Unilag Journal of Mathematics and Applications,
Volume 2 Issue 1 (2022), Pages 23–36.
ISSN: 2805 3966. URL: http://lagjma.edu.ng

PARAMETRIC MODELLING USING BAYESIAN APPROACH

JULIANA CONSUL*, EVANS OSAISAI, BUNAKIYE JAPHETH, AND JOSEPH ERHO

Abstract. In this paper, we focus on the applicability of a Bayesian analysis to survival time of breast cancer data by assuming that the survival times follow a Weibull distribution. This study determines a method of estimating the model parameters in survival analysis. The proportional hazard model is used to relate the hazard function to the covariate values for an individual. The scale parameter of a Weibull distribution is used to incorporate the covariates of the individual and the linear predictor is expressed as a logarithmic link function of the hazard multiplier. The Bayesian approach to survival analysis is used via the Just another Gibbs sampler (RJAGS) program in R language and R functions was used to calculate the prognostic index as a linear predictor on an index from 0 to 100 which is used for predicting the outcome of the patients on the basis of the clinical information. The posterior summaries of interest which were derived from the posterior distribution are provided. The results from the posterior distribution obtained from this study can be used in the calculation of the risk value of the breast cancer patient. Thus, the risk value helps the researcher to have an assess to the patients exposure to breast cancer. The Parametric model was seen to be a very attractive option of modelling and the ease of interpretation of parameters is of benefit especially for clinicians.

1. Introduction

Survival analysis is one of the most important fields of Statistics in Medicine and Health Sciences. The standard statistical techniques can not be applied to survival data because in most cases, the data may include dealing with "incomplete" or "censored" data [18]. Also, normal distribution are usually inappropriate for analysing survival data because times are always positive, have skewed distribution, the variances depend upon covariates and hence, distributions such as Weibull, Gamma etc are used. Censoring affect the likelihood function. The right censoring which occurs when the failure time is known to be larger than some given time is mostly used [25]. More details on censoring are discussed in

2010 *Mathematics Subject Classification.* Primary: 22E30. Secondary: 58J05.
Key words and phrases. prognosis index, modelling, parameters, variables, rjags.

Submitted: August 10, 2021. Revised: December 15, 2021. Accepted: Februaryl 16, 2022.
* Correspondence.

[17] and [18].

The survival function, hazard function and cumulative hazard function are commonly used to describe survival data. The proportional hazard model is mostly known to explore the relationship between the covariates and survival [8]. It is widely used in the analysis of clinical trials. The baseline hazard is a constant term which corresponds to an intercept and is independent of the covariates. In survival analysis, inferences are made for the effects of explanatory variables and the baseline hazards from life time data through maximum likelihood [16]. There are parametric survival models for which the restrictive assumption of hazard is not required. A parametric survival model is one in which the survival time is assumed to follow a known distribution. Computations in survival analysis are done either using the Bayesian or the traditional frequentist approach. The Bayesian approach incorporates the prior knowledge into the current analysis whereas the frequentist approach does not.

Recently, there has been lots of attention to Bayesian analysis to modelling data. For example, [4] applied a spatial frailty model to infant mortality by assuming a parametric Weibull baseline hazard; [1] discussed a class of fully parametric proportional hazard models in which the baseline hazard is assumed to be a power transform of the time scale and the survival times are assumed to follow a Weibull distribution; [2] summarised some of the most popularly used Bayesian survival models; [14] reviewed some parametric and semi parametric approaches to Bayesian survival analysis with focus on proportional hazard models; [19] analysed survival times of patients treated with two different treatment using Weibull parametric model of Bayesian survival analysis; [6] proposed a Bayesian inference to estimate the treatment effect using a proportional hazard model for right censored data. However, all of these examples focused on the proportional hazard models which are the most popular survival models. In this study, we demonstrate the applicability of a parametric modelling of the survival time of breast cancer data and to justify the applicability to the Bayesian approach. This study determines a way of estimating the model parameters in survival analysis. The covariates of the individual are incorporated into the model using the hazard multiplier and the linear predictor is expressed using a logarithmic link function of the hazard multiplier.

In this research, the Bayesian approach to survival analysis is employed via the Just another Gibbs sampler (RJAGS) program in R language [28], [30]. The prognostic index which is usually used for predicting the outcome in patients on the basis of the clinical information of the patients will also be calculated using R functions. A breast cancer data with large number of patients from Saudi Cancer Registry (SCR) was used for illustration in this research.

[24] reviewed some frequentist methods (such as nonparametric methods, semi parametric and parametric methods) and Bayesian approaches to survival analysis. The Bayesian survival analysis is mostly used in biomedical fields due to the

availability of softwares and the ease of interpretation of the research findings. The Bayesian approach is flexible in terms of the uncertainty about the unknown parameters in the model which are usually expressed through the prior distribution. The prior distribution is a belief about the unknown parameter without reference to the data. Bayesian inference (in form of the posterior distribution) is then required to capture all we know about the parameters by combining the likelihood (from the observed data) and the prior experience. The proper choice of the prior distribution also plays an important role in achieving successful objectives in Bayesian survival analysis since the probability of an event is measured as a degree of belief. The Bayesian approach to inference also treats parameters as random and makes direct statements about them and new samples are drawn from the posterior distribution which maintains stability.

2. Materials and Methods

2.1. **The proportional hazard model.** Let the survival time t_i of an individual i be a realisation of a non-negative random variable T_i with cumulative distribution function (cdf) $F_i(t)$ which is the lifetime distribution function of T_i and is given by

$$F_i(t) = \Pr(T_i < t)$$

and the probability density function $f_i(t)$ is given by

$$f_i(t) = \frac{d}{dt} F_i(t)$$

The survival function $S_i(t)$ of an individual i can be defined as the probability that the individual survives longer than some specified time t where t ranges from 0 to ∞ [26[, [9]. This can be expressed mathematically as

$$S_i(t) = \Pr(T_i \geq t) = 1 - F_i(t) = \int_t^{\infty} f_i(t)dt$$

The Hazard function $h_i(t)$ of an individual i can be expressed mathematically as

$$h_i(t) = \frac{f_i(t)}{S_i(t)} \tag{2.1}$$

The survival function can be estimated using Kaplan-Meier estimator, which is also the default method in most statistical packages [25]. Alternatively, Nelson-Aalen estimator is available to estimate the survival function. This research will use a Bayesian approach to describe the survival data with a Weibull modelling.

The proportional hazard and accelerated life models are both ways of relating the covariates to the survival distribution. In this research, we concentrate on using the proportional hazard model.

The proportional hazard model uses the assumption of proportionality to relate the hazard function to the covariate values for an individual. Suppose that we have M covariates for $m = 1, 2, M$ and n individuals for $i = 1, 2, .., n$. The

covariate vector for the i^{th} individual is denoted by $\underline{X}_i$ and given as $\underline{X}_i = (1, x_{i,1}, x_{i,2}, ...x_{i,M})$. The proportional hazard model assumes that any two individuals i and j with hazard function $h_i(t)$ and $h_j(t)$ at time t and covariate vectors $\underline{X}_i = (1, x_{i,1}, x_{i,2}, ...x_{i,M})$ and $\underline{X}_j = (1, x_{j,1}, x_{j,2}, ...x_{j,M})$ have their hazards related by

$$h_i(t) = \lambda_{i,j} \times h_j(t)$$

where $\lambda_{i,j}$ is a constant and does not depend on t. The proportional hazard model can also be written as:

$$h_i(t) = \lambda_i \times h_0(t) \tag{2.2}$$

where $h_0(t)$ is the baseline hazard function which is a function of time t but does not involve the covariates $\underline{X}_i = (1, x_{i,1}, x_{i,2}, ...x_{i,M})$. The quantity λ_i is the hazard multiplier which depends on the covariates of the individual i but not on the time variable t and hence, the scale parameter λ of a Weibull distribution is used to incorporate the covariates of the individual. The value of $\lambda_i \geq 0$. The linear predictor η_i is expressed using a logarithmic link function of the hazard multiplier and is given as

$$\eta_i = g(\lambda_i)$$

where g is a known function called the link function which must be monotonic and differentiable. We make

$$\eta_i = \underline{X}_i^T \underline{\beta}$$

where $\underline{X}_i^T$ is the transpose of the vector $\underline{X}_i$ and the vector of parameters $\underline{\beta} = (\beta_0, \beta_1, \ldots, \beta_M)^T$
where $x_{i,m}$ is the value of covariate m for the i^{th} individual, β_0 is the baseline parameter and β_m is the covariate effect of the m^{th} covariate.
We have that λ_i is expressed as

$$\begin{aligned} \lambda_i &= \exp\{\eta_i\} \\ &= \exp\left\{\underline{X}_i^T \underline{\beta}\right\} \end{aligned}$$

So, $g(\lambda_i) = \log \lambda_i$.

And so, the linear predictor η_i is given as

$$\log \lambda_i = \eta_i = \beta_0 + \sum_{m=1}^{M} \beta_m x_{i,m} \tag{2.3}$$

The linear predictor is used to construct the prognostic index from the proportional hazard model. Hence, the prognostic index is the main product of a proportional hazard model. The Prognostic models are used for predicting the outcome in patients on the basis of the clinical information of the patient usually before treatment. The prognostic indices can be used to estimate the length of an individuals survival. It helps to take clinical decisions and helps doctors choose an appropriate treatment for the patients. The prognostic indices could help in creating clinical risk groups which stratify patients by the severity of the disease. [29] reviewed current practice in methods used to develop and evaluate the performance of prognostic indices and risk groups from the prognostic models. High

values of the prognostic index indicate a worse prognosis or adverse outcome for the event of interest.

2.2. **Bayesian Inference.** Bayesian inference requires the combination of prior experience (which is in the form of prior probability) and the observed data (which is in the form of a likelihood $L(\theta|Y)$) [16]. The prior beliefs about a parameter θ, with no reference to the data, can be expressed in the form of the probability density function $\pi(\theta)$. The posterior distribution combines the likelihood and the prior which then captures all we know about the parameters. The posterior probability density function for the parameter θ, $\pi(\theta|Y)$ summarises our beliefs about θ after seeing the data, Y. Using the Bayes formula, we have that the posterior distribution would be given as

$$\pi(\theta|Y) \propto \pi(\theta) \times L(\theta|Y)$$

and is expressed as

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}.$$

We will note that Bayesian inference often involves calculations which are analytically intractable. These are typically done using Markov chain Monte Carlo methods (MCMC) [12]. Some of the methods include Metropolis and Metropolis-Hastings algorithm, Gibbs sampler and Metropolis within Gibbs algorithm.
We will discuss the likelihood contribution in survival analysis and the prior distribution about the model parameters.

2.2.1. *The likelihood contribution in survival analysis.* Suppose that we have n individuals with lifetimes governed by a survival function $S(t)$, probability density function $f(t)$ and the i^{th} individual has an observation time t_i. The general form of the likelihood, where some observations are right censored is given as

$$L = \prod_{i \in E} f(t_i) \prod_{i \in C} S(t_i)$$

where C is the set of right censored individuals and E is the set of individuals that had the event. The contribution of a right censored observation to the likelihood is $S(t_i)$. This is the probability that the individual is still alive at time t_i.

We suppose that the lifetime random variable has a Weibull distribution with scale parameter (λ) and shape parameter (α). Then, the probability density function of the i^{th} individual is given by

$$f(t|\lambda_i, \alpha) = \lambda_i \alpha t^{\alpha-1} \exp\{-\lambda_i t^{\alpha}\}$$

The survival function for the individual is given by

$$S(t|\lambda_i, \alpha) = \exp\{-\lambda_i t_i^{\alpha}\}$$

Let $D = (T, X, \delta, M, n)$ where:
n is the number of individuals
M is the number of covariates used in the model
$T = (t_1, t_2, ...t_n)^T$, where t_i is the event or censoring time for the i^{th} individual

$\delta = (\delta_1, ...\delta_n)^T$, where δ_i is the event indicator which indicates whether the individual died or was right censored. We have

$$\delta_i = \begin{cases} 1 & \text{if the individual died} \\ 0 & \text{if censored} \end{cases}$$

X is a n by $(M+1)$ matrix such that the i^{th} row of X with $(1, x_{i,1}, \ldots, x_{i,M})$

From Equation 2.1, $h(t) = \frac{f(t)}{S(t)}$ and we have that $f(t) = h(t)S(t)$. Therefore, the likelihood is

$$\prod h_i^{\delta_i}(t) S_i(t) = \left[\prod_E h_i(t)\right] \left[\prod_{E \cup C} S_i(t)\right]$$

where $E \cup C$ is the set of both censored individuals and individuals that had the event (that is all individuals).

If we suppose that the lifetime random variable has a Weibull distribution with parameters (λ, α), the probability density function of the i^{th} individual is denoted by

$$f(t|\lambda_i, \alpha) = \lambda_i \alpha t^{\alpha-1} \exp\{-\lambda_i t^\alpha\}$$

and the survival function for the individual is given by

$$S(t|\lambda_i, \alpha) = \exp\{-\lambda_i t_i^\alpha\}$$

The likelihood contribution from the data is then given by

$$\begin{aligned} L(\beta, \alpha|D) &= \prod_{i \in E} f(t_i|\lambda_i, \alpha) \prod_{i \in C} S(t_i|\lambda_i, \alpha) \\ &= \prod_{i \in E} \lambda_i \alpha t_i^{\alpha-1} \prod_{i \in E \cup C} \exp\{-\lambda_i t_i^\alpha\} \\ = \left[\prod_{i \in E} \lambda_i\right] \alpha^{n_D} \left[\prod_{i \in E} t_i^{\alpha-1}\right] \exp\left\{-\sum_{i \in E \cup C} \lambda_i t_i^\alpha\right\} \end{aligned} \tag{2.4}$$

where n_D is the number of individuals in E.

2.2.2. *Prior distribution for the coefficients of regression parameters.* A Bayesian analysis requires the specification of prior information about the model parameters by expressing beliefs about the parameters in the form of a probability distribution before we look at the observations. The prior distribution should reflect information about the model parameters. The prior information is often an opinion or subjective belief of an "expert" within the field of investigation from whom information is being elicited. This is appropriate, for example, when the purpose of the analysis is to inform a decision which must be made. In other cases, the purpose may be simply to communicate the results of a scientific investigation. In such cases, one or more "reasonable" prior specification may be used [11].

The structures of prior distributions were constructed depending on the type of variable. We follow the construction of a prior for the coefficient of a quantitative covariate in the context of a general linear model following an example in [10]. We incorporate our prior beliefs into the construction of the covariance matrix

of the parameters. The prior distribution for the covariances will be specified by thinking in terms of the coefficient of determination to find the correlation between the parameters.

2.3. **Bayesian survival modelling.** The Weibull distribution [31] is used because the shape parameter accounts for additional possible hazard shapes and the scale parameter. From the likelihood given in Equation 2.4, the logarithm of the likelihood is given by

$$\ell = \log\left\{L(\underline{\beta}, \alpha|D)\right\} = \sum_{i\in E}\left\{\log\left\{\lambda_i\right\} + (\alpha-1)\log\left\{t_i\right\}\right\} + n_D\log\{\alpha\} - \sum_{i=1}^{n}\lambda_i t_i^{\alpha}$$

The vector of regression parameters is given as

$$\underline{\beta} = (\beta_0, \beta_1, \ldots, \beta_M)^T$$

The prior density for the vector of regression parameters $\underline{\beta}$ could be a multivariate normal distribution $N_{M+1}(\underline{\mu}, V)$ where $\underline{\mu}$ is the vector of prior means given as $\underline{\mu} = (\mu_0, ...\mu_M)^T$ and and V is a $M+1$ by $M+1$ covariance matrix. The prior density for the vector of regression coefficient is then given by

$$(2\pi)^{-M/2}|V|^{-1/2}\exp\left\{-\frac{1}{2}\left[(\underline{\beta}-\underline{\mu})^T V^{-1}(\underline{\beta}-\underline{\mu})\right]\right\}$$

For illustration, we may choose to give α and $\underline{\beta}$ independent prior on the grounds that beliefs about $\underline{\beta}$ are beliefs about the effects of covariates on the hazard function. In general, α and $\underline{\beta}$ need not be independent. A special case where the regression coefficients are independent simplifies the logarithm of the multivariate normal prior density to

$$-\frac{M}{2}\log\{2\pi\} - \frac{1}{2}\sum_{m=0}^{M}\log|V_m| - \frac{1}{2}\sum_{m=0}^{M}\frac{(\beta_m-\mu_m)^2}{V_m}$$

given that $V = \text{diag}(V_0, V_1, ...V_M)$
where V_0 is the variance of the intercept regression coefficient and V_m is the variance of the coefficient of the m^{th} covariate since the coefficients are assumed independent. The prior distribution for α could be a gamma distribution $\alpha \sim \text{Ga}(a, b)$ with density given by

$$\begin{aligned}\pi(\alpha|a,b) &= \frac{b^a}{\Gamma(a)}\alpha^{a-1}\exp\left\{-b\alpha\right\} \\ &\propto \alpha^{a-1}\exp\left\{-b\alpha\right\}\end{aligned}$$

Suppose that we denote the joint prior density of the parameters by $\pi(\underline{\beta}, \alpha)$. The posterior density $\pi(\underline{\beta}, \alpha|Y)$ is then given by

$$\begin{aligned}\pi(\underline{\beta}, \alpha|Y) &\propto \text{prior} \times \text{likelihood} \\ &= \kappa\pi(\underline{\beta}, \alpha)L(\underline{\beta}, \alpha|Y)\end{aligned}$$

where κ is the constant of proportionality. The logarithm of the posterior density is then given as

$$\begin{aligned}\log\{\pi(\underline{\beta},\alpha|Y)\} &= \log\kappa+(a-1)\log\alpha-b\alpha-\frac{M}{2}\log\{2\pi\}-\frac{1}{2}\sum_{m=0}^{M}\log\{V_m\}\\ &\quad-\frac{1}{2}\sum_{m=0}^{M}\frac{(\beta_m-\mu_m)^2}{V_m}+\sum_{i\in E}\{\log(\lambda_i)+(\alpha-1)\log(t_i)\}\\ &\quad+n_D\log(\alpha)-\sum_{i=1}^{n}\lambda_i t_i^{\alpha}\end{aligned}$$

The joint posterior density would not have a closed form but we can simulate from it using MCMC techniques.

2.4. **Application: Bayesian survival modelling to Breast cancer data.** A set of data with large number of patients are available for illustration in this research. The data set was provided by the Saudi Cancer Registry (SCR) of the King Faisal Specialist Hospital and Research Centre. The data was collected from the thirteen administrative regions in the Saudi Kingdom which include Riyadh, Makkah, Madinah, Qassim, Hail, Jouf, Tabouk, Najran, Baha, Asir, Jezan, International and the eastern and northern regions. The data set includes survival time, censoring indicator, sex, age, marital status, tumour details (for example, laterality, grade, extent and topography) for 5432 patients with complete covariate values.

The explanatory variables used in the data set are described as follows:
Age: This variable provides the patient age at diagnosis.
Gender: It refers to patients gender with the value which were "1" for male and "2" for female.
Grade: The grade of a tumor describes how abnormal the tumor cell looked. In our data set, we have used the value "1" for Grade I (well differentiated or low grade), "2' for Grade II (moderately differentiated or intermediate grade), "3" for Grade III (Poorly differentiated or high grade) and "4" for Grade IV (undifferentiated or high grade).
Extent: This variable categorises the breast cancer based on the extent of the disease. In this data set, we have used the value "1" for localised, "2" for regional and "3" for Distant Metastasis.
Laterality: This variable identifies the side of a paired organ or of the body on which the tumor originated. In our data, we have used the value "1" as Bilateral involvement, "2" for Left, "3" for Paired, and "4" for right.
Topography: The variable indicates the site of origin of the tumor or where the tumor arose. The breast halves are divided into quarters or quadrants. In our data set, we have used the value "1" for nipple, "2" for Central portion of breast, "3" for Upper-inner quadrant of breast , "4" for Lower-inner quadrant of breast, "5" for Upper-outer quadrant of breast, "6" for Lower outer quadrant of breast, "7" for Auxillary tail of breast, "8" for Overlapping lesion of breast and "9" for Breast, NOS.

Marital status: In the data set, we have used the value "1" for divorced, "2" for married, "3" for single and "4" for widowed.

In the parametric model, the distribution of the event is specified in terms of unknown parameters which in this case are the coefficient of the covariates or the covariate effects. We apply the Bayesian survival modelling to the Breast cancer data. The covariates and their notations are given in in Table 1.

Table 1. Covariates and notations

Covariates	Notation	Covariates	Notation
Age	x_1	Gender	x_5
Topology	x_2	Marital Status	x_6
Laterality	x_3	Extent	x_7
Grade	x_4		

We will suppose that the overall survival lifetime is Weibull(λ_i, α), where the hazard multiplier λ_i depends on the linear predictor η_i which is used to incorporate the covariates and follows from Equation 2.3 and is further expressed as

$$\eta_i = \beta_0+\beta_1 x_{i,1}+\sum_{k=1}^{9}\beta_{2,k}\delta_{i,2,k}+\sum_{k=1}^{4}\beta_{3,k}\delta_{i,3,k}+\sum_{k=1}^{4}\beta_{4,k}\delta_{i,4,k}+\beta_5 x_{i,5}+\sum_{k=1}^{4}\beta_{6,k}\delta_{i,6,k}+\sum_{k=1}^{3}\beta_{7,k}\delta_{i,7,k} \tag{2.5}$$

where $\delta_{i,j,k} = 1$ if $x_{i,j} = k$ and $\delta_{i,j,k} = 0$ otherwise for $j = 2, 3, 4, 6, 7$.

We construct the prior distribution for the parameters of the regression or linear predictor. A categorical variable with p levels will contribute $p-1$ columns to the design matrix. We follow explanations in [10] to construct the prior means and standard deviations for the parameters of linear predictor as given in Table 2.

The prior distribution constructed in Table 2 are chosen in such a way that will quantify one's prior belief about the likely values for the unknown parameters and these values are independent of the data from the current study.

3. Results

The RJAGS package within the R software was used to run the analysis of the Bayesian model using MCMC. The main advantage of RJAGS in comparison of BUGS model specification language [23] is that the former is easier to setup and works faster. It allows the researcher to provide initial values for the parameters.

The Metropolis-Hastings within Gibbs algorithm was applied using R JAGS software [27]. Following a burn-in of 5000 iterations of the sampler, 100000 iterations were taken. Convergence was checked using two chains starting from very different values. Visual inspection of the trace plots of the covariate parameters showed that the mixing appeared very satisfactory. The posterior numerical summaries of the parameters are also given in Table 3.

The posterior summaries given in Table 3 are estimates of the posterior distribution of the parameters of the model obtained using the MCMC algorithm.

TABLE 2. The prior means and standard deviations for the parameters of the linear predictor

Parameter	prior mean	prior standard deviation
β_0 (baseline parameter)	-1.500	0.400
β_1 (Age parameter)	0.040	0.030
$\delta_{2,1}$	0.000	0.140
$\delta_{2,2}$	0.000	0.077
$\delta_{2,3}$	0.000	0.055
$\delta_{2,4}$	0.000	0.141
$\delta_{2,5}$	0.000	0.158
$\delta_{2,6}$	0.000	0.183
$\delta_{2,7}$	0.000	0.224
$\delta_{2,8}$	0.000	0.316
$\delta_{3,1}$	0.000	0.055
$\delta_{3,2}$	0.000	0.770
$\delta_{3,3}$	0.000	0.141
$\delta_{4,1}$	0.000	0.055
$\delta_{4,2}$	0.000	0.770
$\delta_{4,3}$	0.000	0.141
δ_5 (Gender parameter)	0.050	0.150
$\delta_{6,1}$	0.000	0.055
$\delta_{6,2}$	0.000	0.770
$\delta_{6,3}$	0.000	0.141
$\delta_{7,1}$	0.000	0.055
$\delta_{7,2}$	0.000	0.770
α	1	0.5

These values are used in the calculation of the prognostic index or risk value as will be discussed in Section 3.1.

3.1. **Calculation of the prognostic index for the Weibull survival modelling to the Breast cancer data set.** We recall the formula of the linear predictor in Equation 2.3 as follows

$$\log \lambda_i = \eta_i = \beta_0 + \sum_{m=1}^{M} \beta_m x_{i,m} \tag{3.1}$$

The linear predictor has a linear structure with the posterior values of covariate effects $\underline{\beta} = \beta_0, \beta_1, \ldots, \beta_M$ which are given in Table 3. We can obtain the expectation of the linear predictor by substituting the posterior means into the formula. We have that the expectation of the linear predictor for a new individual $i^{'}$ is

TABLE 3. The posterior means and standard deviations for the parameters of the linear predictor

Parameter	posterior mean	posterior standard deviation
β_0 (baseline parameter)	-1.210	0.172
β_1 (Age parameter)	0.010	0.003
$\beta_{2,1}$ (Topology 1 parameter)	-0.214	0.111
$\beta_{2,2}$ (Topology 2 parameter)	- 0.217	0.113
$\beta_{2,3}$ (Topology 3 parameter)	-0.262	0.120
$\beta_{2,4}$ (Topology 4 parameter)	- 0.144	0.167
$\beta_{2,5}$ (Topology 5 parameter)	- 0.206	0.104
$\beta_{2,6}$ (Topology 6 parameter)	-0.223	0.196
$\beta_{2,7}$ (Topology 7 parameter)	0.963	0.429
$\beta_{2,8}$ (Topology 8 parameter)	0.039	0.106
$\beta_{2,9}$ (Topology 9 parameter)	0.263	0.089
$\beta_{3,1}$ (Laterality 1 parameter)	0.011	0.090
$\beta_{3,2}$ (Laterality 2 parameter)	0.012	0.089
$\beta_{3,3}$ (Laterality 3 parameter)	0.036	0.168
$\beta_{3,4}$ (Laterality 4 parameter)	0.059	0.194
$\beta_{4,1}$ (Grade 1 parameter)	-0.139	0.092
$\beta_{4,2}$ (Grade 2 parameter)	- 0.203	0.070
$\beta_{4,3}$ (Grade 3 parameter)	0.131	0.075
$\beta_{4,4}$ (Grade 4 parameter)	0.211	0.192
β_5 (Gender parameter)	0.031	0.103
$\beta_{6,1}$ (Marital 1 parameter)	-0.075	0.064
$\beta_{6,2}$ (Marital 2 parameter)	0.034	0.088
$\beta_{6,3}$ (Marital 3 parameter)	-0.204	0.097
$\beta_{6,4}$ (Marital 4 parameter)	0.244	0.140
$\beta_{7,1}$ (Extent 1 parameter)	-0.182	0.043
$\beta_{7,2}$ (Extent 2 parameter)	0.161	0.045
$\beta_{7,3}$ (Extent 3 parameter)	0.343	0.049
α	1.379	0.040

$$\log \lambda_{i'} = \eta_{i'} = \beta_{0'} + \sum_{m=1}^{M} \beta_{m'} x_{i',m} \tag{3.2}$$

where $\beta_{0'}$ is the posterior expectation of β_0 and $\beta_{m'}$ is the corresponding posterior expectation of β_m. We might prefer to index the linear predictor in a range $(0, 100)$. We do this by finding $100\Phi^{-1}\left(\frac{\eta_i' - \text{ mean}}{\text{standard deviation}}\right)$ where $\Phi()$ is the standard normal distribution function and mean and standard deviation are the sample mean and sample standard deviation of the values of η for all patients in the breast cancer data set using the posterior means of $\beta_0, \ldots, \beta_M$. For instance, a patient with the covariate vector $\underline{x} = (45, 9, 2, 2, 2, 1, 1)^T$ for the covariates age, topology, laterality, grade, gender, marital status and extent respectively will

have an index of 49. This will mean that the patient has an index of 49 on a scale from 0 to 100 and this is an average risk value.

4. Conclusion

This research has focused on the Bayesian paradigm to parametric survival modelling using Weibull Distribution. The Bayesian inference was used in form of the posterior distribution which captured all we know about the parameters by combining the likelihood (from the observed data) and the prior experience (what we think about the parameters before we see the data) . This approach to inference also treated parameters as random and made direct statements about them. The likelihood contribution of the right censored observations from the data were discussed and thereby simplified. The choice of the prior distribution also played an important role in achieving successful inference in this research. The posterior numerical summaries which were derived from posterior samples of the parameters have been displayed. The expectation of the linear predictor was obtained by substituting the posterior means. The prognosis index for any individual was also calculated by indexing the linear predictor in a range$(0, 100)$ from the proportional hazard model.

The estimation of the parametric model is carried out by assuming a distribution of the survival time. The Parametric model was seen to be a very attractive option of modelling as the hazard functions are of primary interest. The ease of interpretation of parameters may be another benefit especially for clinicians.
The Bayesian approach was flexible in terms of the uncertainty about the unknown parameters in the model which are usually expressed through the prior distribution. This approach is very important as it made use of the available information. The Bayesian approach to inference is transparent in making inferences and the posterior summarises can be used for making predictions about future events. The approach used facilitates the implementation of the analysis of higher dimension data and highly realistic models that account for complicating features. The prognostic index will help the researcher get assess to the exposure of breast cancer. Some of the limitations of the Bayesian are: there is a need for software for making inferences and the choice of prior distribution which should explain what is known before collection of data.

Acknowledgment. The authors are grateful to the reviewers for their comments on the manuscript.

Authors Contributions. Conceptualization: Juliana Consul, Bunakiye Japheth.
Formal analysis: Juliana Consul, Evans Osaisai, Bunakiye Japheth, Joseph Erho.
Methodology: Juliana Consul, Evans Osaisai, Bunakiye Japheth, Joseph Erho.
Resources: Juliana Consul, Evans Osaisai, Bunakiye Japheth, Joseph Erho.
Writing, review and editing: Juliana Consul.

Authors' Conflicts of interest. The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding Statement. This research received no specific grant from any agency.

References

[1] K. Abrams, D. Ashby, D. Errington. *A Bayesian approach to Weibull survival modelsApplication to a cancer clinical trial*. Lifetime Data Analysis 2 (1996), 159–174.

[2] D. Alvares, E. Lzaro, V. Gmez-Rubio, C. Armero. *Bayesian survival analysis with BUGS*. Statistics in Medicine 40 (2021), 2975–3020.

[3] E. Avc17. *Bayesian survival analysis: Comparison of survival probability of hormone receptor status for breast cancer data*. International Journal of Data Analysis Techniques and Strategies 9 (2017), 63–74.

[4] S. Banerjee, D. K. Dey. *Semiparametric proportional odds models for spatially correlated survival data*. Lifetime Data Analysis 11 (2005), 175–191.

[5] S. Banerjee, M. M. Wall, B. P. Carlin. *Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota*. Biostatistics 4 (2003), 123–142.

[6] L. Biard, A. Bergeron, V. Lvy, S. Chevret. *Bayesian survival analysis for early detection of treatment effects in phase 3 clinical trials*. Contemporary Clinical Trials Communications 21 (2021).

[7] C. Chatfield. *The Analysis of Time Series, 6th Edition*. Chapman and Hall (2004).

[8] D. R. Cox. *Regression Models and Life-Tables*. Journal of the Royal Statistical Society B 34(1972), 187–220.

[9] M. J. Crowder. *Multivariate survival analysis and competing risks*. Chapman and Hall/CRC, Boca Raton (2012).

[10] M. Farrow. *MAS 8303, Modern Bayesian inference*. Lecture notes, Newcastle University (2011).

[11] P. H. Garthwaite, J. B. Kadane, A. O'Hagan. *Statistical Methods for Eliciting Probability Distributions*. Journal of the American Statistical Association 100(2005), 680–701.

[12] W. R. Gilks. *Markov chain Monte Carlo*. John Wiley and Sons, Ltd (2005).

[13] P. Groeneboom, G. Jongbloed, B. I. Witte. *Maximum smoothed likelihood estimation and smoothed maximum likelihood estimation in the current status model*. Annals of Statistics 38(2010), 352–387.

[14] J. Ibrahim, M. Chen, D. Sinha. *Bayesian survival analysis*. Journal of The American Statistical Association 99(2005).

[15] J. D. Kalbfleisch, R. L. Prentice. *The Statistical Analysis of Failure Time Data. 2nd edition*. New York: John Wiley and Sons (2011).

[16] P. Groeneboom, G. Jongbloed, B. I. Witte. *Analysis of type I and II error rates of Bayesian and frequentist parametric and nonparametric two-sample hypothesis tests under preliminary assessment of normality*. Computational Statistics 22(2020), 101–109.

[17] J. P. Klein, M. L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data. 2nd edition*. New York, Springer (2010).

[18] J. P. Klein, H. C. Van Houwelingen, J. G. Ibrahim, T. H. Scheike. *Handbook of Survival Analysis*. 1st ed. Boca Raton, FL: Chapman & Hall/CRC Press (2013).

[19] Y. Khan, A. A Khan. *Bayesian survival analysis of regression model using Weibull*. International Journal of Innovative Research in Science, Engineering and Technology 2(12),(2013), 7199–7204.

[20] G. Lawrence, M. Wallis, P. Allgood, I. Nagtegaal, J. Warwick, F. Cafferty, N. Houssami, O. Kearins, N. Tappenden, E. O. Sullivan, S. Duffy. *Population estimates of survival in women with screen-detected and symptomatic breast cancer taking account of lead time and length bias*. Breast Cancer Research Treatment 116,(2009), 179–185.

[21] S. Lee, H. Lim. *Review of statistical methods for survival analysis using genomic data*. Genomics and informatics 17(4),(2019).

[22] D. J. Lunn, A. Thomas, N. Best, D. Spiegelhalter. *WinBUGS: A Bayesian modelling framework: Concepts, structure, and extensibility.* Statistics and Computing 10(4),(2000), 325–337.
[23] D. J. Lunn, A. Thomas, N. Best, D. Spiegelhalter. *The BUGS Book: A practical Introduction to Bayesian Analysis.* Boca Raton, London: CRC/ Chapman & Hall (2013).
[24] R. Vinaitheerthan. Overview of Frequentist and Bayesian Approach to Survival Analysis. Applied Medical Informatics 38(1),(2016), 25–38.
[25] Zhongheng Zhang. Statistical description for survival data. Annals of translational medicine 4(20),(2016), 401–410.
[26] E. T. Lee, J. W. Wang. *Statistical Methods for Survival Data Analysis.* Wiley (2003).
[27] M. Plummer. *JAGS Version 3.3.0 user manual (2012).*
[28] M. Plummer. *http://cran.r-project.org/web/packages/rjags/rjags.pdfl (2013).*
[29] S. Mallett, P. Royston, S. Dutton, R. Waters, D. G.Altaman. Reporting performance of prognostic models in cancer: a review. BMC Medicine 8(21),(2010).
[30] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria (2008).
[31] W. Weibull. A Statistical distribution function of wide applicability. Journal of Applied Mechanics 18,(1951) 293–297.

Juliana Consul*
Department of Mathematics / Computer Science, Niger Delta University, Amassoma, Bayelsa State, Nigeria.
E-mail address: ji.consul@ndu.edu.ng

Evans Osaisai
Department of Mathematics / Computer Science, Niger Delta University, Amassoma, Bayelsa State, Nigeria.
E-mail address: fevansosaisai@gmail.com

Bunakiye Japheth
Department of Mathematics / Computer Science, Niger Delta University, Amassoma, Bayelsa State, Nigeria.
E-mail address: jbunakiye@gmail.com

Joseph Erho
Department of Mathematics / Computer Science, Niger Delta University, Amassoma, Bayelsa State, Nigeria.
E-mail address: joseph.erho@mail.ndu.edu.ng