Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://asr.nsps.org.ng/index.php/asr

African Scientific Reports 5 (2026) 374

African Scientific Reports

Evaluation of some Bayesian parametric survival models with application to hypertensive patients

Dorcas Modupe Okewole[a,*], Sunday Samuel Odebode[a], Olumide Sunday Adesina[b], Adedayo Funmi Adedotun[c]

[a]*Department of Mathematics and Statistics, Redeemer's University, Ede, Nigeria*
[b]*Johannesburg Business School, University of Johannesburg, South Africa*
[c]*Department of Statistics, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria*

Abstract

Bayesian parametric survival models provide a flexible framework for time-to-event analysis, particularly when structural assumptions may enhance efficiency. However, performance in low-event settings remains an important practical consideration. We evaluated three Bayesian parametric survival models: Exponential, Weibull, and Lognormal, using data from 155 hypertensive patients, among whom 14 events (9.1%) were observed. Five clinically relevant covariates were included *a priori* in all models. Posterior inference was obtained via Markov Chain Monte Carlo methods. Model comparison was performed using WAIC (Watanabe–Akaike Information Criterion) and LOOIC (Leave-one-out cross-validation). Sensitivity analyses were conducted under alternative prior specifications to assess robustness. The Bayesian Lognormal model (SE = 50.5, LOOIC = 220.4, WAIC = 220.0) outperformed other models. The study showed that occupation is significantly associated with survival time of hypertension patients across all the models. These findings underscore the need for careful consideration of the choice of the parametric model to employ in survival data analysis. The results also suggest the need for the provision of job-related intervention in people's health regarding hypertension.

DOI:10.46481/asr.2026.5.2.374

Keywords: Parametric models, Survival, Bayesian approach, Proportional hazards, Hypertension.

Article History :
Received: 09 October 2025
Received in revised form: 19 December 2025
Accepted for publication: 27 April 2026
Published: 18 May 2026

1. Introduction

Survival analysis is widely applied in medical research to investigate time-to-event outcomes. The Cox proportional hazards (PH) model is the most frequently used framework; however, its validity depends on the proportional hazards assumption. When this assumption is violated, accelerated failure time (AFT) models provide a useful alternative by modelling survival time directly.

Previous studies on hypertension survival have highlighted the limitations of PH models due to time-dependent covariates [1]. Parametric models as accelerated failure time models are known to be effective in analyzing survival data and produce better performance in most cases [2–5]. Bayesian approaches have also been shown to improve inference in complex survival settings [6]. The present study evaluates three parametric AFT models: Exponential, Weibull, and Lognormal within a Bayesian framework, applied to hypertensive patients in Nigeria.

*Corresponding author Tel. No.:
Email address: okewoled@run.edu.ng (Dorcas Modupe Okewole)

The objectives are: (i) to identify the optimal parametric survival model for this dataset, (ii) to determine which covariates significantly affect survival time, and (iii) to assess the robustness of Bayesian AFT models under conditions of limited events. This work contributes to methodological understanding of Bayesian survival modelling and provides insights into occupation-related differences in hypertensive survival outcomes.

Bayesian methods allow incorporation of prior information, which yields full posterior distributions for model parameters when combined with observed data, and enable probabilistic interpretations. Moreover, Bayesian survival models accommodate complex data structures such as time-dependent effects, censoring, and clustering while allowing reliable probabilistic predictions [7, 8]. Although Bayesian models also have strength in their applicability to smaller datasets, they remain constrained by sparse events unless informative priors are carefully specified. In the case of a small number of observed failures relative to the number of model parameters in survival analysis, the likelihood may not contain sufficient information concerning the parameters of interest. The limited information captured in the likelihood makes posterior estimates highly sensitive to the choice of prior. Consequently, it is often required to make careful prior specifications together with sensitivity analysis to ensure robust inference and cautious interpretations [7, 8]. We therefore emphasize both the potential and the limitations of Bayesian survival analysis in this context.

2. Data and methodology

2.1. Data source

Data were obtained retrospectively from Okemesi General Hospital (OGH), Ekiti State, Nigeria, covering hypertensive patients registered between January 1, 2017 and December 31, 2023. Ethical clearance was granted by OGH.

2.2. Study population

A total of 155 patients met inclusion criteria. Of these, 14 experienced the event of interest (death), while 141 were right-censored (transfer, dropout, or end of study). This low event rate (9.1%) represents a key limitation for inference and is explicitly acknowledged in our analysis.

2.3. Variables

The response variable was survival time (days from diagnosis to event or censoring). The covariates included:

(i) Age (years)
(ii) Gender (male/female)
(iii) Marital status (single/married/widowed)
(iv) Occupation (formal/informal/dependent)
(v) Diabetes mellitus status (yes/no).

The coding of variables is given as follows:

- Age – measured in years
- Diabetes Mellitus: Yes = 0, No = 1
- Status: Censored = 0, Dead = 1
- Gender: Male = 1, Female = 0
- Marital status: Single = 0, Married = 1, Widowed = 2
- Occupation: Formal = 0, Informal = 1, Dependents = 2

2.4. Methodological considerations

Given the limited number of events, model complexity was carefully balanced against identifiability. Bayesian AFT models were employed to estimate covariate effects on survival time. Sensitivity analyses with alternative priors were conducted to assess robustness. The Kaplan–Meier estimator was used for descriptive survival curves, and Cox PH assumption checks were performed. Violations of the PH assumption motivated the adoption of AFT models. Model comparison employed WAIC and LOOIC, with diagnostic measures (Pareto-k values) reported to ensure reliability.

3. Method of data analysis

3.1. Descriptive statistics

1. Based on the research variables in OGH, the frequency distribution table provided in Excel and software by R Core Team [9] was utilized to compile the data obtained from the health facility.
2. The Kaplan–Meier technique was employed to estimate and display the survival curves. A formula for the product limit is used to compute the predicted survival probabilities:

$$S_{t+1} = S_t\left(\frac{N_{t+1} - D_{t+1}}{N_{t+1}}\right), \tag{1}$$

where:

N_t = number of participants alive at time t,
D_t = total number of events (e.g. deaths) occurring at time t,
S_t = cumulative survival probability, representing the proportion of the population that survives beyond interval t.

3.2. Survival function comparison

To ascertain whether variations in survival periods existed among the groups of explanatory variables under examination, Kaplan–Meier plots were utilized.

The hypotheses tested are:

H_0 : There is no variation in the survival curves.
H_1 : There is variation in the survival curves.

3.3. Parametric regression models

When examining survival time data, parametric models provide an alternative to the popular Proportional Hazards (PH) models in survival analysis. In parametric models, the direct effects of explanatory variables on survival time are measured.

In this study, three parametric distributions were used to model the survival time variable in order to identify risk factors associated with mortality from hypertension. Parametric survival models are also important in Bayesian survival analysis because many Bayesian investigations in practice are conducted using these models. Parametric modeling provides relatively simple approaches for modeling and analysis.

The details of the Exponential, Weibull, and Lognormal parametric models can be found in the literature and are briefly defined below.

3.3.1. Lognormal distribution

The lognormal distribution is useful in modeling continuous random variables that are greater than or equal to zero. The lognormal distribution is also useful in modeling data that would otherwise be considered normally distributed except for the fact that they may be skewed [10]. Such skewness occurs frequently when means are low, variances are large, and values cannot be negative. Appropriately estimating the parameters of the lognormal distribution is important in the study of these and related subjects.

The lognormal distribution exists in both two-parameter and three-parameter forms. The density function for the two-parameter lognormal distribution is given by

$$f(T \mid \mu, \sigma^2) = \frac{1}{T\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(T) - \mu)^2}{2\sigma^2}\right]. \tag{2}$$

Due to its close relationship with the normal distribution, $\ln(T)$ is normally distributed whenever T follows a lognormal distribution. The parameter μ may therefore be interpreted as the mean of the logarithm of the random variable, while σ represents the standard deviation of the logarithm of the random variable. Additionally, μ is regarded as a scale parameter, whereas σ is regarded as a shape parameter of the lognormal density function.

The hazard and survival functions are respectively given as

$$h(\mu, \sigma) = \frac{1}{T\sigma}\phi\left(\frac{\ln(T) - \mu}{\sigma}\right), \tag{3}$$

$$S(\mu, \sigma) = 1 - \Phi\left(\frac{\ln(T) - \mu}{\sigma}\right). \tag{4}$$

3.3.2. Exponential model

The exponential model is the most fundamental parametric model in survival analysis. The exponential distribution assumes that small values occur more frequently than large values. Consequently, it can be used to model failure times. The density function $f(t)$, survival function $S(t)$, and hazard function $h(t)$ are respectively given as [11]:

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0,\ \lambda > 0, \tag{5}$$

$$S(t) = e^{-\lambda t}, \tag{6}$$

$$h(t) = \lambda, \tag{7}$$

where λ is the scale parameter and represents the constant hazard rate.

3.3.3. Weibull model

The Weibull model is one of the most important parametric models because its hazard function can be increasing, decreasing, or constant depending on the values of the model parameters [11]. The density function of the two-parameter Weibull distribution is given by

$$f(t) = a\lambda(\lambda t)^{a-1} \exp\left[-(\lambda t)^a\right]. \tag{8}$$

The survival and hazard functions are respectively given as

$$S(t) = \exp\left[-(\lambda t)^a\right] \tag{9}$$

and

$$h(t) = a\lambda(\lambda t)^{a-1}, \tag{10}$$

where λ is the scale parameter and a is the shape parameter.

3.4. Survival regression models

Regression models mainly concentrate on the relationship between covariates (predictor variables or risk factors) and time-to-event outcomes. Two primary frameworks are commonly used:

1. Proportional Hazards (PH) models, such as the Cox regression model, which estimate hazard ratios and assume proportionality of hazards over time.
2. Accelerated Failure Time (AFT) models, which directly model survival time by expressing $\log(t)$ as a linear function of covariates.

When the proportional hazards assumption is violated, AFT models provide a useful alternative because they quantify covariate effects in absolute time units (days, months, or years) rather than relative hazard ratios. In this study, AFT models were adopted after diagnostic checks indicated violations of the proportional hazards assumption.

Parametric proportional hazard models. Let the covariate vector be expressed as

$$X = (x_1, x_2, \ldots, x_q)^T.$$

If $h(t \mid x)$ represents the hazard function for a subject at time t, then according to Cox, the proportional hazard model is given by

$$h(t \mid x) = h_0(t) \exp\left(\beta_1 x_1 + \cdots + \beta_q x_q\right), \tag{11}$$

where $h_0(t)$ represents the baseline hazard function, and $\beta_1, \beta_2, \ldots, \beta_q$ are unknown regression coefficients to be estimated.

Accelerated failure time models. In certain situations, the Accelerated Failure Time (AFT) model serves as an alternative to the Proportional Hazards (PH) model and is equally suitable for analyzing survival time data. In AFT models, the direct effect of explanatory variables on survival time is examined rather than the hazard function itself. This makes interpretation easier because the regression parameters directly measure the effect of variables on survival time.

The natural logarithm of survival time in the AFT model can be expressed as

$$\log(t) = X^T\beta + \sigma\varepsilon. \tag{12}$$

Specifically, the fitted model in this study is given as

$$\begin{aligned}\log(t) = \beta_0 + \beta_1 I(MS) + \beta_2 I(G) + \beta_3 \text{Age} \\ + \beta_4 I(O) + \beta_5 I(DM) + \sigma\varepsilon,\end{aligned} \tag{13}$$

where:

$$\begin{aligned} t_i &: \text{death time for individual } i, \\ \beta_0 &: \text{intercept}, \\ \beta_1,\beta_2,\beta_3,\beta_4,\beta_5 &: \text{regression coefficients}, \\ I(MS) &: \text{indicator variable for marital status categories}, \\ I(G) &: \text{indicator variable for gender categories}, \\ I(O) &: \text{indicator variable for occupation categories}, \\ I(DM) &: \text{indicator variable for diabetes mellitus categories}, \\ \varepsilon &: \text{error term}, \\ \sigma &: \text{scaling parameter}. \end{aligned}$$

The effect of the AFT model is interpreted as a shift in the time scale represented by the function $\exp(x_i\beta)$. Survival time increases or decreases depending on whether this factor is greater than or less than 1. The term "accelerated failure time" refers to an expected waiting time until failure that is either accelerated or decelerated by the covariates. Because AFT models predict survival time while proportional hazards models predict hazard, the signs of parameter estimates in AFT models are opposite to those in equivalent proportional hazards models. One major advantage of the AFT approach is that the effects of covariates are expressed in absolute time units such as days, months, or years rather than relative measures such as hazard ratios. The principal measure of association obtained from AFT models is the acceleration factor, which allows assessment of how different factors influence survival time.

3.5. Bayesian survival analysis

Bayesian methods are often preferred in survival analysis because they combine prior knowledge with observed data. Compared to frequentist approaches, Bayesian statistics offer several advantages:

1. Probability is defined in a more intuitive way.
2. Complex models that are difficult for traditional methods can be evaluated.
3. Large sample sizes are not required.
4. Prior knowledge and uncertainty can be incorporated and updated.

In this framework, model parameters are treated as random variables with prior distributions, while the data remain fixed. Fitting survival models can be challenging, but Gibbs sampling and Markov Chain Monte Carlo (MCMC) methods simplify estimation. However, MCMC methods can be slow to converge, and Gibbs sampling requires conjugate priors, thereby limiting flexibility. Modern tools such as BRMS (Bayesian Regression Models using Stan) and Stan's No-U-Turn Sampler (NUTS) improve efficiency by achieving faster convergence even for high-dimensional models without requiring conjugate priors.

A simple Bayesian regression model is given by

$$y_i \sim N(\beta_0 + \beta_i x_i, \sigma^2), \tag{14}$$

where

$$\begin{aligned} y_i &: \text{response variable}, \\ x_i &: \text{independent variables}, \quad i = 1, 2, \ldots, \\ \beta_0,\beta_i &: \text{regression coefficients}, \\ \sigma^2 &: \text{error variance}. \end{aligned}$$

3.6. Stan-based bayesian regression models (BRMS)

Stan is a probabilistic programming language designed for Bayesian inference, offering flexible tools to define complex models and perform efficient sampling using advanced MCMC methods. It is particularly effective for hierarchical, mixture, and high-dimensional models. Users can specify priors, likelihoods, and data relationships using intuitive syntax, while Stan explores posterior distributions with Hamiltonian Monte Carlo (HMC) and its extension, the No-U-Turn Sampler (NUTS). Details are provided in [12, 13]. The NUTS algorithm improves efficiency by:

1. Automatically tuning step sizes and leapfrog steps.
2. Producing accurate samples that avoid random walk behavior.
3. Achieving faster convergence in high-dimensional parameter spaces.

The BRMS package in R builds Bayesian multilevel models using Stan. Posterior expectations and standard deviations can be obtained from posterior marginals for each model component. Convergence is assessed by running multiple Markov chains with different starting values and checking consistency across them. Diagnostic tools such as trace plots, autocorrelation plots, and the Gelman–Rubin statistic are used to confirm convergence and ensure that the chains properly explore the target posterior density.

Bayesian regression using Stan/BRMS generally involves the following steps:

1. Model specification: Define likelihood functions (e.g., regression models) together with prior distributions.
2. Sampling: Use NUTS to efficiently sample from the posterior distribution, thereby obtaining parameter estimates and uncertainty measures.

3.7. Prior

The initial specification of the study adopted diffuse Normal$(0, 100)$ priors for the regression coefficients. However, the limited number of observed events rendered these priors inadequate because they produced unstable posterior estimates. Consequently, the study adopted a weakly informative Normal$(0, 7)$ prior distribution and assessed different specifications of the prior variance, including 2.5, 5, 7, and 10, through prior sensitivity analysis.

3.8. Diagnostic and model selection

Comparing and selecting models is crucial for identifying the most suitable model among several alternatives consistent with the observed data. Using leave-one-out cross-validation (LOO-CV) as the basis for model comparison, this study employed a more comprehensive Bayesian criterion known as the Watanabe–Akaike Information Criterion (WAIC).

Leave-one-out cross-validation (LOO-CV). Model comparison was conducted using Bayesian LOO-CV, which estimates out-of-sample predictive accuracy based on pointwise log predictive densities. Specifically, the leave-one-out information criterion (LOOIC) was computed using Pareto-smoothed importance sampling (PSIS-LOO), as described in Ref. [8]. Diagnostic measures, including Pareto-k values, were reported to assess the reliability of the importance sampling weights. Models with Pareto-k values greater than 0.7 were flagged for potential instability, and alternative refitting strategies were considered when necessary.

The LOOIC is given by

$$\text{LOOIC} = -2\,(\text{LPD} - p_{\text{LOO}})\,, \tag{15}$$

where

$$\begin{aligned} \text{LPD} &: \text{log pointwise predictive density,} \\ p_{\text{LOO}} &: \text{effective number of parameters estimated using LOO-CV.} \end{aligned}$$

Watanabe–akaike information criterion (WAIC). The Watanabe–Akaike Information Criterion (WAIC) is a fully Bayesian, pointwise criterion computed per observation and considered more accurate than earlier criteria. WAIC does not assume approximately normal posterior distributions. The WAIC is defined as

$$\text{WAIC} = 2P_{\text{WAIC}} - 2\,\text{lppd}, \tag{16}$$

where

$$\begin{aligned} P_{\text{WAIC}} &: \text{effective number of parameters (WAIC style),} \\ \text{lppd} &: \text{log pointwise predictive density.} \end{aligned}$$

4. Results and discussion

4.1. Descriptive summary

The data obtained from the 155 patients under follow-up consisted of 100 (64.5%) female and 55 (35.5%) male patients. Among the patients included in the study, 14 (9.1%) experienced the event, while the remaining 141 (90.9%) observations were censored. Approximately 91% of the patients were right censored, while only 9% experienced the event of interest. In terms of marital status, about 141 (90.9%) of the patients were married and living with their spouses, whereas only 14 (9.1%) were widowed, with death proportions of 11 (7.7%) and 3 (25%), respectively. A total of 111 patients were aged 60 years and above, representing 71.6% of the study population, followed by 43 (27.7%) patients within the age range of 30–59 years.

Furthermore, 131 (84.5%) of the patients were diagnosed with diabetes mellitus, and only 3 (2.3%) of these experienced the event. Also, 91 (58.7%) patients were engaged in informal occupations such as farming and trading, while 52 (33.5%) were dependents at the time of the study. The remaining 12 (7.7%) patients had formal occupations. The proportion of deaths varied across occupational categories, with the highest proportion (15.4%) observed among dependent patients. It is noteworthy that none of the patients engaged in formal occupations experienced the event during the study period. The summary statistics are presented in Table 1.

4.2. The kaplan–meier estimate for the covariates in the model

As shown in Figure 1, the overall survival rate at the end of the first month was approximately 94.8%, and by the end of the study period of 84 months, it was 87.5%. The 95% confidence interval at the first 30-day point was (89.2%, 97.4%). Figure 2 indicated that hypertensive patients who were married and living with their spouses had a higher probability of survival than patients who were not living with their spouses, which could possibly be due to separation by divorce or death. The probability of survival decreased for patients who were widowed or divorced. Figure 3 showed that hypertensive patients with formal occupations had a greater likelihood of survival compared to patients in other occupational categories. The survival curve for patients with informal occupations was higher than that of patients who were dependents. The likelihood of survival decreased among dependent hypertensive patients.

4.3. Cox–PH assumption checks

The proportional hazards (PH) assumption can be assessed using statistical tests and graphical diagnostics based on the scaled Schoenfeld residuals. The Schoenfeld residuals are theoretically independent of time. Therefore, a systematic pattern observed over time in graphical displays contradicts the PH assumption. The `cox.zph()` function in the `survival` package provides a convenient approach for verifying whether the proportional hazards assumption holds for each covariate in a fitted Cox regression model. The `cox.zph()` function examines the correlation between the scaled Schoenfeld residuals and time for each covariate in order to determine whether the residuals are independent of time. It also performs both global and covariate-specific tests.

The PH assumption is supported when there is no statistically significant correlation between the residuals and time, whereas a significant association suggests violation of the assumption. The hypotheses tested are:

$$H_0 : \text{All variables satisfy the PH assumption,}$$
$$H_1 : \text{One or more variables violate the PH assumption.}$$

From the global test results presented in Table 2, the Cox proportional hazards assumption failed due to the significance of the global p-value. When modeling survival data, proportional hazards models are the most commonly used approaches. However, accelerated failure time (AFT) models or parametric regression models provide an alternative when the PH assumption is violated. The proportional hazards assumption was violated, as evidenced by the systematic departures from a horizontal line in the scaled Schoenfeld residual plots against transformed time shown in Figure 4 for the covariates. Since the findings indicated that the Cox proportional hazards model was inadequate for fitting the data, parametric regression models were subsequently considered.

4.4. Bayesian survival analysis

The hypertension data set for survival time, where hypertensive individuals were observed, was analyzed using three parametric AFT models. Let β denote the vector of regression coefficients for the covariates in the study, and let β_0 represent the intercept term. Since there were five covariates in the study, all regression coefficients were assigned normal priors with mean 0 and variances in the categories 2.5, 5, 7, and 10 to represent weakly informative priors.

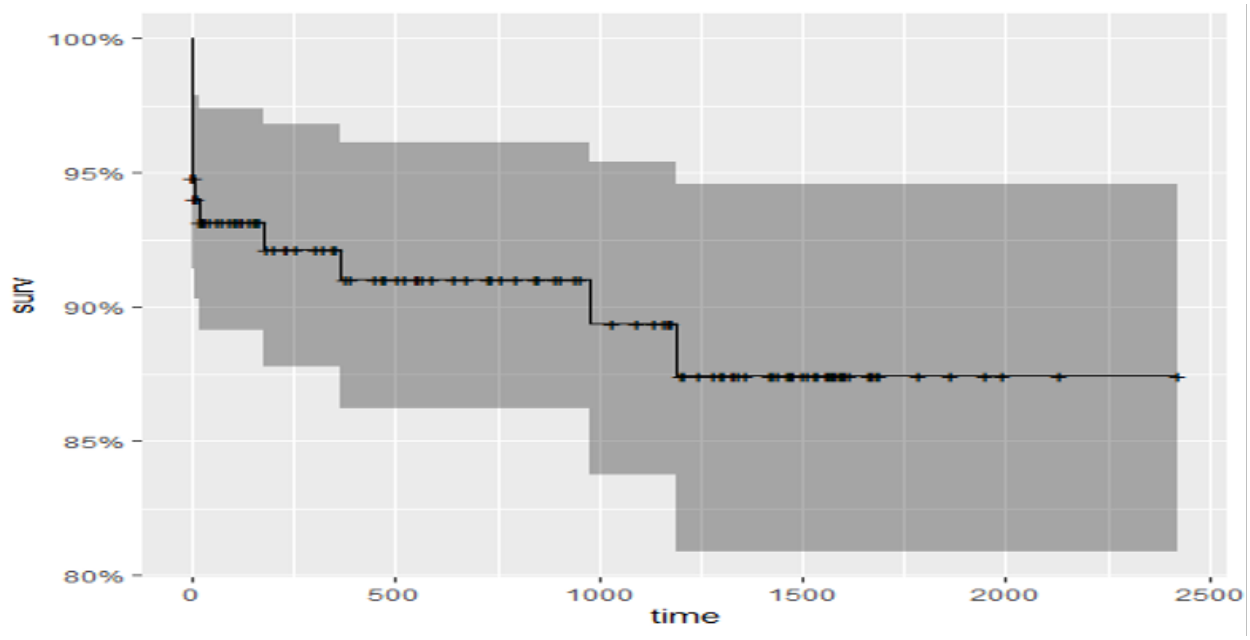

Figure 1: Plot of overall survival function.

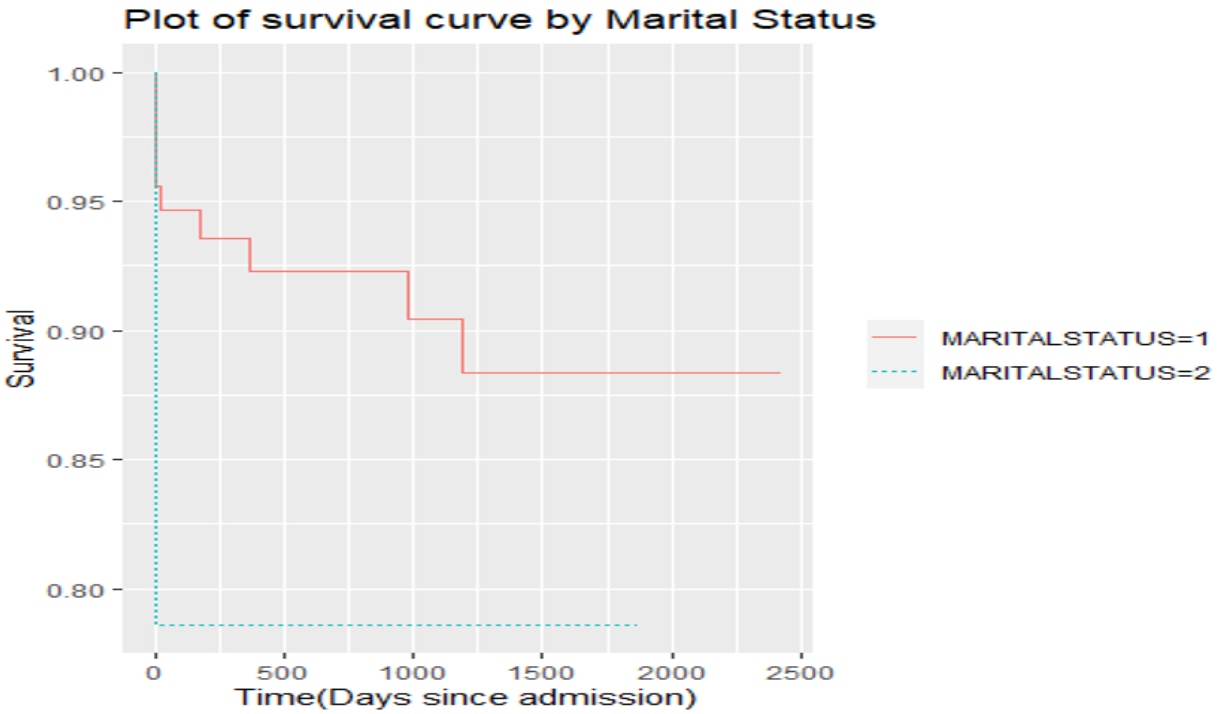

Figure 2: Plot of survival curve by marital status.

4.5. Parametric model comparison results

Table 3 presents the results of the BRMS model comparison study for the three parametric models (Lognormal, Weibull, and Exponential) fitted to the hypertension data set. The models were compared using the Leave-One-Out Information Criterion (LOOIC), the Watanabe–Akaike Information Criterion (WAIC), and their corresponding Pareto-*k* diagnostic values. The model with the smallest criterion value was considered the best-fitting model. The model performance measures are presented in Table 3. Focusing on Table 3, the models providing superior performance relative to the alternatives are indicated by boldface values. The Bayesian lognormal AFT model, with Pareto-*k* values less than 0.7, LOOIC = 220.4, and WAIC = 220.0, was identified as the best model for the survival times of hypertensive patients because it yielded the smallest values among the competing models. Table 4 presents the

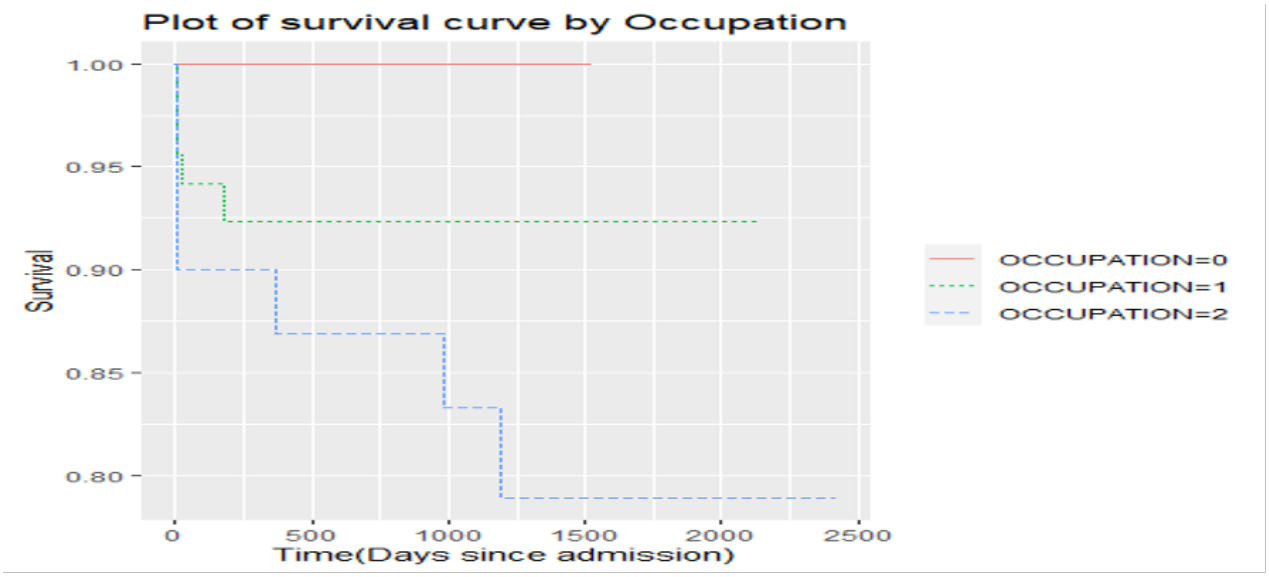

Figure 3: Plot of survival curve by occupation.

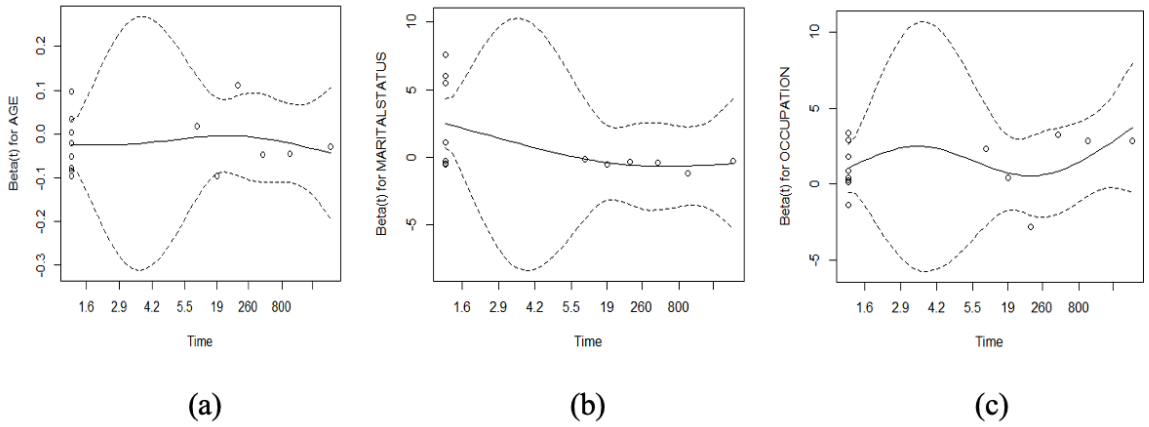

Figure 4: Scaled-Schoenfeld residuals plotted against follow-up time. (a) age (b) marital status (c) occupation.

Table 1: Descriptive summary of the data.

Variable	Mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Marital Status	1.09	0.29	1	1.00	0.00	1	2	1	2.83	6.05	0.02
Gender	0.35	0.48	0	0.32	0.00	0	1	1	0.60	-1.65	0.04
Age	66.57	14.89	66	67.02	17.79	26	97	71	-0.27	-0.63	1.20
Occupation	1.26	0.59	1	1.30	0.00	0	2	2	-0.13	-0.55	0.05
DM	0.15	0.36	0	0.07	0.00	0	1	1	1.89	1.58	0.03

parameter estimates obtained from the Bayesian lognormal model. The findings suggest that the occupation of hypertensive patients has a significant effect on their survival time. Outcomes from the classical lognormal model also gave an indication of the result obtained through the Bayesian model, although at the borderline with marital status as an additional covariate (Table 5).

Posterior sampling and diagnostics in BRMS. The survival model was fitted using four Markov chains, each with 2000 iterations. The first 1000 iterations in each chain were used as warm-up (burn-in), leaving a total of 4000 posterior samples for statistical inference.

Parameter summaries typically include the posterior mean (estimate) and standard deviation, together with the 95% credible

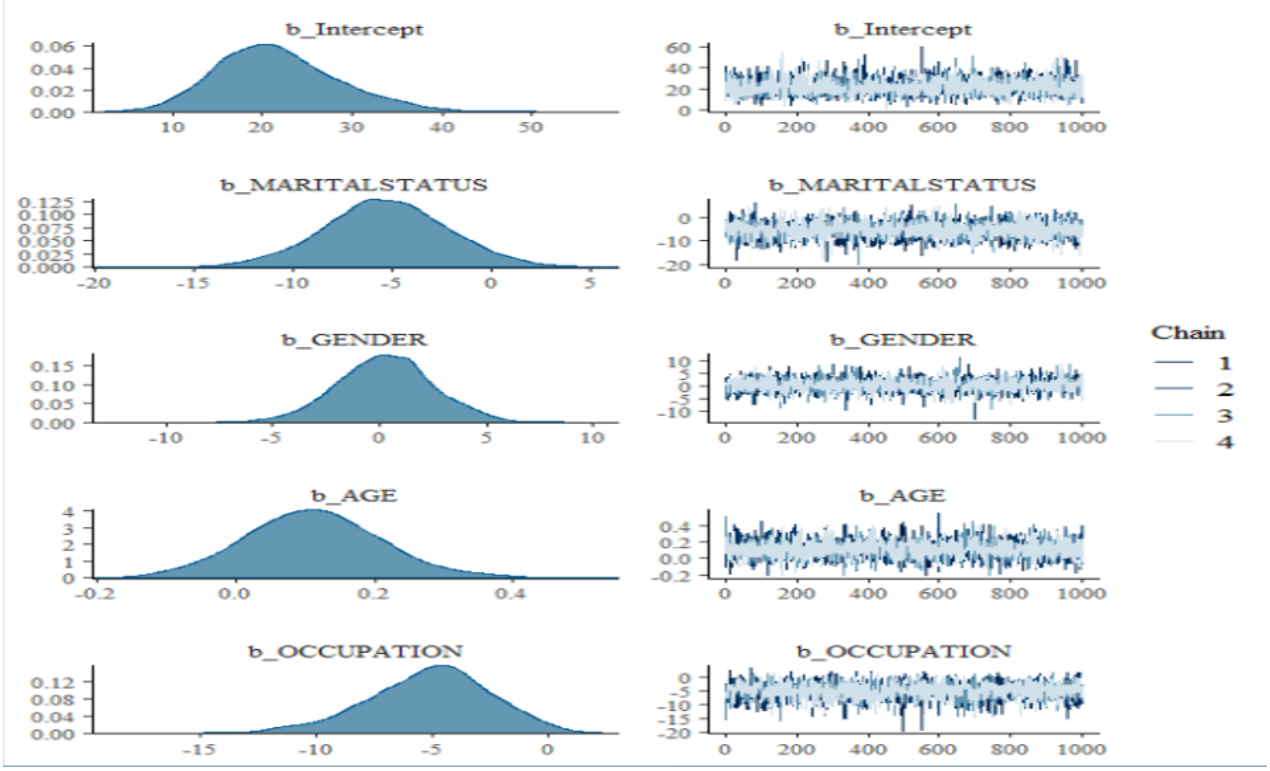

Figure 5: The trace plot showing the Markov chains for the covariates.

Table 2: Cox PH model assumption test.

Covariates	Chi-square	Df	P-value
Marital status	2.240	1	0.134
Gender	4.907	1	0.027
Age	0.158	1	0.691
Occupation	0.500	1	0.479
DM	3.031	1	0.082
Global	12.304	5	0.031

Table 3: Comparison of Bayesian AFT models.

Information Criterion	Exponential	Weibull	Lognormal
Pareto K-values	< 0.7	< 0.7	$< \mathbf{0.7}$
LOOIC	289.0	224.8	220.4
WAIC	287.3	224.2	220.0

intervals (lower 95% CI and upper 95% CI) used to assess parameter significance.

The following model diagnostics were also considered:

- Bulk_ESS and Tail_ESS: These measure the effective sample size and indicate how many independent samples would provide the same level of precision as the correlated MCMC draws.
- Rhat: This is a convergence diagnostic statistic. Values close to 1.00 indicate that the Markov chains have mixed well and converged successfully.

The results presented in Table 4 showed that $\widehat{R} = 1.00$, indicating excellent convergence across all chains.

Table 4: The result for Bayesian lognormal AFT model using BRMS method.

Variable	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Population-Level Effects:							
Intercept	20.86	6.50	9.85	35.09	1.00	4013	2306
MARITALSTATUS	-4.95	2.88	-10.90	0.68	1.00	3752	2787
GENDER	0.33	2.35	-4.31	5.17	1.00	4056	2506
AGE	0.10	0.10	-0.08	0.30	1.00	3097	2722
OCCUPATION	-4.72	2.50	-10.07	-0.13	1.00	2945	2526
DM	-2.07	2.68	-7.16	3.19	1.00	4407	3181
Family Specific Parameters:							
σ	7.46	1.56	5.11	11.17	1.00	1396	2214

Table 5: The result for classical lognormal model using survival function.

Variable	Value	Std. Error	z	p
(Intercept)	22.1667	6.8239	3.25	0.0012
MARITALSTATUS	-6.1829	3.1612	-1.96	0.0505
GENDER	0.1475	2.3728	0.06	0.9504
AGE	0.1243	0.0987	1.26	0.2080
OCCUPATION	-5.5307	2.8800	-1.92	0.0548
DM	-2.7271	2.8156	-0.97	0.3328
Log(scale)	1.9897	0.2200	9.04	$< 2 \times 10^{-16}$
Scale = 7.31				
Log Normal distribution				
Loglik(model)= -103.5 Loglik(intercept only)= -107.9				
Chisq= 8.97 on 5 degrees of freedom, $p = 0.11$				
Number of Newton-Raphson Iterations: 5				
$n = 155$				

Table 6: Comparison of classical and bayesian estimates under different prior specifications.

Parameter	Classical (MLE)	Bayesian N(0,2.5)	Bayesian N(0,5)	Bayesian N(0,7)	Bayesian N(0,10)
Intercept	22.17 (SE 6.82)	17.81 (CI 8.09–29.84)	20.07 (CI 9.38–32.81)	20.86 (CI 9.85–35.09)	21.96 (CI 10.49–36.49)
Marital Status	-6.18 (p=0.0505)	-2.35 (CI -5.96–1.57)	-4.21 (CI -9.17–1.01)	-4.95 (CI -10.90–0.68)	-5.45 (CI -11.51–0.50)
Gender	0.15 (p=0.95)	0.35 (CI -2.90–3.66)	0.37 (CI -3.90–4.68)	0.33 (CI -4.31–5.17)	0.27 (CI -4.48–4.83)
Age	0.12 (p=0.21)	0.05 (CI -0.12–0.20)	0.08 (CI -0.09–0.25)	0.10 (CI -0.08–0.30)	0.11 (CI -0.08–0.31)
Occupation	-5.53 (p=0.0548)	-2.58 (CI -5.91–0.58)	-4.06 (CI -8.66–0.11)	-4.72 (CI -10.07–0.13)	-5.11 (CI -11.06–0.29)
DM	-2.73 (p=0.33)	-1.04 (CI -4.54–2.58)	-1.79 (CI -6.67–3.02)	-2.07 (CI -7.16–3.19)	-2.33 (CI -7.91–3.19)
Scale (σ)	7.31	7.03 (CI 4.84–10.53)	7.30 (CI 5.00–10.59)	7.46 (CI 5.11–11.17)	7.58 (CI 5.17–11.37)
Pareto-k	–	< 0.5	< 0.5	< 0.7	< 0.7

The trace plot shown in Figure 5 illustrates the convergence behavior of the Markov chains. The chains oscillate around a stable equilibrium region, with posterior values plotted on the y-axis against iteration number on the x-axis, demonstrating good mixing and convergence. Results presented in Table 6 show that the Bayesian estimates for the intercept increased steadily as the priors became less restrictive, rising from 17.81 under the tighter prior specification to 21.96 under the wider prior Normal$(0, 10)$, which is closer to the classical estimate of 22.17. The tighter prior shrank the estimates of marital status toward zero. However, as the priors widened, the estimates moved closer to the classical estimate, which exhibited borderline significance (-6.18, $p \approx 0.05$). The Bayesian credible intervals did not include zero for any prior specification, suggesting that the Bayesian approach produced more stable estimates.

The effect of occupation was initially non-significant (-2.58, with credible intervals including zero), but became stronger and eventually excluded zero under the wider priors Normal$(0, 7)$ and Normal$(0, 10)$, approaching the classical estimate, which also

Table 7: Result in terms of AF with N(0,7) prior.

Covariate	Estimate (β)	AF = exp(β)	95% CI for AF	Interpretation
Intercept	20.86	–	–	Baseline log-time, not interpreted as AF.
MARITAL STATUS	-4.95	0.007	(0.000018, 1.97)	Suggests much shorter survival (AF $\ll$ 1), but CI includes 1; effect is uncertain.
GENDER	0.33	1.39	(0.013, 175.9)	Suggests longer survival, but CI is extremely wide; effect is uncertain.
AGE	0.10	1.11	(0.92, 1.35)	Very close to 1; CI includes 1, indicating essentially no credible effect.
OCCUPATION	-4.72	0.009	(0.000043, 0.88)	Strong evidence of shorter survival (AF $<$ 1); CI excludes 1, indicating a credible effect.
DM (Diabetes)	-2.07	0.13	(0.00077, 24.3)	Suggests shorter survival, but CI is wide and includes 1; effect is uncertain.

showed borderline significance (-5.53, $p = 0.0548$). On the other hand, age, gender, and diabetes mellitus (DM) showed no statistically significant effects under any prior specification or under the classical approach. These findings suggest that the parameters representing the effects of age, gender, and diabetes mellitus are robust to prior specification, whereas marital status and occupation are comparatively more sensitive. Model fit assessment using both the Leave-One-Out (LOO) criterion and the Watanabe–Akaike Information Criterion (WAIC) demonstrated high stability across the different prior distributions, as the LOOIC values remained close to 220. This indicates that prior selection did not substantially affect predictive performance.

The robust parameters, namely age, gender, and diabetes mellitus, consistently remained non-significant under both classical and Bayesian approaches. In contrast, the estimates for marital status and occupation varied according to the chosen prior variance. Wider priors yielded estimates closer to the classical estimates and approached borderline significance levels. The intercept and scale parameters were particularly sensitive to prior changes and gradually converged toward the classical estimates as the prior distributions became wider. In summary, the Bayesian results were sensitive to prior choice for marital status and occupation but remained robust for age, gender, and diabetes mellitus. As the priors widened, the Bayesian estimates converged toward the classical results, indicating that the observed data were sufficiently informative to overcome strong prior shrinkage.

4.6. Bayesian lognormal analysis

The study adopted the use of acceleration factors together with the 95% credible intervals from the Bayesian accelerated failure time estimates for interpretation of the results. According to Ref. [14], Bayesian methods allow the incorporation of prior information, provide complete posterior distributions of model parameters, and can be advantageous when dealing with relatively small sample sizes, although the results still depend on the specification of priors and the quality of the data. In the lognormal accelerated failure time (AFT) model, the acceleration factor is obtained as the exponential function of the estimated regression coefficient, as presented in Table 7. The credible interval (CI) for the acceleration factor is computed by exponentiating the lower and upper bounds of the coefficient's credible interval.

From Table 7, occupation was the only variable that demonstrated a reliable effect on survival time. The results showed an acceleration factor (AF) of approximately 0.009, with a credible interval that did not include 1, indicating a significant reduction in survival time relative to the baseline category. This suggests that individuals engaged in occupations requiring strenuous or demanding work conditions may experience earlier mortality compared to other occupational groups. Marital status and diabetes mellitus tended toward shorter survival times, as indicated by acceleration factors less than 1. However, their credible intervals included 1, implying substantial uncertainty regarding their true effects on survival.

For gender, the credible interval was extremely wide despite the acceleration factor being greater than 1, making it difficult to draw any reliable conclusion regarding the effect of gender on survival time [15]. Age showed only a very small effect on survival, with an acceleration factor of approximately 1.11 and a credible interval overlapping 1, indicating essentially no meaningful effect. Overall, the Bayesian lognormal accelerated failure time (AFT) model indicated that occupation had a credible negative effect on survival time, thereby accelerating failure. The remaining covariates, namely marital status, gender, age, and diabetes mellitus, showed possible associations with survival time; however, the evidence supporting these effects remained weak because their credible intervals were wide or overlapped 1. The model therefore suggests that occupation is the primary determinant influencing the survival of hypertensive patients.

4.7. Discussion

The aim of this study was to assess selected parametric models for the survival times of hypertensive patients using data obtained from the General Hospital, Okemesi, through a Bayesian approach. An estimated 1.13 billion individuals worldwide are believed to be living with hypertension, with the majority (about two-thirds) residing in low- and middle-income countries [16]. This highlights the substantial global burden associated with hypertension. The findings from this study suggest that the survival of hypertensive

patients may be influenced by occupational category, with patients engaged in formal employment having a higher likelihood of survival compared with those involved in informal occupations or those who were dependents [17].

Bayesian methods provide several advantages, including the incorporation of prior information, greater flexibility when handling smaller datasets, and the ability to generate full posterior distributions of model parameters. However, Bayesian methods are not inherently superior to frequentist approaches, as both frameworks possess distinct strengths and limitations. In situations involving sparse events, Bayesian inference may still be constrained unless informative priors are carefully chosen. The present study therefore demonstrates the complementary role of Bayesian survival models alongside classical survival analysis approaches. A 95% confidence interval was included in the diagnostic plots of the Bayesian lognormal accelerated failure time (AFT) model, and the posterior density plots for the parameters appeared approximately normally distributed.

4.8. Limitations

To assess the robustness of the findings under sparse event data, sensitivity analyses were conducted using alternative prior specifications. The results indicated that posterior estimates were sensitive to overly diffuse priors but became more stable under weakly informative priors. These analyses emphasize the importance of prior specification in small-sample Bayesian survival models and reinforce the exploratory nature of covariate inference in this dataset. It is important to acknowledge explicitly that the low event rate limits the strength and precision of the conclusions drawn from the analysis. Consequently, the findings should be interpreted with caution. In addition, the use of retrospective hospital records may introduce selection bias into the study.

Future studies involving larger event counts and prospective study designs are therefore recommended in order to provide stronger and more reliable inference.

5. Conclusion

This study evaluated Bayesian parametric accelerated failure time models, namely the Exponential, Weibull, and Lognormal models, applied to survival data obtained from hypertensive patients at Okemesi General Hospital, Ekiti State, Nigeria. Model comparison based on the Watanabe–Akaike Information Criterion (WAIC) and the Leave-One-Out Information Criterion (LOOIC) indicated that the Lognormal model provided the best fit to the data. Occupation emerged as a consistent determinant of survival outcomes, with patients engaged in formal employment exhibiting longer survival times, whereas dependent patients experienced substantially shorter survival times. The findings underscore the importance of socioeconomic factors in hypertension management and suggest that targeted interventions addressing occupational vulnerability may improve patient outcomes. From a methodological perspective, the study demonstrates the usefulness of Bayesian accelerated failure time models when the proportional hazards assumption is violated, since they provide direct estimates of the effects of covariates on survival time. Nevertheless, the analysis was constrained by the relatively small number of observed events (14 deaths among 155 patients), which limited parameter identifiability and reduced estimation precision. Although sensitivity analyses using alternative priors confirmed the robustness of the major findings, the resulting inference remains largely exploratory. Larger datasets with higher event frequencies are required to validate these findings and strengthen the conclusions. Future research should extend Bayesian survival modeling to broader patient populations, incorporate informative priors based on clinical knowledge, and explore hierarchical modeling structures capable of accounting for hospital-level variation. Such studies would contribute to both methodological advancement and the practical application of Bayesian survival analysis in resource-limited settings.

Data availability

Data will be made available on request.

References

[1] A. S. Ayalew, M. A. Erango & K. T. Gergiso, “Survival analysis of factor affects survival time of hypertension patients”, Open Journal of Modelling and Simulation **7** (2019) 177. https://doi.org/10.4236/ojmsi.2019.74010.

[2] D. G. Kleinbaum & M. Klein, *Survival Analysis: A Self-Learning Text*, 3rd Edition, Springer, New York, 2012. https://link.springer.com/book/10.1007/978-1-4419-6646-9.

[3] C. Jackson, “Flexsurv: A platform for parametric survival modeling in R”, Journal of Statistical Software **70** (8) (2016) 1. https://www.jstatsoft.org/article/view/v070i08.

[4] T. S. Hosseini & S. M. Taghi Ayatollahi, “Comparison of Cox regression and parametric models: application for assessment of survival of pediatric cases of acute leukemia in Southern Iran”, Asian Pacific Journal of Cancer Prevention **18** (4) (2017) 981. https://doi.org/10.22034/APJCP.2017.18.4.981.

[5] S. P. Setu, R. Kabir, M. A. Islam, S. Alauddin & M. T. Nahar, “Factors associated with time to first birth interval among ever married Bangladeshi women: A comparative analysis on Cox-PH model and parametric models”, PLOS Global Public Health **4** (12) (2024) e0004062. https://doi.org/10.1371/journal.pgph.0004062.

[6] S. Norouzi, E. Hajizadeh, M. A. Jafarabadi & S. Mazloomzadeh, “Analysis of the survival time of patients with heart failure with reduced ejection fraction: a Bayesian approach via a competing risk parametric model”, BMC Cardiovascular Disorders **24** (2024) 45. https://doi.org/10.1186/s12872-023-03685-y.

[7] I. Paolucci, Y. M. Lin, J. Albuquerque Marques Silva, "Bayesian parametric models for survival prediction in medical applications", BMC Medical Research Methodology **23** (2023) 250. https://doi.org/10.1186/s12874-023-02059-4.
[8] F. Faghiri & A. Kohansal, "Cox proportional hazards model with Bayesian neural network for survival prediction", Scientific Reports **15** (2025) 31581. https://doi.org/10.1038/s41598-025-16993-4.
[9] R Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, Vienna, Austria, 2026. https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf.
[10] E. Limpert, W. A. Stahel & M. Abbt, "Log-normal distributions across the sciences: keys and clues", Bioscience **51** (5) (2001) 341. https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2.
[11] S. I. Wadhah, S. Ibrahim & A. S. Mezher, "Bayesian method estimation for exponential and Weibull survival regression models", Iraqi Statisticians Journal **1** (2024). https://doi.org/10.62933/rrv02w24.
[12] M. D. Hoffman & A. Gelman, "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo", Journal of Machine Learning Research **15** (2014) 1593. https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf.
[13] O. S. Adesina, "Bayesian multilevel models for count data", Journal of the Nigerian Society of Physical Sciences **3** (2021) 224. https://doi.org/10.46481/jnsps.2021.168.
[14] J. G. Ibrahim, M. H. Chen & D. Sinha, *Bayesian Survival Analysis*, Springer-Verlag, New York, 2001. https://link.springer.com/book/10.1007/978-1-4757-3447-8.
[15] D. W. Hosmer, S. Lemeshow & S. May, *Applied survival analysis: regression modeling of time-to-event data*, 2nd Edition, John Wiley & Sons, Hoboken, NJ, 2008. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470258019.
[16] World Health Organization, *Hypertension*, WHO Fact Sheet, World Health Organization, Geneva, Switzerland, 2025. Available online: https://www.who.int/news-room/fact-sheets/detail/hypertension.
[17] NCD Risk Factor Collaboration (NCD-RisC), "Worldwide trends in hypertension prevalence, detection, treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants", The Lancet **398** (2022) 957. https://doi.org/10.1016/S0140-6736(21)01330-1.