of R.O.C. curve methodology when markers are repeated measures. *Procedding of the International Biometric Society (ENAR),* Houston, TX, March 24-27.

[20] Slate, E.H. and Cronin, K.A. (1997). Changepoint modeling of longitudinal PSA as a biomarker for prostate cancer. *Case Studies in Bayesian Statistics III,* eds. C. Gatsonis, J.S. Hodges, R.E. Kass, R. McCulloch, P. Rossi and N.D. Singpurwall, (eds), Springer-Verlag, New York, pp. 444-456.

[21] Slate, E.H. and Clack, L.C. (1999). Using PSA to detect prostate cancer onset: An application of Bayesian retrospective and prospective changepoint identification. *Case Studies in Bayesian Statistics IV,* eds. 511-534.

[22] Slate, E.H. and Turnbull, B.W. (2000). Statistical methods for longitudinal biomarkers of disease onset. *Statistics in Medicine,* 19(4), 617-637.

[23] Spiegelhalter, D.J., Thomsa, A., Best. N.G. and Gilks, W.R. (1996) *BUGS: Bayesian Inference using Gibbs Sampling, Version 0.5.* Cambridge: Medical Research Council Biostatistics Unit.

[24] Zeger, S. and Karim, M.R. (1991). Generalized Linear Models with Random Effects: A Gibbs Sampling Approach. *Journal of the American Statistical Association,* **86,** No. 413, 79-85.