Published in final edited form as:
Comput Stat Data Anal. 2012 February 1; 56(2): 362–369. doi:10.1016/j.csda.2011.07.011.

Bayesian piecewise mixture model for racial disparity in prostate cancer progression

L. Zhao[a,*] and **M. Banerjee**[b]

[a]Biostatistics Unit University of Michigan Comprehensive Cancer Center, University of Michigan, Ann Arbor

[b]Department of Biostatistics University of Michigan, Ann Arbor

Abstract

Racial differences in prostate cancer incidence and mortality have been reported. Several authors hypothesize that African Americans have a more rapid growth rate of prostate cancer compared to Caucasians, that manifests in higher recurrence and lower survival rates in the former group. In this paper we propose a Bayesian piecewise mixture model to characterize PSA progression over time in African Americans and Caucasians, using follow-up serial PSA measurements after surgery. Each individual's PSA trajectory is hypothesized to have a latent phase immediately following surgery followed by a rapid increase in PSA indicating regrowth of the tumor. The true time of transition from the latent phase to the rapid growth phase is unknown, and can vary across individuals, suggesting a random change point across individuals. Furthermore, some patients may not experience the latent phase due to the cancer having already spread outside the prostate before undergoing surgery. We propose a two-component mixture model to accommodate patients both with and without a latent phase. Within the framework of this mixture model, patients who do not have a latent phase are allowed to have different rates of PSA rise; patients who have a latent phase are allowed to have different PSA trajectories, represented by subject-specific change points and rates of PSA rise before and after the change point. The proposed Bayesian methodology is implemented using Markov Chain Monte Carlo techniques. Model selection is performed using deviance information criteria based on the observed and complete likelihoods. Finally, we illustrate the methods using a prostate cancer dataset.

Keywords

random change point; mixture distribution; PSA profiles; MCMC; DIC

1. Introduction

Prostate cancer is the most frequently diagnosed cancer in men, and the second leading cause of cancer-specific death among men in the United States. For reasons that remain unclear, prostate cancer incidence rates are significantly higher in African Americans (AA)

*Corresponding author: Tel.: 7347636898, zhaolili@umich.edu (L. Zhao), mousumib@umich.edu (M. Banerjee).

than in Caucasians (CC). Mortality from prostate cancer is also two to three times greater among AA than among similarly aged CC (American Cancer Society, 2010).

Since the advent of the PSA screening era, men are more often diagnosed at an earlier stage of disease. This has led to an increase in surgery over the past two decades (American Cancer Society, 2010; Lu-Yao et al., 1993). However, approximately 30-35% of these surgically treated patients will have pathologically detected disease outside the prostate and 25-30% will ultimately develop disease progression (Freedland et al., 2003; Lu-Yao et al., 1996). Outcome data of men treated with surgery demonstrate more advanced tumors and higher recurrence rates among AA compared to similarly aged CC (Moul et al., 1996; Powell et al., 1999). The observed differential in the recurrence rates following surgery could be a manifestation of an earlier or a more aggressive disease progression in AA compared to CC.

In the medical literature, recurrence of prostate cancer after surgery is defined as a progressive or sustained elevation of serum prostate-specific antigen (PSA); and is based on an elevated PSA level of 0.4 ng/ml. Current studies in the medical literature are based on analyses of time to recurrence, focusing only on the time from surgery to the time when the PSA measurement exceeded the threshold (0.4 ng/ml.). Given that PSA is a moving target, serial measures of PSA in time describe a dynamic process rather than a static one, and can contribute towards understanding the pattern by which the threshold was attained. This paper develops a Bayesian piecewise mixture model to characterize PSA progression over time in AA and CC. Each individual's PSA trajectory is hypothesized to have a latent phase immediately following surgery followed by a rapid increase in PSA indicating regrowth of the tumor. The true time of transition from the latent phase to the rapid growth phase is unknown, and can vary across individuals; suggesting a random change point across individuals. The PSA profiles of AA may be characterized by an earlier change point or a steeper rise during the tumor regrowth phase or both.

Bayesian change point methods have been successfully applied to CD4 counts to predict the timing of HIV viral rebound (Kiuchi et al., 1995), to pre-diagnosis PSA profiles for predicting cancer onset (Bellera et al. 2008; Slate and Lark, 2001), and to model cognitive ability over time preceding diagnosis of dementia (Hall et al., 2003). Among patients undergoing surgery for prostate cancer, nearly one-thirds of the men will have pathologically detected disease that had already spread outside the prostate gland (Freedland et al., 2003). For these men, surgery will be useless and these patients may experience a rapid increase in PSA immediately following surgery, without going through a latent phase. We propose a two-component mixture model to accommodate patients with, as well as, without a latent phase. Within the framework of this mixture model, patients who do not have a latent phase are allowed to have different rates of PSA rise; patients who have a latent phase are allowed to have different PSA trajectories, represented by subject-specific change point, PSA value at the change point and rates of PSA rise before and after the change point. We use simulation-based approach that exploits recent advances in Markov Chain Monte Carlo (MCMC) techniques, to implement the proposed Bayesian methodology. Section 2 describes the model formulation. In section 3 we present details of the Bayesian analysis, specifically prior models for the various parameters, and details of the MCMC

technique for posterior analyses. Section 4 presents results of data analyses of a cohort of prostate cancer patients. Finally, section 5 contains some concluding remarks and an appraisal of the approach adopted in the current article in contrast to other approaches used in this context.

2. Piecewise mixture model

2.1. Study description

The specific study involved patients who were surgically treated between 2001 and 2006 for clinically localized prostate cancer at Harper Hospital in Detroit, Michigan. Postoperative follow-up included serum PSA level measurements (along with digital rectal examination) every 3 months for the first two years, and every 6 months thereafter. We retained 94 men who had an undetectable serum PSA level (defined as an assay value of 0.05 ng/ml) right after surgery, no preoperative or postoperative hormonal and/or radiation therapy, race CC or AA, and who had experienced recurrence on or before 2009. In this study, 68% were CC; 47% had local, and 53% had regional stage disease. Mean pre-surgery PSA was 14.5 ng/ml (range 2.8-62.8 ng/ml). Median follow-up for these patients was 36 months (range 8-97 months). All PSA measurements recorded at or before the first recurrence were included in the analysis.

2.2. PSA trajectories

After surgery, PSA levels are undetectable and go through a latent phase (characterized by very slow change in PSA) and then start to increase again at variable rates across individuals, with close to exponential patterns once PSA rise begins. For this reason, we applied a base 2 logarithmic transformation to the PSA measurements. As pointed out by Bellera et al. (2008), this transformation has several advantages. First, it allows one to model the individual patient trajectories as piecewise linear before and after the change point within individuals. Second, the trajectories tend to be smoother than when plotted on the original PSA scale. The base 2 logarithmic transformation also has scientifically useful interpretation; the $\log_2$PSA growth rate is equivalent to the number of PSA doublings per month, and its reciprocal corresponds to the PSA doubling time, a quantity of interest to clinicians.

Figure 1 presents a $\log_2$PSA trajectory over time for a given individual *i*; the time origin is at surgery. We are primarily interested in estimating the $\log_2$PSA rate of change during the latent phase (i.e. slope of the first line), the $\log_2$PSA rate of rise during the rapid phase (i.e. slope of the second line), the change point or time of transition, and the variability associated with these parameters in the population. Furthermore, we are interested in assessing whether these parameters differ by race, after adjusting for cancer stage at diagnosis and pre-surgery PSA.

2.3. Model formulation

A piecewise linear model was used to model the slow change in $\log_2$PSA before the transition time and rapid change in $\log_2$PSA after the transition time. Each patient is allowed to have a different $\log_2$PSA trajectory, represented by a subject-specific slope before the

transition point, a subject-specific transition point, a subject-specific $\log_2$PSA value at the transition point and a subject-specific slope after the transition point. However, not all patients have a latent phase. Some patients may have pathologically detected disease that had already spread outside the prostate gland. For these patients, surgery will be useless and they may experience a rapid increase in PSA immediately following surgery, without going through a latent phase. A straight line model for $\log_2$PSA was used to model the rapid phase for such patients. Therefore, we used a two component mixture ($M = 2$) model to model the $\log_2$PSA growth; specifically, we used mixture probabilities p and $1 - p$ to accommodate patients both with and without a latent phase, respectively.

Let y_{ij}, $i = 1, \ldots, n$; $j = 1, \ldots, n_i$, be the $\log_2$ PSA value for subject i at follow-up j, and t_{ij} the time (in months) at which the jth follow-up occurred. Allowing for the mixed trajectories, we consider the following model for y_{ij}

$$y_{ij}=a_i+b_{1i,z_i}(t_{ij}-r_{i,z_i})_-+b_{2i}(t_{ij}-r_{i,z_i})_++e_{ij} \quad (1)$$

where the functions x_- and x_+ correspond to min(x, 0) and max(x, 0), respectively. For subject i, z_i is an unobserved indicator variable that takes the value 1 if subject i has a latent phase and 0 otherwise. If a subject has a latent phase ($z_i = 1$), the term $b_{1i,1}(t_{ij} - r_{i,1})$ corresponds to the linear model for $\log_2$ PSA before the transition time, and $b_{2i}(t_{ij} - r_{i,1})$ corresponds to the linear model for $\log_2$ PSA after the transition time; thereby giving rise to a piecewise linear model. Specifically, $r_{i,1}$ is the change point or time of transition, a_i is the $\log_2$ PSA at the change point, $b_{1i,1}$ and b_{2i} are the $\log_2$ PSA rates of change before and after the change point. In contrast, if subject i has rapid PSA growth right after surgery without a latent phase ($z_i = 0$), only $b_{2i}(t_{ij} - r_{i,0})$ is used to model the trajectory, and $r_{i,0} = 1$ represents that PSA rise started immediately after surgery. The term $b_{1i,0}(t_{ij} - r_{i,0})_- = 0$ since $t_{ij} \geq r_{i,0} = 1$ for all j, and $b_{1i,0}$ is set to 1 for identifiability.

We assume that the $y_{ij}|z_i$'s are normally distributed, with mean $\mu_{z_i ij}$ and variance σ^2, where

$$\mu_{1ij}=a_i+b_{1i,1}(t_{ij}-r_{i,1})_-+b_{2i}(t_{ij}-r_{i,1})_+ \text{ when } z_i=1$$
$$\mu_{0ij}=a_i+b_{2i}(t_{ij}-1) \text{ when } z_i=0$$

3. Bayesian analysis

Let $a = (a_1, \cdots, a_n)'$, $b_1 = (b_{11,1}, \cdots, b_{1n,1})'$, $b_2 = (b_{21}, \cdots, b_{2n})'$, and $r = (r_{1,1}, \cdots, r_{n,1})'$.

Assume

$$\begin{gathered} b_{1i,1}|\mu_{b_1},\sigma^2_{b_1}\sim N(\mu_{b_1},\sigma^2_{b_1}), \\ a_i|\mu_a,\sigma^2_a\sim N(\mu_a,\sigma^2_a), \\ r_{i,1}|\mu_r,\sigma^2_r\sim N(\mu_r,\sigma^2_r), \\ b_{2i}|\mu_{b_2},\sigma^2_{b_2}\sim N(\mu_{b_2},\sigma^2_{b_2}), \\ \text{and } z_i|p\sim Bernoulli(p). \end{gathered}$$

The observed data associated with this model is $\mathbf{y} = (y_1, \cdots, y_n)'$, where y_i is a $n_i \times 1$ vector of serial $\log_2$PSA values for subject i, and the corresponding latent indicator $\mathbf{z} = (z_1, \cdots, z_n)$ constitutes missing data. Following the EM terminology, the likelihood $l(\mathbf{y}|\boldsymbol{\theta})$ is often called the *observed likelihood*, while $l(\mathbf{y},\mathbf{z}|\boldsymbol{\theta})$ is called the *complete likelihood*. The observed likelihood function can be written as

$$l(\mathbf{y}|\boldsymbol{\theta})=\prod_{i=1}^{n}\prod_{j=1}^{n_i}[(1-p)\phi(y_{ij}|\mu_{0ij},\sigma^2)+p\phi(y_{ij}|\mu_{1ij},\sigma^2)] \quad (2)$$

where $\boldsymbol{\theta} = (a, b_1, r, b_2, p, \sigma^2)'$ and $\varphi(\cdot|\mu_{z_i ij}, \sigma^2)$ is the normal pdf. By introducing the latent indicator $\mathbf{z} = (z_1, \cdots, z_n)$, the above can be interpreted as a missing data model and the corresponding complete likelihood can then be written as

$$l(\mathbf{y},\mathbf{z}|\boldsymbol{\theta})=\prod_{i=1}^{n}\prod_{j=1}^{n_i}[\{(1-p)\phi(y_{ij}|\mu_{0ij},\sigma^2)\}^{1-z_i}+\{p\phi(y_{ij}|\mu_{1ij},\sigma^2)\}^{zi}] \quad (3)$$

In order to carry out the Bayesian analysis, next we describe the prior models for the various parameters. Let $\eta=(\mu_a,\mu_{b_1},\mu_{b_2},\mu_r,\sigma_a^2,\sigma_{b_1}^2,\sigma_{b_2}^2,\sigma_r^2)'$ denote the vector of hyperparameters characterizing the $\log_2$ PSA trajectory. The mixture parameter p is also thought to be a hyperparameter characterizing the distribution of $z_i's$. We selected non-informative priors for these hyperparameters, specifically we chose $\mu_a \sim N(-2, 100)$, $\mu_{b_1} \sim N(0.05,100)$, $\mu_{b_2} \sim N(0.1,100)$, $\mu_r \sim N(0,10000)$, and non-informative Jeffrey's prior for p ($p \sim Beta(0.5,0.5)$). We further assumed that $\sigma_a^2,\sigma_r^2,\sigma_{b_1}^2,\sigma_{b_2}^2,\sigma^2 \sim IG(0.001, 1000)$ where $IG(\cdot,\cdot)$ denotes the Inverse Gamma distribution. Additionally, all group of parameters described above are assumed to be stochastically independent.

Elicitation of the prior parameters has remained an inherent part of the prior selection process. Since a strong prior information can drive the direction of analysis in a significant manner, it is imperative that reasonably accurate information on the prior parameters be obtained wherever possible. In our context, one naive but often useful prior elicitation method may consist of extracting summary information (e.g. mean, variance) on rate of PSA rise, and transition time from historical data or previous case studies and then translating the information into parameter values for the prior distributions. In situations where such information is unavailable, one acceptable strategy is to make the prior information sufficiently “diffused” (by assuming a large variance) relative to the likelihood, so that the prior has minimal impact on the posterior distribution. This is the strategy we adopted in our prior selection process.

Regression Models

The model in (1) can be extended to include covariates. In our analyses, we considered the following three models:

Model 1. Uses the model specification in (1). No covariate adjustment is considered.

Model 2. This model incorporates race as a covariate. Specifically, the parameters corresponding to transition time, $\log_2$PSA value at the transition time, $\log_2$PSA rates of change before and after the transition time, and mixture probability were allowed to vary by race.

Model 3. This model additionally adjusts for cancer stage at diagnosis, and pre-operative PSA. Specifically, the parameters corresponding to transition time, $\log_2$PSA value at the transition time, and $\log_2$PSA rates of change before and after the transition time were allowed to vary by stage and pre-operative PSA.

The regression coefficients are typically modeled as independent Normal or *t* variables, unless there is apriori reason to envision them as skewed. In the latter case skew-normal or skew-t are reasonable prior choices for the regression parameters. We chose non-informative normal priors for the regression coefficients in the above models.

3.1. Posterior Analysis

The simple conjugate form for all the parameters allowed us to write the full conditional posterior distribution in closed form (Hall et al., 2003). We adopted the simulation-based Markov-chain Monte Carlo (MCMC) technique that relies on generating random draws from the relevant full conditional distributions in an iterative manner. For the mixture model, we obtained draws from the model parameters given the indicator, z_i. Given the indicator z_i, the mixture model reduces to a regular hierachical model.

3.1.1. Model Selection—Formal model selection was carried out using the *Deviance Information Criterion* popularized by Spiegelhalter et al. (2002). The deviance $D(\theta) = -2 \log p(\mathbf{y}|\theta)$ is used as a general measure of goodness of fit. It is natural to estimate the expected deviance function $E_{\theta|\mathbf{y}}(D(\theta))$, evaluated under the posterior distribution and consider it as a Bayesian goodness of fit diagnostic. In a Bayesian numerical integration scheme with direct generation of posterior observations, this quantity can be estimated by the average deviances over the posterior realizations $\{\theta^{(l)} : l = 1, \cdots, L\}$, expressed as,

$$\overline{D}=\frac{1}{L}\sum_{l=1}^{L}D(\theta^{(l)})=\frac{1}{L}\sum_{l=1}^{L}\{-2\log p(\mathbf{y}|\theta^{(l)})\}.$$

The effective number of parameters (p_D) in a model can be thought of as $\overline{D} - D(\hat{\theta})$, where $\hat{\theta}$ is the posterior mean, $E[\theta|\mathbf{y}]$. A corresponding *Deviance Information Criterion* (DIC) propagated by Spiegelhalter et al. (2002) based on this construction is DIC = $\overline{D} + p_D$. DIC can be easily computed in a MCMC setting, in terms of the readily available posterior quantities. In comparing across models, smaller values of DIC are preferred.

In developing DIC, Spiegelhalter et al. (2002) mostly focused on the case of generalized linear models, although they concluded their seminal paper with a discussion of the possibilities of extending this notion to models like mixtures of distributions. The ensuing discussion of their article pointed out the possible difficulties of defining DIC precisely in these scenarios. In particular, Celeux et al. (2000) and DeIorio and Robert (2002) described

some possible inconsistencies in the definition of a DIC for mixture models. For instance, in the mixture case, the above definition of DIC quite often leads to a negative value of p_D. Celeux et al. (2006) proposed alternative definitions of DIC for mixture models. Unlike the mixture model case in Celeux et al. (2006)'s paper, the $\mu_{z_i ij}$ in our study is a function of z_i, the transition time, $\log_2$PSA at the transition time, $\log_2$PSA rates of change before and after the transition time, and times t_{ij} for subject i.

In this study, we consider three different DIC constructions. The first is based on Spiegelhalter et al.'s approach (we call this DIC_0). Additionally, we consider adaptations of Celeux et al. (2006) that rely on the observed (DIC_1) and complete (DIC_2) likelihoods in our setting. For computing DIC_0,

$$\overline{D}=-\frac{2}{L}\sum_{l=1}^{L}\log\phi(y|\boldsymbol{\theta}^{(l)},\boldsymbol{z}^{(l)})=-\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{n}\sum_{j=1}^{n_i}\{(1-z_i^{(l)})\log\{\phi(y_{ij}|\mu_{0ij}^{(l)},\sigma^{2(l)})\}+z_i^{(l)}\log\{\phi(y_{ij}|\mu_{1ij}^{(l)},\sigma^{2(l)})\}\}$$

where $\mu_{1ij}^{(l)}=a_i^{(l)}+b_{1i,1}^{(l)}(t_{ij}-r_{i,1}^{(l)})_-+b_{2i}^{(l)}(t_{ij}-r_{i,1}^{(l)})_+$, $\mu_{0ij}^{(l)}=a_i^{(l)}+b_{2i}^{(l)}(t_{ij}-1)$, and $z_i^{(l)}, a_i^{(l)}, b_{1i,1}^{(l)}, b_{2i}^{(l)}, r_{i,1}^{(l)}$ and $\sigma^{2(l)}$ the results from the lth MCMC iteration.

We define $D(\hat{\boldsymbol{\theta}})$ using the posterior mode of the unobserved indicator z_i. Since z_i is either 1 or 0, the posterior mode is 1 if $\mathbf{1}_{\{\bar{z}_i\geq 0.5\}}$, and $D(\hat{\boldsymbol{\theta}})$ is computed as

$$D(\hat{\boldsymbol{\theta}})=-2\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\{\mathbf{1}_{\{\bar{z}_i<0.5\}}\phi(y_{ij}|\bar{\mu}_{0ij},\bar{\sigma}^2)+\mathbf{1}_{\{\bar{z}_i\geq 0.5\}}\phi(y_{ij}|\bar{\mu}_{1ij},\bar{\sigma}^2)\}$$

where $\bar{\mu}_{1ij} = \bar{a}_i + \bar{b}_{1i,1}(t_{ij} - \bar{r}_{i,1})_- + \bar{b}_{2i}(t_{ij} - \bar{r}_{i,1})_+$, $\bar{\mu}_{0ij} = \bar{a}_i + \bar{b}_{2i}(t_{ij}-1)$, and $\bar{a}_i$, $\bar{b}_{1i,1}$, $\bar{b}_{2i}$, $\bar{r}_{i,1}$, $\bar{z}_i$ and $\bar{\sigma}^2$ are the posterior means for the corresponding parameters.

For computing DIC_1, we rely on the observed likelihood, $l(\mathbf{y}|\bar{\boldsymbol{\theta}})$, under the assumption that it can be computed in closed form. Note that $\bar{D}$ is clearly and uniquely defined, even though it requires MCMC simulations to be computed approximately, and

$$\overline{D}=-\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\{(1-p_i^{(l)})\phi(y_{ij}|\mu_{0ij}^{(l)},\sigma^{2(l)})+p_i^{(l)}\phi(y_{ij}|\mu_{1ij}^{(l)},\sigma^{2(l)})\}.$$

In general for mixture models, the definition of $D(\hat{\boldsymbol{\theta}})$ may not be unique because $\boldsymbol{\theta}$ is not always identifiable and the posterior mean can then be a very poor estimator. However, in our case, the subject-specific trajectories uniquely identify z_i and $D(\hat{\boldsymbol{\theta}})$ is obtained as

$$D(\hat{\boldsymbol{\theta}}) = -2\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\{(1-\overline{p}_i)\phi(y_{ij}|\overline{\mu}_{0ij},\overline{\sigma}^2)+\overline{p}_i\phi(y_{ij}|\overline{\mu}_{1ij},\overline{\sigma}^2)\},$$

where $\bar{p}_i$ is the posterior mean for p_i.

A third version of DIC (DIC_2) is based on using the complete likelihood. Using $l(\mathbf{y},\mathbf{z}|\boldsymbol{\theta})$, we can set $\bar{D}$ as the posterior expected value (over the missing data) of the joint deviance, and $\bar{D}$ can be obtained using

$$\overline{D} = -\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{n}\sum_{j=1}^{n_i}[P(z_i=0|\boldsymbol{\theta},\mathbf{y})\log\{(1-p^{(l)})\phi(y_{ij}|\mu_{0ij}^{(l)},\sigma^{2(l)})\}+P(z_i=1|\boldsymbol{\theta},\mathbf{y})\log\{p^{(l)}\phi(y_{ij}|\mu_{1ij}^{(l)},\sigma^{2(l)})\}].$$

Since the z_i is identified by the entire trajectory of subject i, we use the likelihood of subject i to define $P(z_i = 1 |\boldsymbol{\theta}, \mathbf{y})$ such that

$$P(z_i=1|\boldsymbol{\theta},y) = \frac{p_i^{(l)}\prod_{j=1}^{n_i}\phi(y_{ij}|\mu_{1ij}^{(l)},\sigma^{2(l)})}{p_i^{(l)}\prod_{j=1}^{n_i}\phi(y_{ij}|\mu_{1ij}^{(l)},\sigma^{2(l)})+(1-p_i^{(l)})\prod_{j=1}^{n_i}\phi(y_{ij}|\mu_{0ij}^{(l)},\sigma^{2(l)})}.$$

In addition to the difficulty of choosing $\hat{\boldsymbol{\theta}}$ as mentioned above, we now have the problem of defining the fixed point deviance, $D(\hat{\boldsymbol{\theta}})$, in connection with the missing data structure. Using similar motivations as in Celeux et al. (2006), we first define a complete data DIC, by defining the complete data estimator $E[\boldsymbol{\theta}|\mathbf{y},\mathbf{z}]$, and then integrate this quantity to obtain DIC. This requires the computation of a posterior expectation for each value of $\mathbf{z}$. Specifically then, for each lth MCMC iteration, we obtain $\mathbf{z}^{(l)}$, calculate the posterior mean $\bar{\boldsymbol{\theta}}^{(l)} = E[\boldsymbol{\theta}|\mathbf{y}, \mathbf{z}^{(l)}]$, and define $\widehat{D(\boldsymbol{\theta})}$ as

$$D(\hat{\boldsymbol{\theta}}) = -\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{n}\sum_{j=1}^{n_i}[(1-z_i^{(l)})\log\{1-\overline{p}_i^{(l)})\phi(y_{ij}|\overline{\mu}_{0ij}^{(l)},\overline{\sigma}^{2(l)})\}+z_i^{(l)}\log\{\overline{p}_i^{(l)}\phi(y_{ij}|\overline{\mu}_{1ij}^{(l)},\overline{\sigma}^{2(l)})\}]$$

where $\overline{p}_i^{(l)}, \overline{\mu}_{0ij}^{(l)}, \overline{\mu}_{1ij}^{(l)}, \overline{\sigma}^{2(l)}$ are the posterior means given $(\mathbf{y}, \mathbf{z}^{(l)})$.

4. Data Analysis

As described in section 2.1, the analytic cohort comprised of PSA trajectories of 94 patients who met the inclusion criteria. We considered as covariates race (African American vs. Caucasian), stage at diagnosis (local or regional), and serum PSA (in ng/ml) prior to surgery (continuous variable), all of which are established predictors of prostate cancer recurrence. The analysis was carried out in WinBUGS version 1.4.1 (Lunn et al., 2000). The choice of priors (described in section 3) was fairly non-informative for the mean as well as the variance parameters. We found that the choice of hyperparameters in the prior distributions

had little influence on the marginal posterior distributions. We ran two independent parallel chains with different starting values. The chains were run with a burn in of 20000 iterations, and additional 70000 iterations for inference. With a thinning interval of 50, posterior distributions for the parameters were based on 2000 iterations. Point estimates of parameters were calculated using the mean and median of the posterior distribution; we also obtained 95% credible intervals. The convergence of the Markov chain was assessed visually by track plots and autocorrelation plots, and by the Gelman-Rubin Diagnostic criterion.

4.1. Results

Table 1 presents the posterior summary statistics for the model parameters. Results from model 2 describe the unadjusted effect of race, and that from model 3 the adjusted (*i.e*. stage and pre-operative PSA adjusted) effect of race on PSA trajectories. Based on both models 2 and 3, race does not significantly affect PSA trajectory. During the latent phase, PSAs remain virtually unchanged for both AA and CC. The average transition time for AA is 18.1 months after surgery (95% credible interval: 12.5 to 23.6 months) and that for CC is 18.6 months (95% credible interval: 15.0 to 22.2 months). The average rate of rise in $\log_2$ PSA in the rapid phase is also similar for both groups. Posterior distributions of all the parameters turned out to be fairly symmetric.

Table 2 presents the DIC values for model comparison. In computing DIC_0, an alternative to p_D for measuring model complexity is half the posterior variance of the deviance (p_V), which can be obtained based on posterior simulations as $p_V = \frac{1}{2}\frac{1}{L-1}\sum_{l=1}^{L}\left(D(\boldsymbol{\theta}^{(l)}) - \overline{D}\right)^2$. Gelman et al. (2004) recommends using p_V based on robustness and accuracy considerations. The current version of winBUGS also uses p_V for computing DIC in the case of mixture models. For the sake of completeness, in addition to DIC_1 and DIC_2, we therefore present in Table 2 two versions of DIC_0 using p_D and p_V. Both $DIC_0(p_V)$ and DIC_1 favor model 1 over the other two models, but DIC_2 favors model 2. $DIC_0(p_d)$, in contrast picks model 3 as the best. The corresponding p_D values associated with all three models are negative (specifically, −80, − 22, and −67, for models 1, 2, and 3, respectively). Thus $DIC_0(p_D)$ may not be a good choice here.

Our model also allowed us to obtain posterior distribution of the indicator variable, $z_i (i = 1, \cdots, n)$. Subjects with large posterior means are more likely to be classified as having a latent phase. We used the cutoff of 0.8 to classify subjects as having a latent phase or not; *i.e*. if the posterior mean of z_i is larger than 0.8, subject i is said to have a latent phase. Otherwise, subject i is said to have rapid PSA rise immediately after surgery. Figure 2 shows the subject-specific predicted $\log_2$PSA profiles based on model 1 using the threshold of 0.8 for twelve randomly selected patients. The predicted trajectories fit the observed data very well for both groups of patients who may or may not have a latent phase. It is evident though that most patients experience a latent phase (point estimate of p based on model 1 is 90%, 95% credible interval: 80-100%).

5. Concluding Remarks

In this article, we have proposed a Bayesian piecewise mixture model to analyze PSA trajectories after surgery for prostate cancer. We believe this model can also be used after other treatments, such as radiotherapy, because PSA levels after radiotherapy typically decrease and then start to increase again at variable times and rates across individuals. Our model allows us to account for between as well as within-subject variabilities. In addition, it allows us to accommodate complex features, such as the presence of a random change point across subjects, as well as mixture of patient populations with and without a latent phase.

To the extent we trust our model specification, our approach offers some flexibility in modelling. In the clinical literature, serial PSA data is often handled by using a simplistic approach that independently estimates individual PSA profiles; population estimates are then obtained by pooling individual estimates using a simple average (D'Amico et al., 2003; Pollack et al., 1994). This approach does not adequately account for the PSA variability. By using the observed change point as the true change point, the variability is completely ignored. Another approach uses a smooth function of time to model the longitudinal PSA values (Ulm, 1991). However, the major drawback of this approach is that the model does not accord with the underlying biological process. A quadratic model allows for a smooth transition, however, it assumes that the PSA pattern is symmetric about the change point, which is in contradiction with the typical PSA profile after surgery. Use of a higher degree polynomial is also biologically implausible, since post surgery PSA values are typically characterized by a single inflection point prior to giving any additional therapy. Our model also allows the flexibility to accommodate patients for whom surgery might have been useless (cancer spread outside the prostate gland that can only be confirmed pathologically). These patients may experience a rapid increase in PSA immediately following surgery, without going through a latent phase. The two-component mixture model proposed in this paper allows us to accommodate patients both with and without a latent phase, while estimating the change point for patients experiencing a latent phase, within the same modelling framework. Patients with a latent phase are allowed to have different PSA trajectories, represented by a subject-specific rate of PSA rise before the change point, a subject-specific change point, and a subject-specific rate of rise after the change point. The Bayesian paradigm provides a unified framework for carrying out estimation and predictive inference. Further, it lends itself readily to straightforward model selection procedures. The possibility of entertaining a flexible modelling environment and the ease of implementation make this, in our view, an appealing approach to analyze serial PSA data.

References

American Cancer Society. Cancer Facts and Figures 2010. Atlanta: American Cancer Society; 2010.

Bellera CA, Hanley JA, Joseph L, Albertsen PC. Hierarchical change point models for biochemical markers illustrated by tracking postradiotherapy prostate-specific antigen series in men with prostate cancer. Ann Epidemiol. 2008; 18:270–82. [PubMed: 18374279]

Celeux G, Forbes F, Robert CP, Titterington DM. Deviance information criteria for missing data models. Bayesian Analysis. 2006:651674.

Celeux G, Hurn M, Robert CR. Computational and Inferential Difficulties With Mixture Posterior Distributions. Journal of the American Statistical Association. 2000; 95:957–970.

D'Amico A, Cote K, Loffredo, et al. Determinants of Prostate Cancer Specific Survival following Therapy during the Prostate Specific Antigen Era. Journal of Urology. 2003; 170:42–47. [PubMed: 12796640]

DeIorio M, Robert CP. Discussion of Spiegelhalter et al. Journal of the Royal Statistical Society Series B. 2002; 64:629630.

Freedland SJ, Presti JC Jr, Amling CL, Kane CJ, Aronson WJ, Dorey F, Terris MK. Time trends in biochemical recurrence after radical prostatectomy: results of the SEARCH database. Urology. 2003; 61:736–41. [PubMed: 12670557]

Gelman, A.; Carlin, JB.; Stern, HS.; Rubin, DB. Bayesian Data Analysis. second. Chapman and Hall; New York: 2004.

Hall CB, Ying J, Kuo L, Lipton RB. Bayesian and profile likelihood changepoint methods for modeling cognitive function over time. Computational statistics and data analysis. 2003; 42:91–109.

Kiuchi AS, Hartigan JA, Holford TR, Rubinstein P, Stevens CE. Change Points in the Series of T4 Counts Prior to AIDS. Biometrics. 1995; 51:236–48. [PubMed: 7766779]

Lu-Yao GL, McLerran D, Wasson J, Wennberg JE. An assessment of radical prostatectomy: time trends, geographic variation and outcomes. Journal of the American Medical Association. 1993; 269:26332636.

Lu-Yao GL, Potosky AL, Albertsen PC, Wasson JH, Barry MJ, Wennberg JE. Follow-up prostate cancer treatments after radical prostatectomy: a population based study. Journal of the National Cancer Institute. 1996; 88:166173.

Lunn DJ, Thomas A, Best N, Spiegelhalter D. *WinBUGS* – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing. 2000; 10:325–337.

Moul JW, Douglas TH, McCarthy WF, McLeod DG. Black race is an adverse prognostic factor for prostate cancer recurrence following radical prostatectomy in an equal access health care setting. Journal of Urology. 1996; 155:1667–1673. [PubMed: 8627850]

Pollack A, Zagars G, Kavadi V. Prostate Specific Antigen Doubling Time and Disease Relapse after Radiotherapy for Prostate Cancer. Cancer. 1994; 74:670–678. [PubMed: 7518341]

Powell IJ, Banerjee M, Sakr W, et al. Should African-American men be tested for prostate carcinoma at an earlier age than white men? Cancer. 1999; 85:472477.

Slate E, Lark L. Using PSA to detect prostate cancer onset: an application of Bayesian retrospective and prospective changepoint identification. Journal of Educational and Behavioral Statistics. 2001; 26:443–468.

Spiegelhalter DJ, Best NG, Carlin BP, Van de Linde A. Bayesian measures of model complexity and fit (with discussion). Journal of the Royal Statistical Society. 2002; 64:583–639.

Ulm K. A Statistical Method for Assessing a Threshold in Epidemiological Studies. Statistics in Medicine. 1991; 10:341–349. [PubMed: 2028118]

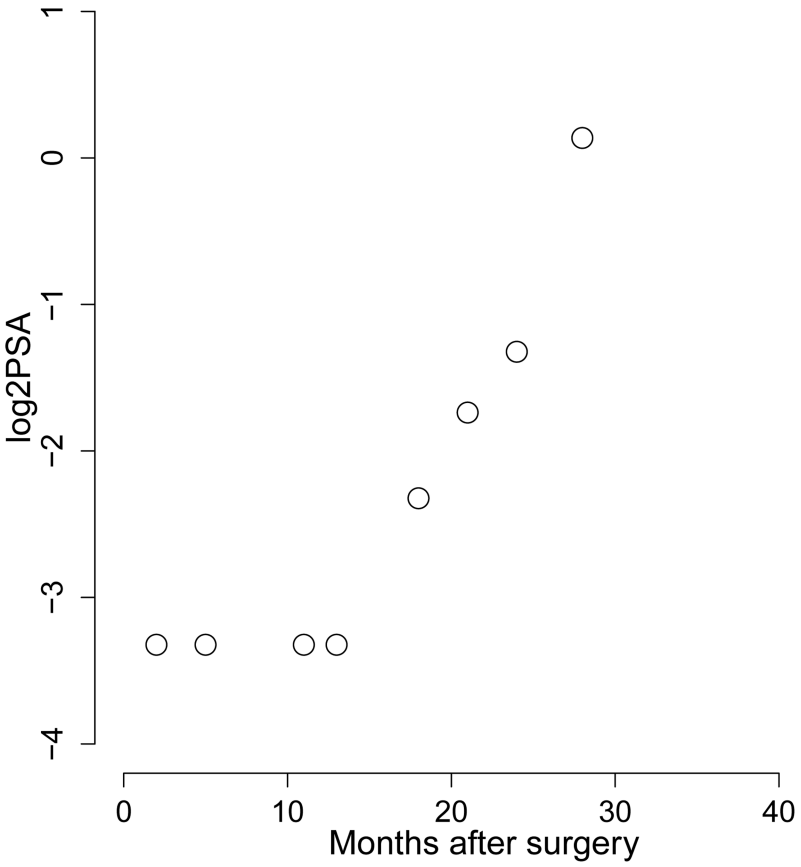

Figure 1.

Circles indicate observed $\log_2$PSA values for an individual. The horizontal axis represents time (in months) after surgery.

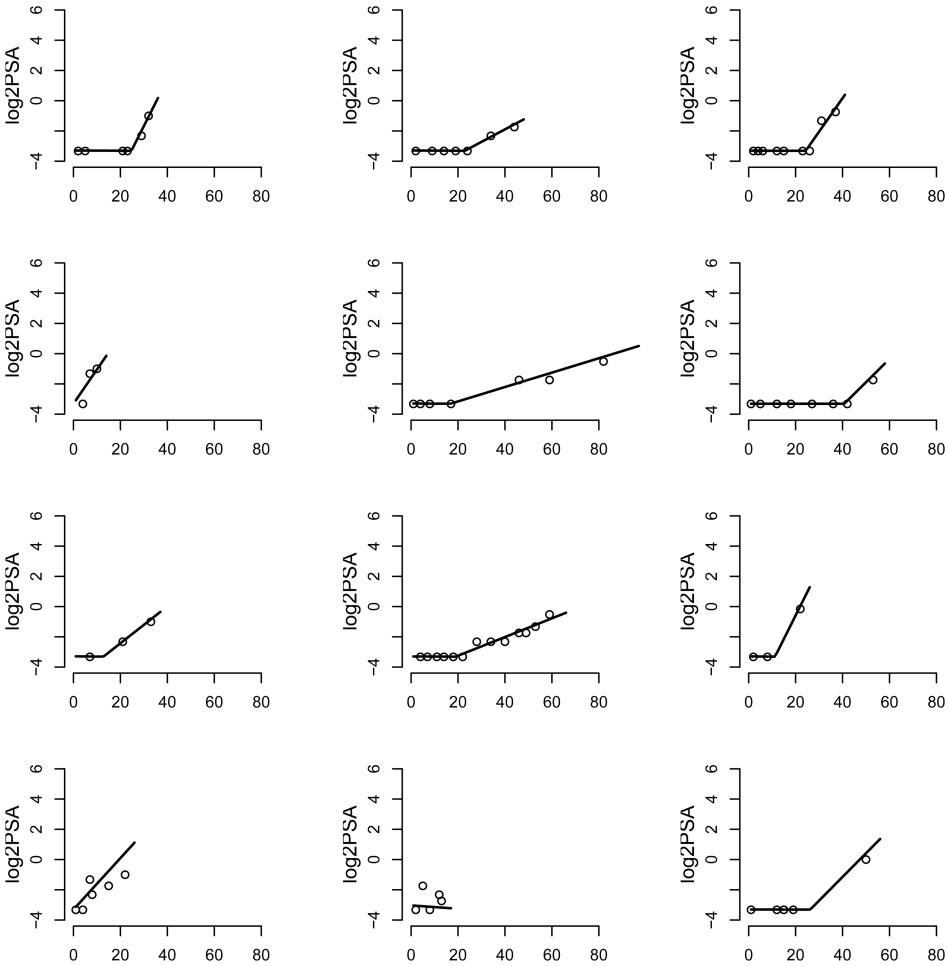

Figure 2.

Predicted $\log_2$PSA profiles (solid line) based on model 1 of twelve randomly selected subjects. Circles indicate observed $\log_2$PSA values. The horizontal axis represents time (in months) after surgery.

Table 1

Posterior mean, median, and credible intervals for the model parameters

	Model 1		Model 2		Model 3	
	Mean(Median)	**95% CI**	**Mean(Median)**	**95% CI**	**Mean(Median)**	**95% CI**
p	0.9(0.9)	(0.8, 1.0)				
AA			0.9(0.9)	(0.7,1.0)	0.9(0.9)	(0.7,1.0)
CC			1.0(1.0)	(0.9,1.0)	1.0(1.0)	(0.9,1.0)
AA vs. CC			0.1(0.1)	(-0.1,0.3)	0.1(0.1)	(-0.1,0.3)
μ_a	-3.3(-3.3)	(-3.4, -3.2)				
AA			-3.3(-3.3)	(-3.4,-3.2)	-3.4(-3.4)	(-3.6,-3.2)
CC			-3.3(-3.3)	(-3.4,-3.2)	-3.4 (-3.4)	(-3.5,-3.2)
AA vs. CC			-0.01(-0.0)	(-0.16,0.14)	0.01 (0.01)	(-0.14,0.17)
stage					0.006(0.004)	(-0.15,0.17)
pre-PSA					0.003(0.003)	(-0.004,0.011)
μ_{b1}	-0.001(-0.001)	(-0.008, 0.005)				
AA			0.0(0.0)	(-0.011,0.011)	-0.003(-0.003)	(-0.017,0.012)
CC			-0.001(-0.001)	(-0.009,0.006)	-0.003(-0.003)	(-0.015,0.008)
AA vs. CC			-0.001(-0.001)	(-0.015,0.011)	-0.001(-0.001)	(-0.013,0.013)
stage					0.002(0.002)	(-0.01,0.01)
pre-PSA					0.0(0.0)	(-0.001,0.001)
μ_{b2}	0.20(0.21)	(0.17, 0.24)				
AA			0.19(0.20)	(0.13,0.26)	0.18(0.18)	(0.10,0.28)
CC			0.21(0.21)	(0.16,0.25)	0.20(0.20)	(0.14,0.27)
AA vs. CC			0.02(0.02)	(-0.06,0.09)	0.02(0.02)	(-0.06,0.10)
stage					-0.04(-0.04)	(-0.11,0.03)
pre-PSA					0.002(0.002)	(-0.002,0.005)
μ_r	18.5(19.0)	(15.6, 21.5)				
AA			18.1(18.7)	(12.5,23.6)	19.8(19.8)	(12.4,27.3)
CC			18.6(19.1)	(15.0,22.2)	20.1(20.2)	(14.0,26.0)
AA vs. CC			0.5(0.5)	(-6.1,7.1)	0.3(0.3)	(-6.3,7.0)
stage					-2.6(-2.6)	(-8.6,3.4)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

	Model 1		Model 2		Model 3	
	Mean(Median)	**95% CI**	**Mean(Median)**	**95% CI**	**Mean(Median)**	**95% CI**
pre-PSA					0.02(0.01)	(-0.3,0.3)

Table 2

DIC for model comparison

	$DIC_0(p_D)$	$DIC_0(p_V)$	DIC_1	DIC_2
Model 1	332	765	474	653
Model 2	290	783	480	642
Model 3	251	785	481	654