Published in final edited form as:
J Appl Stat. 2011 March ; 38(2): 591–603. doi:10.1080/02664760903521476.

Bayesian Parametric Accelerated Failure Time Spatial Model and its Application to Prostate Cancer

Jiajia Zhang[a,*] and **Andrew B. Lawson**[b]

[a]Department of Epidemiology and Biostatistics University of South Carolina, Columbia, SC 29208, USA

[b]Division of Biostatistics and Epidemiology Medical University of South Carolina, Charleston, SC 29425, USA

Abstract

Prostate cancer is the most common cancer diagnosed in American men and the second leading cause of death from malignancies. There are large geographical variation and racial disparities existing in the survival rate of prostate cancer. Much work on the spatial survival model is based on the proportional hazards model, but few focused on the accelerated failure time model. In this paper, we investigate the prostate cancer data of Louisiana from the SEER program and the violation of the proportional hazards assumption suggests the spatial survival model based on the accelerated failure time model is more appropriate for this data set. To account for the possible extra-variation, we consider spatially-referenced independent or dependent spatial structures. The deviance information criterion (DIC) is used to select a best fitting model within the Bayesian frame work. The results from our study indicate that age, race, stage and geographical distribution are significant in evaluating prostate cancer survival.

Keywords

Spatial; Accelerated failure time model; Deviance information criterion (DIC); Bayesian; Likelihood

1. Introduction

In public health and population-based biomedical studies, data are often collected by geographic regions, such as the district or postal code of the residence of individuals. Often the adjacent neighborhoods may be more alike than those from distant region due to similar environmental and social factors. Failing to account for the correlation within neighborhoods may lead to biased statistical inference.

Recently, there has been lots of attention paid to the analysis of geographical patterns of survival times, in addition to the impact of other covariates. For example, Henderson et al. [10] modeled spatial variation in survival of acute myeloid leukemia patients in northwest England; Banerjee et al. [2] applied a spatial frailty model to infant mortality in Minnesota by assuming a parametric Weibull baseline hazard and geostatistical or Gaussian Markov random field priors for the spatial component; Li and Ryan [17] analyzed the effect of risk factors on the onset of childhood asthma with spatial data from the East Boston Asthma

*Corresponding author. jzhang@mailbox.sc.edu .

Study; and Hennerfeind et al. [12] applied a geoadditive survival model to data on waiting times for coronary artery bypass grafting. However, all of these examples focus on the proportional hazards (PH) model or its extensions which measure the spatial effects on the hazard scale. For example, Henderson et al. [10] used the PH model because the initial survival analysis indicates that the PH assumption was satisfied in their data set. There are some discussions about the spatial survival models based on the non-proportional cases, such as the multivariate adaptive regression spline (MARS) model [19], the additive hazard survival model with frailty [20], dynamic survival models with spatial frailty [3], and the normal transformation model for spatial correlated data [16]. In the normal transformation model, the survival outcome marginally follows a PH model, and in their discussions, they proposed an extension to the accelerated failure time (AFT) model. The AFT model is widely accepted as an alternative approach when the PH assumption does not hold, but there are few studies using semiparametric Bayesian analysis in the AFT model [1,8,21]. A semi-Bayesian analysis of the AFT model is given by [5], where they utilized a Dirichlet process to estimate the survival distribution. Walker and Mallick [24] proposed a fully Bayesian approach for the median regression model by the Polya tree prior. Recently, Komarek and his colleagues [14,15] proposed a normal mixture as the error distribution in the AFT model.

In this paper, we propose an AFT spatial model by adding a random effect to the AFT model for investigating risk effects of prostate cancer (PrCA). Prostate cancer diagnosis data from the state of Louisiana from the Surveillance, Epidemiology, and End Results (SEER) program [11] of the National Cancer Institute (NCI) is used as an example. The purpose of this application is to investigate whether PrCA is much more aggressive in African-Americans than in Whites and whether there exists regional environmental difference. The estimation procedure is developed from a Bayesian perspective with parametric assumption. A semiparametric AFT estimation approach [5,7,14,15,24] could be considered for this study, which requires more effort on programming than the parametric AFT model. In order to provide a clear picture of the AFT spatial model and its application, we will analyze the PrCA data using WinBUGS code [18]. Then, the DIC [22] is applied to choose the best fit parametric model. In the appendix, the WinBUGS code for the parametric AFT spatial model or the parametric AFT frailty model is provided.

The remainder of this paper is organized as follows: Section 2 describes the data that motivate this study. The standard survival analysis and possible issues are described in Section 3. Section 4 outlines the AFT spatial model. The estimation procedure is discussed in Section 5. Section 6 illustrates the application of the proposed approach to the PrCA data from Louisiana. Finally, Section 7 summarizes and discusses the results.

2. Motivating Data

Prostate cancer (PrCA) is a major public health problem, which over a lifetime will affect an estimated one in five American men. Since PrCA is the number one incident cancer and the number two cause of cancer deaths among US men, the data in PrCA from the SEER program are particularly important for researchers, clinicians, policy makers, and citizens in understanding this disease. The SEER program has 17 registries, which include San Francisco-Oakland, Connecticut, Detroit, Hawaii, Iowa, New Mexico and Utah for period 1973-2004, Seattle for period 1974-2004, Atlanta for period 1975-2004, Alaska, San Jose-Monterey, Los Angeles and Rural Georgia for period 1992-2004, Great California, Kentucky, Louisiana and New Jersey for period 2000-2004. We extract the PrCA data from the SEER cancer incidence public-use data base. Observations with missing values on race, age, county of residence, stage and marital status at diagnosis are excluded in this analysis. According to patient's medical records, race includes White, Black, or Other, with Black being the designator for African American. In this study we are only interested in the

disparities between white and black, and so we remove other races. Stage of cancer has four categories: local, regional, distant and unstaged. Unstaged means information is not suffcient to assign a stage for the cancer. So, we exclude the unstaged cases.

In order to investigate large geographical variation and racial disparities in the survival rate of PrCA, we need to select the registry with a relatively large proportion of African-American males. After checking all registries, we focused our study on the SEER data set from Louisiana, which has 64 counties and whose ratio of black men is 29.34%. Note, the data from Louisiana can not represent the whole population due to the limit of the observation period, but it does represent the status of the incidence for the five year period 2000-2004 in Louisiana.

The individual-specific information for a patient that is used in this study are: age (age of the patient at diagnosis in complete years), race (White and Black), county (patient's county of residence at the time of diagnosis), stage (SEER summary stage, localized/regional and distant), marital status at diagnosis (single, married and other), and survival time after diagnosis (including censoring time). It is worthwhile pointing out that in the definition of the stage, localized tumors are confined to the prostate gland, regional tumors are spread to contiguous organs or lymph nodes, and distant tumors are spread to remote organs. Clinically localized tumors are frequently upstaged to regional stage after surgery, so there is an extra category (localized/regional) only for PrCA in SEER data. We have 446 observations from localized category, 103 observations from the regional category, and 15 132 observations from the localized/regional category, so we combined both localized and regional staged cancers into the localized/regional category in our data analysis. Table 1 provides a summary of the characteristics of the PrCA patients included in this study.

3. Modeling Issues

Commonly, the PH model and the AFT model are the most popular survival models and the nonparametric Kaplan-Meier (KM) survival curve is used as a rule of thumb to choose between them [13]. After visual inspection, the test based on the Schoenfeld residuals [23] is applied. For illustration purpose, we select nine different counties from the different parts of Louisiana in order to detect the difference between locations. Among these, Caddo, Bossier and Webster are in the northwest; Sabine, Grant and Avoyelles are in the midwest; Calcasieu, Vermilion and Acadia are in the southwest. In each county, we fitted the KM survival curves for white and black respectively (Figure 1). In each plot, the y-axis presents the survival probability for white or black and x-axis denotes the time period after the diagnosis of PrCA.

From Figure 1, we find that the survival rate does change markedly with the location and it appears that the middle west part of the state tends to have higher survival rates than the other two parts. So, considering spatial effects in survival models should improve the estimates of risk effects.

The Schoenfeld residual test is used to check the PH assumption. If the P-value from the Schoenfeld residual test is less than significant level (such as, 0.05), it indicates that the PH assumption is not satisfied. Investigating the survival curve with respect to different races, we find that some of the KM survival curves cross over in Figure 1, such as in county Webster, Avoyelles and Acadia. Through the Schoenfeld residual test, we find that the P-value are significantly less than 0.05 in county Bossier (P-value=0.0261), Avoyelles (P-value= 0.043), Vermilion (P-value 0.00892), and Acadia (P-value=0.0428). The P-value is not significant enough in county Webster (P-value= 0.103). Thus, we doubt the accuracy of the PH assumption and consider the AFT spatial model for this data set.

4. Accelerated Failure Time Spatial Model

Let T_{ij} denote the survival time after diagnosis for patient j in county i, and x_{ij} denote possible risk effects corresponding to T_{ij}, where $j = 1, \ldots, n_i$, $i = 1, \ldots, n$. The AFT model can be expressed as:

$$\log\left(T_{ij}\right) = \mu + \beta \mathrm{x}_{ij} + \sigma\varepsilon_{ij} \quad ,$$

where β is the unknown coefficient, ε_{ij}'s are independent random errors, μ and σ are the shape parameter and scale parameter. Letting $n_i = 1$, we obtain the regular AFT model.

The spatial structure can be considered by adding a random effect to the AFT model and the AFT spatial model is specified as:

$$\log\left(T_{ij}\right) = \mu + \beta \mathrm{x}_{ij} + W_i + \sigma\varepsilon_{ij} \quad , \tag{1}$$

where W_i's are spatial random effects. The advantage of the AFT spatial model is that the interpretation of risk/spatial effects on the failure time are easy since the AFT spatial model simply regresses the logarithm of the survival time over covariates and random spatial effects. In this paper, we consider county specific random effects.

Let $f(\cdot)$ denote the density function of T and $f_0(\cdot)$ denote the density function of ε. $S(\cdot)$ and $S_0(\cdot)$ denote the survival functions, and $h(\cdot)$ and $h_0(\cdot)$ represent the hazard functions corresponding to $f(\cdot)$ and $f_0(\cdot)$. Then, we have

$$f\left(t_{ij}|W_i\right) = \frac{1}{\sigma t_{ij}} f_0\left(\frac{\log\left(t_{ij}\right) - \lambda\left(\mathrm{x}_{ij}\right)}{\sigma}\right),$$

$$S\left(t_{ij}|W_i\right) = S_0\left(\frac{\log\left(t_{ij}\right) - \lambda\left(\mathrm{x}_{ij}\right)}{\sigma}\right),$$

$$h\left(t_{ij}|W_i\right) = \frac{1}{\sigma t_{ij}} h_0\left(\frac{\log\left(t_{ij}\right) - \lambda\left(\mathrm{x}_{ij}\right)}{\sigma}\right),$$

where $\lambda(\mathrm{x}_{ij}) = \mu + \beta \mathrm{x}_{ij} + W_i$. From the relationship between survival functions, we can see that the spatial random effects have a direct effect on the survival probability. Note that the hazard rate keeps changing over time even when the spatial random effect is fixed in the AFT spatial model, while it stays at the same rate given the specific region in the PH spatial model. For some data set, we believe it is more reasonable to assume the hazard rate changes over time even in the same location.

It is common to assume that $S_0(\cdot)$ comes from the standard normal distribution, the standard extreme value distribution, or the logistic distribution. The $S_0(\cdot)$ expressions and their corresponding $S(\cdot)$'s are summarized in Table 2, where $\varphi(\cdot)$ denotes the cumulative density

function from the standard normal distribution. Corresponding to the distribution of ε, the survival distribution of T follows the lognormal distribution, Weibull distribution or loglogistic distribution.

We consider survival data (t_{ij}, δ_{ij}, $\mathbf{x}_{ij}$), where δ_{ij} is the censoring indicator. We assume that the censoring is independent and noninformative. Let $\mathbf{W} = (W_1, \ldots, W_n)$, $\mathbf{O}$ denote the observations and $\phi = \{\mu, \sigma, \beta\}$ denote the parameters to be estimated. Given the spatial random effect, the likelihood function can be written as:

$$\begin{aligned} L(t|\phi, \mathbf{W}) &= \prod_{i=1}^{n}\prod_{j=1}^{n_i} f(t_{ij})^{\delta_{ij}} S(t_{ij})^{1-\delta_{ij}}, \\ &= \prod_{i=1}^{n}\prod_{j=1}^{n_i} \left[\frac{1}{\sigma t_{ij}} f_0\left(\frac{\log(t_{ij}) - \lambda(\mathbf{x}_{ij})}{\sigma}\right)\right]^{\delta_{ij}} S_0\left(\frac{\log(t_{ij}) - \lambda(\mathbf{x}_{ij})}{\sigma}\right)^{1-\delta_{ij}}. \end{aligned} \tag{2}$$

The spatial random errors can be correlated or not among counties. We refer to mutually uncorrelated county-specific effects as spatially uncorrelated heterogeneity and model this situation with independent Gaussian distributions defined as $W_i \sim \text{Normal}\left(0, v_s^2\right)$, where v_s^2 denotes the variance of spatial random effect. It is worthwhile pointing out that the AFT spatial model under the independent correlation is similar to the AFT frailty model with normal random effects, albeit with frailty effects at the county level rather than individual level. In the correlated situation, we can consider the conditional autoregressive (CAR) model. The CAR model, first introduced by Besag et al. [4], is widely used not only for smoothing in image processing but also in disease mapping. This formulation permits correlation among the random effects according to a neighborhood structure:

$$W_i|\left(W_k, v_s^2\right) \sim \text{Normal}\left(\bar{W}_i, \sigma_i^2\right),$$

where

$$\bar{W}_i = \frac{\sum_k W_k g_{ik}}{\sum_k g_{ik}}$$

$$\sigma_i^2 = v_s^2 / \sum_k g_{ik}$$

$$g_{ik} = 1 \quad (\text{if region } i \text{ and } k \text{ and adjacent } i \neq k); 0 \quad (\text{otherwise})$$

For the model's identifiability, it is common to assume that $\sum_i W_i = 0$. From the specification of the CAR distribution, it can be seen that in the ith region $\bar{W}_i$ depends on the corresponding values in their neighborhood regions and the number of neighborhoods in the ith region, hence exhibiting spatial correlation. We call W_i with the CAR model prior specification the spatially correlated heterogeneity.

We also include both spatially correlated and uncorrelated random effects in a single model to permit a trade-off between independence and a purely local spatially structured dependence of the random effects [4], that is $W_i = W_{i1} + W_{i2}$, where

$$W_{i1}|\left(W_{k1}, v_{s1}^2\right) \sim \text{Normal}\left(\bar{W}_{i1}, \sigma_i^2\right),$$

where

$$\bar{W}_{i1} = \frac{\sum_k W_{k1} g_{ik}}{\sum_k g_{ik}}$$

$$\sigma_i^2 = v_{s1}^2 / \sum_k g_{ik}$$

$$g_{ik} = 1 \quad (\text{if region} \quad i \quad \text{and} \quad k \quad \text{are adjacent} \quad i \neq k); 0 \quad (\text{otherwise})$$

and

$$W_{i2} \sim \text{Normal}\left(0, v_{s2}^2\right).$$

This combined spatially correlated and uncorrelated heterogeneity is called a convolution prior, and v_{s1}^2 and v_{s2}^2 are used to control the variability of the correlated and uncorrelated heterogeneity separately. Apparently, the convolution prior maintains the correlation between adjacent counties, but the correlation is weakened by the uncorrelated structure.

5. Estimation Procedure

The parameter estimates for the AFT spatial model in this study are obtained by posterior sampling based on a Markov chain Monte Carlo (McMC) simulation method. Let $p(\phi)$ denote the prior distribution for ϕ and $p(v_s)$ denote the prior for the variance of the spatial random effects. The posterior distribution can be expressed as

$$p(\phi, \mathbf{W}, v_s|\mathbf{t}) \propto L(\mathbf{t}|\phi, \mathbf{W})\, p(\mathbf{W}|v_s)\, p(\phi)\, p(v_s). \tag{3}$$

To conduct data analysis from the Bayesian perspective, we must specify the prior distributions for each parameter in the model. Because we have little prior information for all the parameters to be estimated, we want our data information to dominate the prior distribution by assuming reasonably non-informative priors for all parameters in this model. For all regression coefficients β and the shape parameter μ, we assume independent vague normal priors with mean 0 and variance 1×10^6. The scale parameter σ in the model is given non-informative priors by gamma distribution with shape parameter 1 and scale parameter 0.001 (with mean 1000, variance 1×10^6). For parameters v_s^2, v_{s1}^2 and v_{s2}^2, which control the variability of the spatial random effects, we assign the vague proper gamma prior

distribution with shape parameter 0.001 and scale parameter 0.001 for their reciprocals (precision parameters for the random effects). Posterior sampling of the AFT spatial model can proceed from the definition of the posterior distribution in (3).

Given the likelihood function Eq (2), the posterior distribution can be factored into different components:

- for $\sigma|\mathbf{W},\beta : \propto \prod_{i=1}^{n}\prod_{j=1}^{n_i} f\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{\delta_{ij}} S\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{1-\delta_{ij}} P(\sigma)$
- for $\mathbf{W}| \quad \sigma,\beta,v_s : \propto \prod_{i=1}^{n}\prod_{j=1}^{n_i} f\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{\delta_{ij}} S\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{1-\delta_{ij}} P(\mathbf{W}|v_s)$
- for $\beta|\mathbf{W},\mu,\sigma : \propto \prod_{i=1}^{n}\prod_{j=1}^{n_i} f\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{\delta_{ij}} S\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{1-\delta_{ij}} P(\beta)$
- for $\mu|\beta,\mathbf{W},\sigma : \propto \prod_{i=1}^{n}\prod_{j=1}^{n_i} f\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{\delta_{ij}} S\left(t_{ij}|\mathbf{W},\phi,v_s\right)^{1-\delta_{ij}} P(\mu)$

where $P(\sigma) \sim Gamma(a, b)$, $P(\mathbf{W}|v_s)$ where W_i is defined conditionally as $W_i|v_s^2 \sim \text{Normal}\left(\bar{W}_i, \sigma_i^2\right)$, $\beta_j \sim Normal(0, c)$, and $P(\mu) \sim Normal(0, c)$ where $a = b = 0.001$, $c = 10^6$. To evaluate competing models, we have run each model with the sampler using multiple chains with overdispersed starting points. Trace plots, the Brooks-Gelman-Rubin diagnostic [6], and autocorrelations within chains are used to assess the convergence of the iterations based on the multiple chains.

Remark: WinBUGS is used to run the whole program with a "zero trick" employed since the joint likelihood function can not be expressed directly by a standard density function. This trick allows arbitrary sampling distributions to be used, and is particularly suitable when, say, dealing with truncated distributions.

6. Real Data Analysis

The risk effects we consider in the AFT or AFT spatial model include the e ect of race, marital status, age and stage, which are very common in PrCA analysis.

We will assume ε follows the standard normal distribution, the standard extreme value distribution and the logistic distribution in both the AFT model and the AFT spatial models. In the AFT spatial model, we will consider three different cases according to different spatial correlations, which are summarized as:

- Case 1: W_i is spatially uncorrelated. That is W_i follows an independent normal distribution.
- Case 2: W_i is spatially correlated. That is W_i follows the CAR model.
- Case 3: $W_i = W_{i1} + W_{i2}$, where W_{i1} is the spatially correlated random effect and W_{i2} denotes the spatially uncorrelated random effect. This case consider both spatially correlated and uncorrelated effects.

In our algorithm we ran two, initially overdispersed, parallel McMC chains for 20 000 iteration each. Then, we discarded the first 10 000 iterations as pre-convergence burn-in and retained 10 000 as the posterior analysis. In the Bayesian framework, model assessments and choices of the best-fitting model can be performed using the DIC [22], which is a Bayesian analog of the Akaike's information criterion (AIC). pD represents the effective number of parameters, which reflects the model complexity or degrees of freedom. If a Bayesian hierarchical model has negligible prior information, pD will approximate the actual number

of parameters, and DIC will approximate AIC. Lower values of DIC indicate a better-fitting model. Spiegelhalter et al. [22] suggest that models with DIC values within 1 or 2 units of the "best" model deserve consideration, those with values within 3-7 units of the "best" are only weakly supported and models with a DIC value more than 7 units higher than the "best" model are substantially inferior. The DIC and pD are listed in Table 3.

From Table 3, we can see that the model with spatially correlated random effects under the normal baseline is the best among these cases, which has the smallest DIC value (11 930). Therefore, we believe that the survival probability is affected by the geographical region. The pD value for the normal distribution with spatial correlation (Normal+Case 2) is 17.59, which indicates the complexity of the model.

The estimated parameters for the normal distribution with spatial correlation is summarized in Table 4.

The exponential of coeffcient illustrates the effect of covariate on survival time. For example: one unite change of age decreases the survival time by $e^{-0.0336} = 0.967$. From the table we can see that the age, race and stage have significant influence on the survival probability of the PrCA (as indicated by * in the table). Marital status does not display significance. More than 75% of patients in this study are married, so there may not be enough evidence to show the effect of the marital status.

In order to show the spatial effect, we present the median of the posterior spatial random effect in Figure 2, which displays considerable spatial structure in the middle eastern area of the state. It is worthwhile pointing out that the survival time is affected by the exponential of spatial random effects. The larger the spatial random effects indicates the longer the survival time. Note, the cut point in this figure is generated automatically by the default in WinBUGS, which is based on the absolute value of the variable to be mapped and are chosen to give equally spaced intervals.

For illustration purpose, we compare the estimated survival curves between five regions indicated in Figure 2 since the survival curve for each region will be effected by the spatial random effects. The estimated survival curves for the different races based on the AFT spatial model and the KM approach according to the different regions are displayed in Figure 3. In the KM approach, we only consider the race effect. In the AFT spatial model, we consider the median value of the age, marital status and stage for each race and the median of the estimated spatial random effects in each region. The survival curves are illustrated in Figure 3.

We can see that the survival curves from the AFT spatial model are similar to those from the KM approach, which indicates that the AFT spatial model fits the data set well. As appointed out by a referee, the survival curve corresponding to "Black" for the AFT spatial model does not fit very well to that for the KM approach when the time is around 20 to 30 in Region 1, 2, or 3. In order to solve this issue, we may relax the parametric assumption. For example, we may assume that the survival time follows the generalized gamma distribution or piecewise exponential distribution. Non parametric approach can also be considered. The survival probability is different in each region, which may indicate the geographical effect on PrCA. For example, the survival rate for black men at 60 months is around 0.6 in region 1 from the AFT spatial model, 0.67 from the KM approach, while in region 5 it is around 0.8 from the AFT spatial model and 0.83 from the KM approach. Similar effects can be found in other regions.

7. Discussion and Conclusion

In this paper, we investigate the PrCA data of Louisiana from the SEER program by the Bayesian parametric AFT spatial model. The spatially correlated and un-correlated heterogeneity were considered to characterize the spatial distribution pattern for better understanding the geographic features of survival probability of PrCA. WinBUGS is used to analyze the PrCA data via the parametric AFT spatial model and the DIC criterion is used to select the appropriate assumption for the error term distribution and correlation structure. The estimated survival curve from the parametric AFT spatial model is compared to that from the nonparametric approach. Finally, we concluded that the normal AFT spatial model with correlated spatial random effects is the best fitting model to analyze the PrCA data of Louisiana. The results from our study indicate that the age, race, stage and geographical distribution have significant impact in evaluating PrCA survival in Louisiana.

However, we have focused on a parametric specification of the AFT spatial model, so the DIC criteria is applied to check the model fitting. The model can be more flexible if we release the parametric assumption, which will increase the computational diffculties [5,7,24]. For the spatial component a CAR model is a common choice [9]. We demonstrate that different forms of spatial model have variable success in describing the Louisiana data, but a CAR model yields the best fit. Even the semiparametric structure is more flexible than the parametric model, the parametric spatial survival model is recommended in this paper because it can be conducted easily in WinBUGS in practice. In addition, we fit this data set by the PH spatial model with the lognormal distribution. Under the PH spatial model with the lognormal distribution, the DIC and pD are 12 010 and 19.40 for the spatially uncorrelated case, 12 006 and 19.98 for the spatially correlated case, and 12 007 and 20.14 for the convolution prior. Thus the lognormal PH spatial model has a better fit under the spatial correlation structure. However, comparing with the AFT spatial model, we find that the AFT spatial model with the normal distribution (DIC=11 930; pD=17.59) is still a better fit than a PH spatial model with the lognormal distribution (DIC=12 006; pD= 19.98).

It is also worthwhile pointing out that this model could be extended in a number of ways. First, more complex spatial structures could be included. Second, some smoothing terms rather than linear terms could be allowed for covariate modeling, such as $\log(T_{ij}) = \mu + \sum f_k(x_k) + W_i + \sigma\varepsilon_{ij}$, where $f_k(\cdot)$ is a unknown function. Finally latent spatial structure could be important and this might be a fruitful path to pursue in application to such survival data.

Acknowledgments

We sincerely thank the referees for the valuable comments that lead to greatly improved presentation of our work. The project described was supported by Award Number R03CA139538 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the offcial views of the National Cancer Institute or the National Institutes of Health.

Appendix

```
    # s0[i]<-1-phi(temp[i])
    # f0[i]<-pow(2*3.14, -0.5)*exp(-0.5*pow(temp[i],2))

    #survival distribution and density function of  extreme value distribution
    #s0[i]<-exp(-exp(temp[i]))
    #f0[i]<-s0[i]*exp(temp[i])

    #survival distribution and density function of  logistic distribution
    s0[i]<-1/(1+exp(temp[i]))
    f0[i]<-exp(temp[i])*pow(s0[i],2)

   #Loglikelihood Function
L[i]<-status[i]*log(f0[i]/(sigma*(time[i]+0.01)))+(1-status[i])*log(s0[i])

#poisson zero trick
zeros[i]<-0
new[i]<- -L[i]
zeros[i]~dpois(new[i])
}

#Car model specification
#for(j in 1:sumNum) { weights[j] <- 1}
#W[1:regions] ~ car.normal(adj[], weights[], num[], tau)
#W.mean <- mean(W[])

#spatial from independent normal distribution
#for (i in 1:regions){ W[i] ~ dnorm(0, tau)
 #                      intercept[i]<-beta0+W[i]}

#spatial from independent normal distribution plus the car model effects
for(j in 1:sumNum) { weights[j] <- 1}
W1[1:regions] ~ car.normal(adj[], weights[], num[], tau)
W1.mean <- mean(W1[])
for (j in 1:regions){ W2[j] ~ dnorm(0, tau1)
                      W[j]<-W1[j]+W2[j]
                      intercept[j]<-beta0+W[j]
                       }

#Parameter Prior for the parameters in the AFT model
beta0 ~ dnorm(0.0,0.001)
 for(i in 1:4) {beta[i] ~ dnorm(0.0, 0.001)}
 inversesigma~dgamma(0.001,0.001)
sigma<-1/inversesigma

    #parmaters in the car model or normal model
#tau ~ dgamma(0.001, 0.001)

#parameters in the mix car model and normal model
tau ~ dgamma(0.001, 0.001)
tau1~dgamma(0.001,0.001)
}
```

References

[1]. Banerjee S, Dey DK. Semiparametric proportional odds models for spatially correlated survival data. Lifetime Data Anal. 2005; 11:175–191. [PubMed: 15938545]

[2]. Banerjee S, Wall MM, Carlin BP. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. Biostatistics. 2003; 4:123–142. [PubMed: 12925334]

[3]. Bastos LS, Gamerman D. Dynamic survival models with spatial frailty. Lifetime Data Anal. 2006; 12:441–460. [PubMed: 17031498]

[4]. Besag J, Mollie A, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics (with discussion). Ann. Inst. Stat. Math. 1991; 43:1–59.

[5]. Christensen R, Johnson W. Modelling accelerated failure time with a Dirichlet process. Biometrika. 1988; 75:693–704.

[6]. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences (with discussion). Statist. Sci. 1992; 7:457–511.

[7]. Hanson T, Johnson WO. A Bayesian semiparametric aft model for interval-censored data. J. Comput. Graph. Statist. 2004; 13:341–361.

[8]. Hanson T, Yang M. Bayesian Semiparametric Proportional Odds Models. Biometrics. 2007; 63:88–95. [PubMed: 17447933]

[9]. Henderson R, Diggle P, Dobson A. Identification and efficacy of longitudinal markers for survival. Biostatistics. 2002; 3:33–50. [PubMed: 12933622]

[10]. Henderson R, Shimakura S, Gorst D. Modeling spatial variation in leukemia survival data. J. Amer. Statist. Assoc. 2002; 97:965–972.

[11]. National Cancer Institute. Surveillance and Epidemiology and End Results (SEER) Program. www.seer.cancer.govLimited-Use DataDCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2008, based on the November 2007 submission, 1973-2005

[12]. Hennerfeind A, Brezger A, Fahrmeir L. Geoadditive survival models. J. Amer. Statist. Assoc. 2006; 101:1065–1075.

[13]. Klein, JP.; Moeschberger, ML. Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag Inc; New York: 1997.

[14]. Komárek A, Lesaffre E. Bayesian accelerated failure time model for correlated interval-censored data with a normal mixture as error distribution. Statist. Sinica. 2007; 17:549–569.

[15]. Komárek A, Lesaffre E, Hilton JF. Accelerated failure time model for arbitrarily censored data with smoothed error distribution. J. Comput. Graph. Statist. 2005; 14:726–745.

[16]. Li Y, Lin X. Semiparametric normal transformation models for spatially correlated survival data. J. Amer. Statist. Assoc. 2006; 101:591–603.

[17]. Li Y, Ryan L. Modeling spatial survival data using semiparametric frailty models. Biometrics. 2002; 58:287–297. [PubMed: 12071401]

[18]. Lunn D, Thomas A, Best N, Spiegelhalter D. Winbugs–a bayesian modelling framework: concepts, structure, and extensibility. Statist. Comput. 2000; 10:325–337.

[19]. Mallick BK, Denison DGT, Smith AFM. Bayesian survival analysis using a mars model. Biometrics. 1999; 55:1071–1077. [PubMed: 11315050]

[20]. Silva GL, Amaral-Turkman MA. Bayesian analysis of an additive survival model with frailty. Comm. Statist. A. 2004; 33:2517–2533.

[21]. Sinha D, Dey DK. Semiparametric Bayesian analysis of survival data. J. Amer. Statist. Assoc. 1997; 92:1195–1212.

[22]. Spiegelhalter DJ, Best NG, Carlin BP, Lindevan der A. Bayesian measures of model complexity and fit (Pkg: P583-639). J. Roy. Statist. Soc. Ser. B. 2002; 64:583–616.

[23]. Therneau, TM.; Grambsch, PM. Modeling Survival Data: Extending the Cox Model. Springer-Verlag Inc; New York: 2000.

[24]. Walker S, Mallick BK. A Bayesian semiparametric accelerated failure time model. Biometrics. 1999; 55:477–483. [PubMed: 11318203]

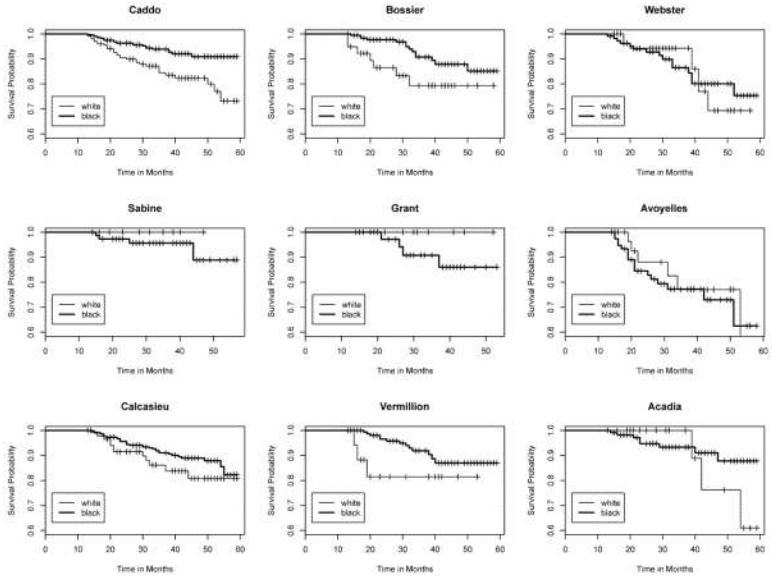

Figure 1.
KM survival curves according to race for PrCA in nine different counties, Louisiana.

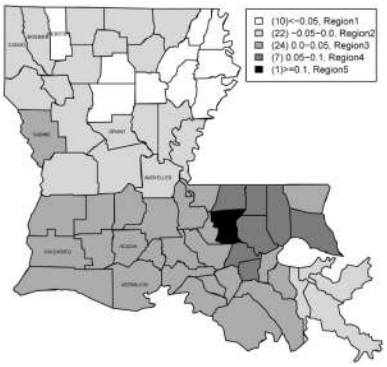

Figure 2.
Posterior mean of spatial random effects from the best fitted model, Louisiana counties.

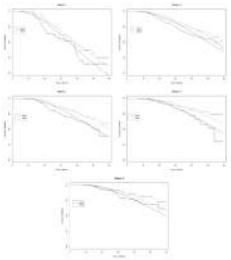

Figure 3.
Fitted survival curves from the KM approach and the AFT spatial model, Region 1-5, Louisiana counties. Step line represents the estimated survival curve from the KM approach and smoothed line represents the estimated curve from the normal AFT spatial model.

Table 1

Summary characteristics of prostate cancer patients: Louisiana, 2000-2004

Covariate	N	Patients(%)
Race		
Black	3 006	29.34
White	7 240	70.66
Marital Status		
Single (Never married)	939	9.16
Married	7 752	75.66
Other (Separated, divorced)	1 555	15.18
Cancer Stage		
Localized/regional	9 870	96.33
Distant	376	3.67
Vital status		
Alive	9 274	90.51
Dead	972	9.49

Table 2

Common distributions in the AFT spatial model, where $\lambda(\mathbf{x}_{ij}) = \mu + \beta\mathbf{x}_{ij} + \mathrm{W}_i$

Distribution	$S_0(\cdot)$	$S(\cdot)$
Normal	$1 - \Phi(\varepsilon_{ij})$	$1 - \Phi\left(\frac{\log(t_{ij}) - \lambda(\mathrm{x}_{ij})}{\sigma}\right)$
Extreme value	$\exp(-\exp(\varepsilon_{ij}$	$\exp(-\exp(-\lambda(\mathrm{x}_{ij}))t_{ij})^{\frac{1}{\sigma}}$
Logistic	$\frac{1}{1+\exp(\varepsilon_{ij})}$	$\frac{1}{1+(\exp(-\lambda(\mathrm{x}_{ij}))t_{ij})^{1/\sigma}}$

Table 3

A comparison of goodness-of-fit (DIC, pD) for the AFT model and the three cases of the AFT spatial model between the three baseline survival distributions (normal, extreme value and logistic)

Distribution of $S_0(\cdot)$	DIC	pD
Normal AFT	11 950.0	6.133
Normal+Case 1	11 960.0	65.42
Normal+Case 2	11 930.0	17.59
Normal+Case 3	12 000.0	85.01
Extreme value AFT	12 090.0	12.05
Extreme+Case 1	12 090.0	66.51
Extreme+Case 2	12 060.0	19.25
Extreme+Case 3	12 200.0	117.60
Logistic AFT	12 000.0	5.768
Logistic+Case 1	12 020.0	68.93
Logistic+Case 2	11 990.0	17.27
Logistic+Case 3	12 090.0	99.14

Table 4

The best tting AFT spatial model (Normal+Case 2): mean parameter estimates, sample standard deviations, and quantiles

	mean	sd	2.5%	50%	97.5%
age	−0.0336*	0.00184	−0.0374	−0.0335	−0.0301
marital	−0.0096	0.0299	−0.0688	−0.0100	0.0494
race	0.1843*	0.0360	0.1171	0.1835	0.2553
stage	−0.9643*	0.0569	−1.080	−0.9617	−0.8578

*denotes that the coefficient is significant at 0.95 level.