

# Posterior Inference for White Hispanic Breast Cancer Survival Data

Hafiz MR Khan<sup>1\*</sup>, Anshul Saxena<sup>2</sup> and Alice Shrestha<sup>1</sup><sup>1</sup>Department of Biostatistics, Robert Stempel College of Public Health & Social Work Florida International University, Miami, FL 33199, USA<sup>2</sup>Department of Health Promotion & Disease Prevention, Robert Stempel College of Public Health & Social Work, Florida International University, Miami, FL 33199, USA**Abstract**

The purpose of this paper is to develop a statistical probability model and to obtain posterior inference for the parameters given the survival times of the White Hispanic female cancer patients. Stratified random sample of White Hispanic female patients' survival data was used to derive a best fit statistical probability model. The study sample was extracted from the Surveillance Epidemiology and End Results (SEER) cancer registry database. Three model building criterions were utilized; Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Deviance Information Criteria (DIC) to measure the goodness of fit. We found that the Exponentiated Weibull model fits the survival times better as compared to other widely known statistical probability models. The Bayesian approach is employed to derive the posterior inference for the parameters.

**Keywords:** Breast cancer survival data; Statistical inference; Bayesian inference

**Introduction**

Breast cancer is a cancer that begins in the tissues of the breast. It develops as a result of uncontrolled growth of altered cells of the breast [1]. These cells form a tumor, which can be palpated and felt by some females as a lump or mass during breast self-examination. Recent studies have indicated that one in eight women in the United States (U.S) will develop invasive breast cancer during their lifetime [2]. Regardless of what race and ethnicity women belong to, breast cancer in U.S. is the most common cancer in women excluding some kinds of skin cancer [3]. It accounts for a very high prevalence of 16% among all cancer types [4]. In 2008, breast cancer claimed 458,400 lives [5]. Moreover, approximately 60% of deaths due to breast cancer occurred in developing nations, which contrasted the common belief among cancer researchers that breast cancer is prevalent in developed countries. In 2010, 206,966 women were diagnosed with breast cancer and 40,996 women died from the disease in the US [6]. The incidence rate of breast cancer among White females is 119.5 per 100,000 population, which is highest among all races [3].

Breast cancer is the most common cancer among American women accounting for the highest overall incidence rate of 123.1 per 100,000 population among all cancers [1,6]. Some major known risk factors of breast cancer are age, smoking, excessive alcohol drinking, obesity, lactation, and family history [7,8]. It was believed that level of awareness, fewer mammograms, socio-economic factors, and lack of access to health care are strong risk factors for breast cancer. However, recent studies indicate that ethnic differences could also be an important factor associated with breast cancer mortality and incidence rates [1].

In 2013, an estimated 232,340 new cases of invasive breast cancer diagnosis were expected to be diagnosed among US women along with an estimated 64,640 additional cases of in situ breast cancer. Approximately 39,620 women were expected to die from breast cancer [9]. The incidence rate of breast cancer among all racial and ethnic groups remained stable from 2004-2008. The death rates among all ethnic groups began decreasing in the early 1990s except for the American Indians/Alaska Natives [2]. According to the American Cancer Society, sufficient evidence supports the fact that there are disparities among breast cancer death rates by state, socioeconomic status, and race/ethnicity [2]. Research focusing on each of these determinants is necessary in order to have an understanding of the

breast cancer epidemic among American women. Despite the fact that age remains the strongest risk factor for breast cancer, race and ethnicity also contributes to the increased probability of developing breast cancer [1]. White women are more likely to develop breast cancer compared to any other racial group in the United States [1].

According to the US Census Bureau, 50.0 million Americans, or 16% of the total population identified themselves as Hispanic or Latino in 2010 [4]. About 17,100 diagnosis among Hispanic women were expected to be diagnosed and 2400 were expected to die from breast cancer in 2012 [4]. Among all ethnicities in US, the overall death rates of breast cancer are highest among Hispanic women [3]. Lack of awareness and inaccessibility to health care are possible factors that explain the elevated risk of breast cancer among Hispanic women as compared to other ethnic groups [1].

Although previous studies suggest that Mexican women have a lower risk of developing breast cancer, current research indicates an increase in the incidence rates of breast cancer among Mexican-American women. The Arizona Cancer Center and three Mexican universities have collaborated in the Ella Binational Breast Cancer Study (EBBCS) to gather data that can provide insight on the breast cancer differences between Mexican native and Mexican-American women [1]. According to the EBBCS, Mexican women who live in the US have increased risk of breast cancer due to lifestyle and reproductive factors. Studies conducted on all other U.S. born Hispanic women correspond to the EBBCS findings. Physical inactivity, early menarche, late menopause, postmenopausal obesity, and alcohol consumption are responsible for the increase in the risk of breast cancer among the Mexican-American women [4]. On the other hand, having more children breast-feeding for a longer period of time, active lifestyle, and

**\*Corresponding author:** Hafiz MR Khan, Department of Biostatistics, Robert Stempel College of Public Health & Social Work Florida International University, Miami, FL 33199, USA, Tel: +001-305-348-4908; Fax: +001-305-348-4901; E-mail: [hmkhan@fiu.edu](mailto:hmkhan@fiu.edu)

**Received** November 18, 2013; **Accepted** January 11, 2014; **Published** January 18, 2014

**Citation:** Khan HMR, Saxena A, Shrestha A (2014) Posterior Inference for White Hispanic Breast Cancer Survival Data. J Biomet Biostat 5: 183. doi:[10.4172/2155-6180.1000183](https://doi.org/10.4172/2155-6180.1000183)

**Copyright:** © 2014 Khan HMR, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

more fiber consumption lowers the risk of getting breast cancer among women born in Mexico [1].

Cancer survival data is recorded and stored at various hospitals and cancer registries, so that it can be used for future analysis. There is a high demand for novel statistical analysis and methods to understand such type of data in a scientific manner. Statistical analysis can give an idea of inferences about the exiting survival data and its probability model.

The main goals of this paper are (i) to review certain right skewed models; (ii) to give a justification that the given sample data set follows a specific model by using model selection criterions for goodness of fit tests; and (iii) to perform a Bayesian analysis of the posterior distribution for the parameters.

This paper is organized as follows. A real breast cancer survival data example related to goodness of fit and reparameterization, and posterior inference for the model parameters of white ethnicity are presented in Section 2. Finally, conclusion is added in Section 3.

## Probability Model Testing

The data extracted from healthcare experiments may follow several statistical probability models, for example, exponential, gamma, lognormal, Weibull, exponentiated exponential (EE), exponentiated Weibull (EW), and beta generalized exponential (BGE), etc. Statistical methodologies are immensely necessary to understand and make scientific conclusions from such type of data.

There are many statistical probability models have been used in modeling survival data. In this paper we consider exponentiated exponential model (EEM), beta generalized exponential (BGEM), exponentiated Weibull model (EWM), and beta inverse-Weibull (BIW) because for specific values of the parameters they reduces to certain statistical probability models.

The exponentiated exponential model (EEM) is used in modeling the data from engineering and biomedical sciences. The EEM has two parameters, scale and shape. A random variable  $x$  is said to have an exponentiated exponential distribution if its probability density function (pdf) is given by

$$p(x) = \alpha \lambda \exp\{- (\lambda x)\} \left(1 - \exp\{- (\lambda x)\}\right)^{\alpha-1},$$

where  $\alpha > 0$  and  $\lambda > 0$  are the shape and scale parameters, respectively.

The probability density function of beta generalized exponential model is given by

$$p(x) = \frac{\alpha \lambda}{B(a, b)} \exp\{- (\lambda x)\} \left(1 - \exp\{- (\lambda x)\}\right)^{a-1} \left(1 - \left(1 - \exp\{- (\lambda x)\}\right)^a\right)^{b-1},$$

where the shape parameter,  $\alpha > 0$  and the scale parameter,  $\lambda > 0$ . There are two additional parameters,  $a > 0$  and  $b > 0$  whose role is to introduce skewness and to vary tail weight [10]. The BGE model generalizes some well-known models; beta exponential and generalized exponential models are the special cases.

Mudholkar and Srivastav [11] presented the first exponentiated Weibull model (EWM). The probability density function for the exponentiated Weibull model is given by

$$p(x) = \alpha \beta \lambda x^{\beta-1} \exp\{- (\lambda x^\beta)\} \left(1 - \exp\{- (\lambda x^\beta)\}\right)^{\alpha-1},$$

where  $\alpha > 0$  and  $\beta > 0$  are the shape parameters, and  $\lambda > 0$  is the scale parameter.

The beta inverse-Weibull (BIW) model is one of the widely used distributions for problems in medicine and reliability. It shows a good fit to several data sets such as the times to breakdown of an insulating fluid and subject to constant tensions [12]. The probability density function of beta inverse-Weibull model is given by

$$p(x) = \frac{\beta x^{-(\beta+1)}}{B(a, b)} \exp\{- (x^{-\beta})\} \left(\exp\{- (x^{-\beta})\}\right)^{a-1} \left(1 - \exp\{- (x^{-\beta})\}\right)^{b-1},$$

where  $\beta > 0$  is the shape parameter, and two additional parameters,  $a > 0$  and  $b > 0$  whose role is to introduce skewness and to vary tail weight.

A Bayesian method is used to explore the posterior probability for the parameters from the EEM, BGEM, EWM, and BIWM. The purpose of Bayesian method is to develop the posterior inference for the parameters given a set of observed data. For further information about Bayesian method, the readers can refer to several published works [13-19]. Additional applications of Bayesian method for predictive inference have been discussed by a number of authors [20-26].

## Example of Breast Cancer Survival Data

We used the breast cancer data (N=657,712) from Surveillance, Epidemiology and End Results (SEER, 1973-2009) cancer registry website in the USA [27]. In the USA, there are twelve states that collect breast cancer patients' information. The total SEER data were by gender: males=4,269 and females=653,443. Among the total females=608,032, White Hispanic=22,639, White non-Hispanic=531,562. Since there is a small chance that breast cancer will occur in males, they were not considered in this study. Stratified random sampling scheme was used to randomly select a sample of nine states out of twelve data-recorded states to represent White Hispanic breast cancer cases. Exclusion of three states will allow other researchers to perform external validation of our findings, since the information-theoretic criteria are essentially internal validations. However, external validation is beyond the scope of the present study. Finally, a simple random sampling (SRS) method was then used to select 2,000 White Hispanic patients from the selected nine states.

## Goodness of fit and reparameterization

The most commonly used methods to measure the goodness of fit for the models are Akaike Information Criterion (AIC), Deviance Information Criterion (DIC), and Bayesian Information Criterion (BIC). Among these methods, DIC is the most widely used method. It is a Bayesian measure of fit, which is used for comparison of different models where the samples of the posterior distribution of parameters are obtained by Markov chain Monte Carlo (MCMC) methods, for example, the use of public data by Congdon [28,29]. The values of DIC can be both positive and negative. Model with lower DIC value is considered better than others. BIC is an asymptotic result assumed that the distribution of data is an exponential family. Similar to AIC, given any two estimated models, the model with a lower value of BIC is preferred. We used multiple information-theoretic criteria in this study to see whether or not each criterion agrees in selecting a best fitting model. Achcar et al. [30] used a re-parameterization for certain skewed models. A re-parameterization method may apply considering the log-likelihood functions based on  $\mathbf{x}=(x_1, x_2, \dots, x_n)$  from the models described earlier, which are given below:

The log-likelihood function from the EE model which is given by

$$l(\alpha, \lambda | \mathbf{x}) = n \log(\alpha) + n \log(\lambda) + (\alpha - 1) \sum_{i=1}^n \log(1 - \exp\{- (\lambda x_i)\}) - \lambda \sum_{i=1}^n x_i.$$

Assume  $\rho_1 = \log(\alpha)$  and  $\rho_2 = \log(\lambda)$ . We further assume that  $\rho_1$  and  $\rho_2$  are independently distributed. To obtain non-informative prior for  $\rho_1$  and  $\rho_2$ , let a uniform prior distribution for  $\rho_i$  be  $U(-c_i, c_i), \forall i = 1, 2$ .

The log-likelihood function from the beta generalized exponentiated model, which is given by

$$\ell(\alpha, \lambda, a, b | \mathbf{x}) = n \log(\alpha / B(a, b)) - \lambda \sum_{i=1}^n x_i + (a\alpha - 1) \sum_{i=1}^n \log(1 - \exp\{-\lambda x_i\}) + (b - 1) \sum_{i=1}^n \log\{1 - [1 - \exp\{-\lambda x_i\}]^a\}$$

Assume  $\rho_1 = \log(a)$ ;  $\rho_2 = \log(b)$ ;  $\rho_3 = \log(\alpha)$ ; and  $\rho_4 = \log(\lambda)$ . We further assume that  $\rho_1, \rho_2, \rho_3$ , and  $\rho_4$  are independently distributed. To obtain non-informative prior for  $\rho_1, \rho_2, \rho_3$ , and  $\rho_4$  let a uniform prior distribution for  $\rho_j$  be  $U(-d_j, d_j), \forall j = 1, 2, 3, 4$ .

The log-likelihood function from the EW model which is given by

$$\ell(\alpha, \beta, \lambda | \mathbf{x}) = n \log(\alpha) + n \log(\beta) + n \log(\lambda) + (\alpha - 1) \sum_{i=1}^n \log(1 - \exp\{-\lambda x_i^\beta\}) - \lambda \sum_{i=1}^n x_i^\beta + (\beta - 1) \sum_{i=1}^n \log(x_i)$$

Assume  $\rho_1 = \log(a)$ ;  $\rho_2 = \log(\beta)$ ; and  $\rho_3 = \log(\lambda)$ . We further assume that  $\rho_1, \rho_2$  and  $\rho_3$  are independently distributed. To obtain non-informative prior for  $\rho_1, \rho_2$  and  $\rho_3$ , let a uniform prior distribution for  $\rho_k$  be  $U(-e_k, e_k), \forall k = 1, 2, 3$ .

The log-likelihood function from the beta inverse Weibull model, which is given by

$$\ell(\beta, a, b | \mathbf{x}) = n \log(\beta / B(a, b)) - a \sum_{i=1}^n x_i^{-\beta} + (\beta - 1) \sum_{i=1}^n \log(x_i) + (b - 1) \sum_{i=1}^n \log(1 - \exp\{-x_i^{-\beta}\})$$

Assume  $\rho_1 = \log(\beta)$ ;  $\rho_2 = \log(a)$ ; and  $\rho_3 = \log(b)$ . We further assume that  $\rho_1, \rho_2$  and  $\rho_3$  are independently distributed. To obtain non-informative prior for  $\rho_1, \rho_2$  and  $\rho_3$  let a uniform prior distribution for  $\rho_g$  be  $U(-f_g, f_g), \forall g = 1, 2, 3$ .

Table 1 presents the results of the measures of goodness of fit. The posterior distributions for the parameters and their results are reported in Tables 2-5. The posterior kernel densities for the parameters are given in Figures 1-4. Specifications for the kernel density estimation were the WinBugs defaults.

### Results of goodness of fit tests and posterior inference for the parameters from the White Hispanic survival data

The following AIC, BIC, and DIC values are calculated and the posterior inference for the parameters with their corresponding kernel densities are obtained.

Table 1 consists of AIC, BIC, and DIC values for the EE, EW, BGE, and BIW models. Lower values of AIC, BIC, and DIC infer better model fit. The data fits EW distribution better than other models. The estimated value of AIC is the lowest (19425.700) while the DIC value is

Model criterions	AIC	BIC	DIC
Exponentiated exponential	19430.400	19441.602	19430.426
Exponentiated Weibull	19425.700	19442.001	19423.700
Beta generalized exponential	19433.300	19455.703	19429.300
Beta inverse Weibull	19444.400	19465.700	19442.300

Table 1: Selection of the best model for White Hispanic females on the basis of AIC, BIC, and DIC criterions.

Node	Mean	SD	MC error	Median	95% CI	Sample
alpha	8.152	0.3771	0.005046	8.144	(7.44, 8.909)	50,000
lambda	0.0351	7.19E-04	9.50E-06	0.0351	(0.03369, 0.03652)	50,000
rho1	2.097	0.04627	6.18E-04	2.097	(2.007, 2.187)	50,000
rho2	-3.35	0.02049	2.71E-04	-3.35	(-3.39, -3.31)	50,000

Table 2: Summary results of the posterior parameters in the case of exponentiated exponential for White Hispanic females breast cancer patients (n=2,000).

Node	Mean	SD	MC error	Median	95% CI	Sample
alpha	6.338	0.3607	0.0171	6.295	(5.744, 7.164)	50,000
beta	1.099	0.01956	0.001189	1.103	(1.052, 1.124)	50,000
lambda	0.02083	0.002238	1.38E-04	0.02019	(0.01839, 0.02642)	50,000
rho1	1.845	0.05629	0.002639	1.84	(1.748, 1.969)	50,000
rho2	0.09379	0.01798	0.001095	0.0981	(0.05065, 0.1166)	50,000
rho3	-3.877	0.1015	0.006212	-3.903	(-3.996, -3.634)	50,000

Table 3: Summary results of the posterior parameters in the case of exponentiated Weibull (EW) for White Hispanic females breast cancer patients (n=2,000).

Node	Mean	SD	MC error	Median	95% CI	Sample
a	4.741	1.319	0.08746	4.483	(2.928, 7.226)	50,000
alpha	1.811	0.4869	0.03249	1.777	(1.093, 2.665)	50,000
b	1.042	0.02052	2.94E-04	1.044	(1.003, 1.071)	50,000
lambda	0.03407	8.50E-04	1.64E-05	0.03404	(0.03246, 0.03579)	50,000
rho1	1.518	0.2774	0.01853	1.5	(1.074, 1.978)	50,000
rho2	0.04052	0.01977	2.83E-04	0.04303	(0.002838, 0.0688)	50,000
rho3	0.5564	0.2759	0.01844	0.5747	(0.08872, 0.9801)	50,000
rho4	-3.38	0.02494	4.80E-04	-3.38	(-3.428, -3.33)	50,000

Table 4: Summary results of the posterior parameters in the case of beta generalized exponentiated for White Hispanic females breast cancer patients (n=2,000).

Node	Mean	SD	MC error	Median	95% CI	Sample
a	1.031	0.01789	9.39E-05	1.031	(1.002, 1.06)	50,000
b	1.047	0.02722	1.25E-04	1.046	(1.002, 1.092)	50,000
beta	403.2	0.2026	0.001536	403.3	(402.7, 403.4)	50,000
rho1	6	5.03E-04	3.81E-06	6	(5.998, 6.0)	50,000
rho2	0.03037	0.01736	9.11E-05	0.03054	(0.001552, 0.05857)	50,000
rho3	0.04527	0.02601	1.20E-04	0.0453	(0.002265, 0.08786)	50,000

Table 5: Summary results of the posterior parameters in the case of beta inverse Weibull for White Hispanic females breast cancer patients (n=2,000).

the least (19423.700) in the case of EWM. The estimated value of BIC (19442.001) is very close to the lowest (19441.602) while the DIC value is the lowest (19423.700) in the case of EWM. Comparing the estimated values of AIC, BIC, and DIC for the models, the EWM fits better for the survival days because it produces smaller values of AIC, BIC, and DIC.

Table 2 indicates summary results of the posterior distribution of the parameters from the exponentiated exponential by using the White Hispanic breast cancer patient's survival data. By generating the values of the  $\rho_1$  and  $\rho_2$  from the data, the results of the posterior distribution parameters  $\alpha$  and  $\lambda$  are estimated using the MCMC method. The log-likelihood function is derived from the exponentiated exponential model and then its parameter values are assigned to the appropriate theoretical probability distributions. The WinBugs software is used to obtain the summary results of the parameters. After removing the burn-in samples, the remaining samples are treated as if the samples are from the original distribution. The procedure was conducted by 50,000 Monte Carlo repetitions to produce the inference for the posterior parameters in Table 2. Figure 1 displays the graphical representation of the parameters behavior. After 50,000 Monte Carlo repetitions, it

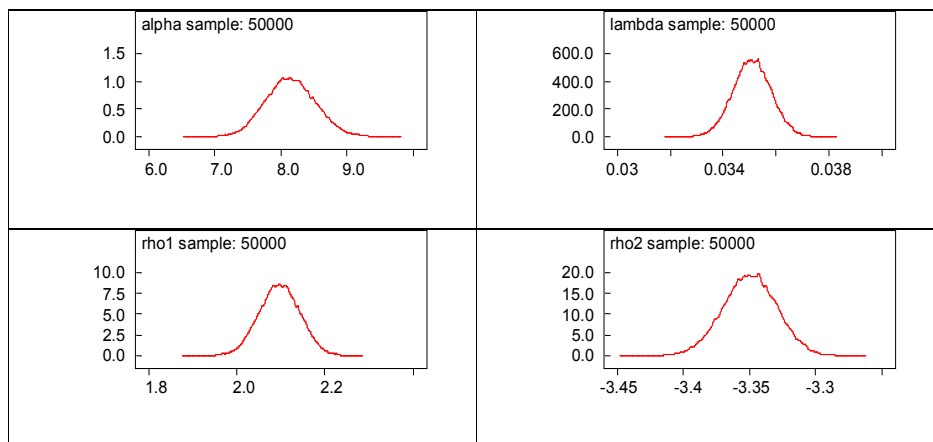


Figure 1: Kernel density of the posterior parameters in the case of exponentiated exponential for White Hispanic females breast cancer patients (n=2,000).

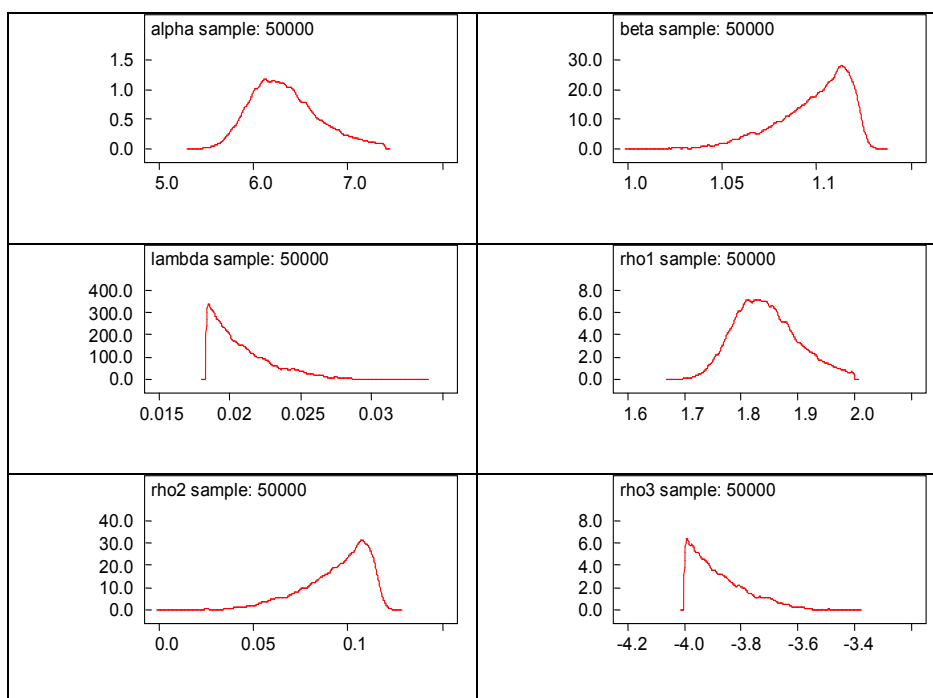


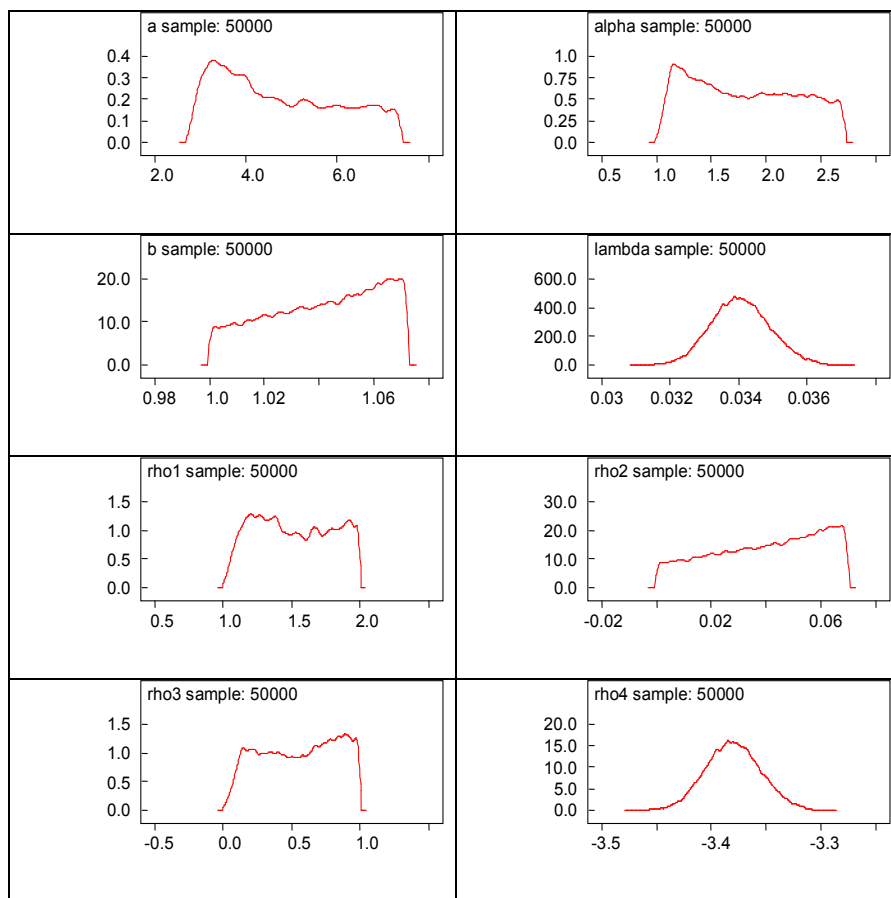
Figure 2: Kernel density of the posterior parameters in the case of exponentiated Weibull for White Hispanic females breast cancer patients (n=2,000).

is noted that the kernel densities for both shape and scale parameters exhibit approximately symmetric distribution.

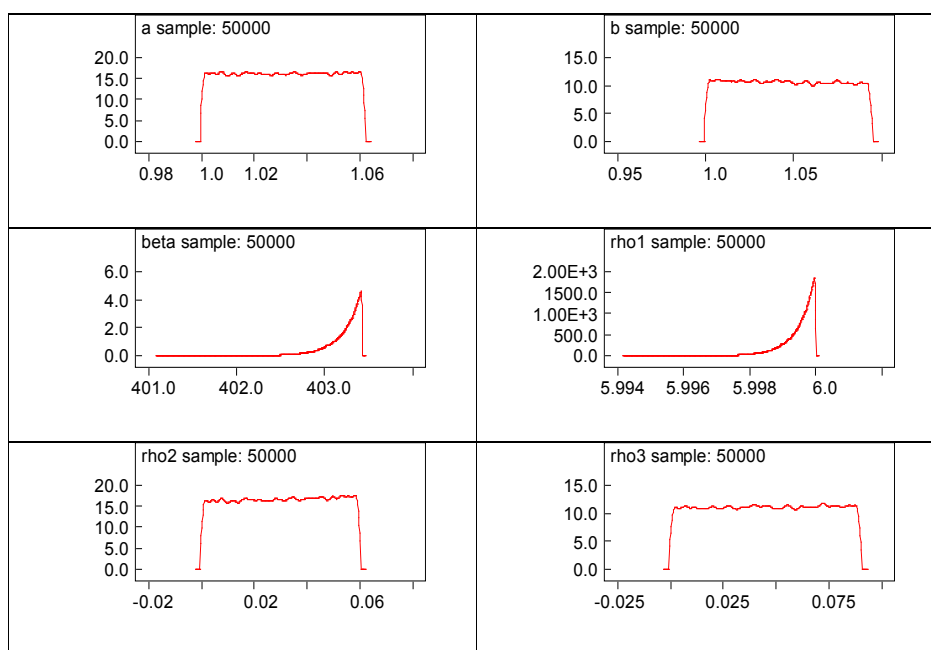
Table 3 indicates the summary results of the posterior distribution of the parameters from the exponentiated Weibull by using the White Hispanic female breast cancer patients' survival data. By setting the generated values  $\rho_1$ ,  $\rho_2$  and  $\rho_3$  from the data, the results of the posterior distribution parameters  $\alpha$ ,  $\beta$ , and  $\lambda$  are estimated using the MCMC methods. The log-likelihood function is derived from the exponentiated Weibull model and then by its parameter values which are assigned to appropriate probability distributions. The WinBugs software is used to obtain the summary results of the parameters. Figure 2 displays the graphical representation of the distributions of the parameters behaviors. It is noted that the distribution of the shape parameter  $\alpha$

deviates from symmetric distribution, and other model parameters  $\beta$  and  $\lambda$  are distributed as skewed models. The parameter  $\rho_1$  deviates from the normal; and both  $\rho_2$  and  $\rho_3$  form skewed models.

Table 4 indicates the summary results of the posterior distribution of the parameters from the beta generalized exponentiated model by using the White Hispanic female breast cancer patients' data. The WinBugs software is used to obtain the summary results (Mean, SD, MC Error, Median, and Confidence Intervals) of the parameters. Figure 3 displays the graphical representations of the parameters for female in the case of beta generalized exponential. It is noted that the parameters  $\lambda$  and  $\rho_4$  from the beta generalized exponential exhibit approximate normal distribution. The other parameters tend not only non-uniform, they are often not even remotely symmetric distributions.



**Figure 3:** Kernel density of the posterior parameters in the case of beta generalized exponential for White Hispanic females' breast cancer patients (n=2,000).



**Figure 4:** Kernel density of the posterior parameters in the case of beta inverse Weibull for White Hispanic females' breast cancer patients (n=2,000).

Table 5 indicates the summary results of the posterior distribution of the parameters from the beta inverse Weibull model by using the White Hispanic female breast cancer patients' survival data. The WinBugs software is used to obtain the summary results (Mean, SD, MC Error, Median, and Confidence Intervals) of the parameters. Figure 4 displays the graphical representations of the parameters for female in the case of beta inverse Weibull model. It is noted that the parameters  $\beta$  and  $\rho_1$  from the beta inverse Weibull exhibit skewed distribution, and other parameters remains approximately uniform distributions.

## Conclusion

Several statistical models were used to identify the best fit model for the White Hispanic female breast cancer patients' survival data. In the case of goodness of fit analysis, the breast cancer survival sample for the ethnicity followed exponentiated Weibull distribution. The lowest DIC value of White Hispanic is 19423.700. In the case of EWM, Mean  $\pm$  SD for  $\alpha$ ,  $\beta$ , and  $\lambda$  values are  $6.338 \pm 0.3607$ ,  $1.099 \pm 0.01956$ , and  $0.02083 \pm 0.002238$ , respectively.

We determined the inference for posterior parameters given breast cancer survival model by using the Bayesian method. By using less Markov Chain errors, the inferences for the posterior parameters are reported in Tables 2-5. The dynamic kernel densities for each of the parameters are reported in Figures 1-4 so that one can observe the shape of the kernel density.

Statistical probability models are very important to describe inferences for posterior model parameters. To develop the best statistical probability model for White Hispanic, we used model selection criteria, AIC, BIC, and DIC. The summary results of the posterior parameters are reported. The results are obtained after running 50,000 Monte Carlo repetitions. The results of the posterior distribution of parameters using the breast cancer patients' survival data will contribute a new addition to White Hispanic ethnicity. WinBugs software was used to check the goodness of fit tests, to obtain the summary results of the posterior parameters, to determine the kernel densities of the parameters, and also to carry out all related calculations.

## Acknowledgements

The authors would like to thank the editor-in-chief and the referees for their valuable comments and suggestions.

## References

1. National Cancer Institute (2012) What You Need To Know About™ Cancer: Understanding Cancer.
2. National Cancer Institute (2012). SEER Stat Fact Sheets: Breast Cancer.
3. American Institute for Cancer Research (2012). Breast cancer.
4. CDC (2012) Breast Cancer Rates by Race and Ethnicity.
5. American Cancer Society (2012-2014) Cancer Facts and Figures for Hispanics/Latinos.
6. Bray F, Ren JS, Masuyer E, Ferlay J (2013) Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer* 132: 1133-1145.
7. CDC (2012) Breast Cancer. Statistics.
8. Zhao G, Li C, Okoro CA, Li J, Wen XJ, et al. (2013) Trends in modifiable lifestyle-related risk factors following diagnosis in breast cancer survivors. *J Cancer Survivorship* 7: 563-569.
9. Antoniou AC, Easton DF (2006) Models of genetic susceptibility to breast cancer. *Oncogene* 25: 5898-5905.

10. Barreto-Souza W, Santos AHS, Cordeiro GM (2008) The beta generalized exponential distribution.
11. Mudholkar GS, Srivastava DK (1993) Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Trans. Reliability* 42: 299-302.
12. Nelson W (1982) *Applied Life Data Analysis*. New York: John Wiley and Sons.
13. Ahsanullah M, Ahmed SE (2001) Bayes and empirical Bayes estimates of survival and hazard functions of a class of distributions. *Lecture Notes in Statistics* 148: 81-87.
14. Gelman (2004) Parameterization and Bayesian modeling. *Journal of the American Statistical Association* 99: 537-545.
15. Baklizi A (2004) Preliminary test estimation in the rayleigh distribution using minimax regret significance levels. *Pertanika J Sci & Techno* 12: 45-50.
16. Baklizi A (2008) Estimation of  $\Pr(X \leq Y)$  using record values in the one and two parameter exponential distributions. *Communications in Statistics - Theory and Metho* 37: 692-698.
17. Bernardo JM, Smith AFM (1994) *Bayesian Theory*. Wiley, New York.
18. Berger JO (1985) *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, USA.
19. Geisser S (1993) *Predictive Inference: An Introduction*. Chapman and Hall, New York, USA.
20. Khan HMR, Haq MS, Provost SB (2004) Predictive distributions for responses from the Weibull life testing model. *Journal of Statistical Theory and Applications* 3: 53-73.
21. Khan HMR, Albatineh AN, Alshahrani S, Jenkins N, Ahmed NU (2011) Sensitivity analysis of predictive modeling for responses from the three-parameter Weibull model with a follow-up doubly censored sample of cancer patients. *Computational Statistics & Data Analysis* 55: 3093-3103.
22. Thabane L (1998) *Contributions to Bayesian Statistical Inference*. Ph.D. Thesis, the University of Western Ontario, Canada.
23. Thabane L, Haq MS (2000) Bayesian prediction from a compound statistical model: An actual application. *Pak J Statist* 16: 110-125.
24. Ali-Mousa MAM, Al-Sagheer SA (2006) Statistical inference for the Rayleigh model based on progressively type-II censored data. *Statistics* 40: 149-157.
25. Raqab MZ (1995) On the maximum likelihood prediction of the exponential distribution based on double type II censored samples. *Pakistan J Statist* 11: 1-10.
26. Raqab MZ (2009) Distribution-free prediction interval for the future current record statistics. *Stat papers* 50: 429-439.
27. Surveillance, Epidemiology and End Results, SEER (2012) Cancer of the Breast-SEER Stat Fact Sheets.
28. Congdon P (2007) Mixtures of spatial and unstructured effects for spatially discontinuous health outcomes. *Computational Statistics and Data Analysis* 51: 3197-3212.
29. Congdon P (2005) Bayesian predictive model comparison via parallel sampling. *Comput Stat Data Anal* 48: 735-753.
30. Achcar JA, Coelho-Barros EA, Cordeiro GM (2009) Beta distributions and related exponential models: A Bayesian approach.