

Article

Ernest Zamanah* and Suleman Nasiru

Bayesian Analysis of Tuberculosis Cases in Bolgatanga Municipality, Ghana–West Africa

<https://doi.org/10.1515/spp-2025-0023>

Received April 6, 2025; accepted August 20, 2025; published online September 19, 2025

Abstract: Tuberculosis (TB) represents a major threat to global health security and is currently the world most deadly infectious disease. TB is noted to affect people of all age groups and exposure to the disease has been shown to exhibit variation in the risk of infection across different factors. This study was therefore carried out to identify possible disease associations of some risk factors in relation to time until onset of TB among TB infected individuals in the Bolgatanga Municipality of the Upper East Region of Ghana. To this end, secondary data on time until onset of TB was obtained from the Bolgatanga Regional Hospital of the Upper East Region of Ghana and used for statistical analysis. In order to obtain appropriate underlying disease associations, the Bayesian approach to the modelling of survival time was employed in the statistical analyses. Thus, different parametric models including; skew normal, lognormal, shifted lognormal, exponential, Weibull and Fréchet models were fitted to the TB data within the Bayesian framework. The Weibull was the overall best fitted model. Based on the estimation results of the Bayesian Weibull regression model; individuals with no education, individuals with diabetes and males were identified as high-risk groups with shorter time frame for the onset of TB in the Bolgatanga Municipality. The study was concluded by indicating the need for Ghana Health Service to institute targeted interventions aimed at preventing and controlling the TB with special emphasis to the high-risk groups in the Bolgatanga Municipality.

Keywords: Bayesian framework; posterior distribution; widely applicable information criterion; Weibull model

***Corresponding author: Ernest Zamanah**, Department of Biometry, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana,
E-mail: ezamanah@yahoo.com. <https://orcid.org/0000-0002-2481-4127>

Suleman Nasiru, Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana,
E-mail: sulemanstat@gmail.com, snasiru@cktutas.edu.gh

1 Introduction

Tuberculosis (TB) is a communicable disease that mostly affects the lungs with mycobacterium tuberculosis, a type of bacteria being the causative agent. As an airborne disease, the bacteria are transmitted when infected people with the disease release bacteria into the air through coughing, sneezing or spitting (Kiani et al. 2021).

TB represents a major public health problem worldwide and poses a significant threat to health security. The incidence of tuberculosis has been shown to exhibit an increasing trend since 2021 with recorded cases exceeding 10 million every year. In 2023, the total number of TB recorded cases was about 10.8 million around the globe with 6.0 million cases attributed to men, 3.6 million cases being women and about 1.3 million cases attributed to children (Deku et al. 2024; World Health Organization [WHO] 2024). Globally, 1.25 million mortalities resulted from TB infections with 1.09 million deaths attributed to HIV-negative people and 161,000 deaths associated with HIV-positive people (WHO 2024). Although some level of progress has been made in the fight against TB, TB has replaced COVID-19 as the current global leading cause of death from a single infectious agent after COVID-19 held the position from 2019 to 2021 (Berida and Lindsley 2024; WHO 2024).

Currently, TB is found everywhere in the world. In 2023, the South-East Asia Region accounted for 45 % of all TB cases in the world. The second largest incidence of TB cases occurred in the African Region accounting for 24 % of all cases whilst Western Pacific Region accounted for 17 % (WHO 2024).

TB has been an old age major health related problem in Africa and Ghana in particular. In Africa, TB is noted as one of the top ten causes of death, outranking HIV/AIDS (WHO 2020). In 2021, the incidence of TB in Ghana was shown to be 136 per 100,000 with a mortality rate of 36 per 100,000 people (WHO 2023). On annual bases, about 44,000 cases of TB are recorded in Ghana with an estimated mortality of 10,000 cases. Noted as the seventh leading cause of mortality in Ghana, the distribution of TB cases has been shown to exhibit regional variations in Ghana with Upper East Region established as one of the regions with an increased likelihood of TB infection (Iddrisu et al. 2024).

Globally, TB affects people of all age groups and exposure to the disease has been shown to exhibit variation in the likelihood of infection across different risk factors. People with disease conditions such as HIV/AIDS or diabetes that compromise the immune system are noted to be at risk of getting infected with TB (Umeta et al. 2022). Some other covariates include alcohol use, cigarette smoking and educational level among others (Umeta et al. 2022).

Considerable progress has been made in relation to the treatment and elimination of TB in Ghana (Deku et al. 2024). However, there is paucity of data in relation to assessing the age of onset of TB across different risk factors in the Bolgatanga Municipality of the Upper East Region of Ghana. Bayesian estimation procedure is a useful method in modelling the time until onset of infectious diseases (Rastegar et al. 2024). Thus, the primary objective of this study was to evaluate possible risk factors associated with the time frame until onset of TB and determine possible disease associations in the Bolgatanga Municipality of the Upper East Region of Ghana by applying the Bayesian approach.

The structure for the rest of the article is as follows: In Section 2, the materials and methods are presented whilst Section 3 outlines the results and discussions. Section 4 finally presents the conclusions and recommendations of the study.

2 Materials and Methods

2.1 Study Design

Secondary data on tuberculosis was obtained for this study through a retrospective cohort study design.

2.2 Study Area

The Bolgatanga Municipality of the Upper East Region of Ghana represented the study area of this work. As the administrative capital of the Upper East Region of Ghana, the Bolgatanga Municipality is home to the Bolgatanga Regional Hospital. The hospital serves as a referral center for all private health facilities within the Municipality and also provides healthcare services to all patients within the Municipality and beyond.

2.3 Study Population

The participants representing the study population included recorded history of tuberculosis patients from 2015 to 2024 at the Bolgatanga Regional Hospital.

2.4 Inclusion and Exclusion Criteria

The medical records of all tuberculosis cases from 2015 to 2024 were included in the study. Medical records of tuberculosis cases were however excluded from the study where information on patients age were unavailable.

2.5 Data Collection

Secondary data on tuberculosis were obtained from the Bolgatanga Regional Hospital of the Upper East Region of Ghana by reviewing the medical histories of all tuberculosis cases recorded from 2015 to 2024. The Records Department of the Bolgatanga Regional Hospital is staffed with qualified and well trained personnel. Thus, the review of the medical records was done by the staff of the Records Department of the hospital. In all, the data were obtained from the records of 343 patients infected with TB in the Bolgatanga Municipality. The time frame representing the age of the patients from birth until onset of tuberculosis was the variable of interest in the study. Other variables such as diabetes status of patients at the time of TB diagnosis, sex of patient, marital status, employment status, religion and educational level were used as independent variables in the study.

2.6 Ethical Approval

The data were provided and approved for its use by the Records Department of the Bolgatanga Regional Hospital upon written request for the data. Personal information and identification numbers of patients were excluded from the data given out for this study. This, ensured confidentiality in the use of the data that were obtained for the study.

2.7 Study Variables

2.7.1 Dependent Variable

Time until the onset of TB represented the response variable for the study. Thus, the age of patient in years from birth until onset of TB was the variable of interest.

Starting time: Starting time is the age at birth of the TB patient.

End time: is the age at which the event occurred. That is, the age of TB diagnosis.

2.7.2 Independent Variables

Sex: As a prognostic factor, sex refers to male and female categories of the patients.

Marital Status: Three categories (married, single and divorced/widowed) were defined for marital status of the patients as an independent variable.

Employment Status: Two categories (Employed and Unemployed) were defined for the employment status of the patients.

Religion: Three categories (Muslim, Christian and Traditionalist) were defined for the religion of patients.

Educational Level: Four categories (No education, Primary, Secondary and Tertiary) were defined for educational level of the patients.

Diabetes Status: For the diabetes status of the patients, two categories (Yes and No) were defined.

2.8 Statistical Analysis

2.8.1 Bayesian Inference

Consider a random sample of independently and identically distributed survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, with a probability function denoted by $P(t_j|\eta)$ and survival function denoted by $S(t_j)$. The likelihood function of the sample according to Umeta et al. (2022) given that there are censored observations in the data is defined by

$$L = \prod_{j=1}^n P(t_j|\eta)^{\delta_j} S(t_j)^{(1-\delta_j)},$$

where

$$\delta_j = \begin{cases} 0, & \text{censored} \\ 1, & \text{observed} \end{cases}$$

is an indicator variable.

The Bayesian approach utilizes prior distributions of the unknown parameters η of the likelihood function in the modelling process. Before the sample t_j is observed, the prior distribution gives a reflection of existing knowledge about the parameter space η . Considered as stochastic, appropriate distributions known as prior distributions are assigned to all the parameters η within the Bayesian framework (Blangiardo and Cameletti 2015).

In a posterior inference, a probability distribution known as the posterior distribution is produced by combining the likelihood function of the sample and prior distributions (Blangiardo and Cameletti 2015). Given the observed data, the posterior distribution, denoted by $P(\eta|t_j)$ is the conditional distribution of the parameters. Defined as the product of the likelihood function and the prior distributions $P(\eta)$, the posterior distribution is given by

$$P(\eta|t_j) = \frac{P(t_j|\eta)P(\eta)}{\int_{\eta} L(t_j|\eta)P(\eta)d\eta}. \quad (1)$$

Alternatively, the posterior distribution can be expressed as

$$P(\eta|t_j) \propto P(t_j|\eta)P(\eta), \quad (2)$$

with $\int_{\eta} L(t_j|\eta)P(\eta)d\eta$ being a normalizing constant. By assigning a distribution for each parameter under the assumption of the parameter estimates being stochastic could produce informative priors resulting in misleading statistical inferences (Hahn 2014). Conjugate or non-informative priors allow for more efficiency in the parameter estimates of the posterior distributions within the Bayesian framework (Robert and Rousseau 2023). Thus, non-informative priors were utilized in this study.

To assess the adequacy of the Bayesian model, trace and density plots of the parameter estimates of posterior distributions were utilized through Markov Chain Monte Carlo (MCMC) procedures to determine model convergence in this study.

2.8.2 Bayesian Analysis Using Weibull Model

Consider that the distribution of the random sample of survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the Weibull distribution with parameters γ and η . The PDF and cumulative distribution function (CDF) of the Weibull distribution due to Weibull (1951) are given respectively by

$$f(t_j) = \frac{\gamma}{\eta} \left(\frac{t}{\eta} \right)^{\gamma-1} e^{-\left(\frac{t}{\eta}\right)^{\gamma}}, t_j \geq 0, \quad (3)$$

and

$$F(t_j) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\gamma}}, t_j \geq 0. \quad (4)$$

The survival function of the Weibull distribution is given by

$$S(t_j) = e^{-\left(\frac{t_j}{\eta}\right)^\gamma}, t_j \geq 0. \quad (5)$$

In building the Weibull regression, the log link function is used to introduce covariates to the regression structure such that

$$E(t_j) = \exp(W\varphi),$$

where W' is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$\eta\Gamma(1 + 1/\gamma) = \exp(W\varphi).$$

Thus,

$$\eta = \frac{\exp(W\varphi)}{\Gamma(1 + 1/\gamma)}.$$

This therefore gives the likelihood function as

$$L = \prod_{j=1}^n \left[\frac{\gamma}{\exp(w_j\varphi)/\Gamma(1 + 1/\gamma)} \times \left(\frac{t_j}{\exp(w_j\varphi)/\Gamma(1 + 1/\gamma)} \right)^{\gamma-1} \times e^{-\left(\frac{t_j}{\exp(w_j\varphi)/\Gamma(1 + 1/\gamma)} \right)^\gamma} \right]^{\delta_j} \\ \times \left[e^{-\left(\frac{t_j}{\exp(w_j\varphi)/\Gamma(1 + 1/\gamma)} \right)^\gamma} \right]^{(1-\delta_j)}.$$

(6)

In building the posterior distribution, the prior distributions for the regression coefficients and shape parameter of the model are specified. In the current study, a Gaussian prior was chosen for the regression coefficients such that $\varphi_i \sim N(0, \sigma_0^2)$ with a half Cauchy prior chosen for the shape parameter such that $\gamma \sim HC(0, k_0)$.

By applying the Bayes theorem, the joint posterior distribution is given by

$$P(\gamma, \varphi | W, t_j) \propto L(t_j | \gamma, \varphi, W) P(\gamma) P(\varphi). \quad (7)$$

That is,

$$\begin{aligned}
 P(\gamma, \varphi | W, t_j) &\propto \prod_{i=1}^n \left[\frac{\gamma}{\exp(w_j \varphi) / \Gamma(1+1/\gamma)} \times \left(\frac{t_j}{\exp(w_j \varphi) / \Gamma(1+1/\gamma)} \right)^{\gamma-1} \times e^{-\left(\frac{t_j}{\exp(w_j \varphi) / \Gamma(1+1/\gamma)} \right)^{\gamma}} \right]^{\delta_j} \\
 &\times \left[e^{-\left(\frac{t_j}{\exp(w_j \varphi) / \Gamma(1+1/\gamma)} \right)^{\gamma}} \right]^{(1-\delta_j)} \times \frac{2 \times k_0}{P(\gamma^2 + k_0^2)} \times \prod_{i=1}^I \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left(-\frac{1}{2 \times \sigma_o^2} \varphi_i^2\right).
 \end{aligned}
 \tag{8}$$

2.8.3 Bayesian Analysis Using Exponential Model

Consider that the distribution of the random sample of survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the exponential distribution with parameter η . The PDF of the exponential distribution as cited by Rather and Rather (2017) and Umeta et al. (2022) is given by

$$f(t_j) = \eta \exp(-\eta t_j), t_j \geq 0. \tag{9}$$

The CDF and survival function of the exponential distribution are given respectively by

$$F(t_j) = 1 - e^{-(\eta t_j)}, t_j \geq 0. \tag{10}$$

and

$$S(t_j) = e^{-(\eta t_j)}, t_j \geq 0. \tag{11}$$

Using the log link function, the exponential regression is defined such that

$$E(t_i) = \exp(W\varphi),$$

where W is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$1/\eta = \exp(W\varphi).$$

Thus,

$$\eta = \frac{1}{\exp(W\varphi)}$$

$$= \exp(-W\varphi).$$

This therefore gives the likelihood function as

$$L = \prod_{j=1}^n \left\{ e^{-[w_j \varphi + t_j \exp(-w_j \varphi)]} \right\}^{\delta_j} \times \left[e^{-t_j \exp(-w_j \varphi)} \right]^{(1-\delta_j)} \quad (12)$$

By choosing a Gaussian prior for the regression coefficients such that $\varphi_i \sim N(0, \sigma_0^2)$, the joint posterior distribution is obtained by applying the Bayes theorem. This gives

$$P(\varphi|W, t_j) \propto L(t_j|\varphi, W)P(\varphi). \quad (13)$$

That is,

$$P(\varphi|W, t_j) \propto \prod_{j=1}^n \left\{ e^{-[w_j \varphi + t_j \exp(-w_j \varphi)]} \right\}^{\delta_j} \times \left[e^{-t_j \exp(-w_j \varphi)} \right]^{(1-\delta_j)} \\ \times \prod_{i=1}^I \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{1}{2 \times \sigma_0^2} \varphi_i^2\right). \quad (14)$$

2.8.4 Bayesian Analysis Using Lognormal Model

Consider that the distribution of the random sample of independently and identically distributed survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the lognormal with parameters μ and σ . The PDF and CDF of the lognormal distribution as cited in Umeta et al. (2022) are given respectively by

$$f(t_j) = (2\pi)^{-1/2} (t_j \sigma)^{-1} \exp\left\{ \frac{-1}{2\sigma^2} [\log(t_j) - \mu]^2 \right\}, \mu \in \mathbb{R}, \sigma \geq 0, \quad (15)$$

and

$$F(t_j) = \Phi\left[\frac{\log(t_j) - \mu}{\sigma} \right]. \quad (16)$$

The survival function of the lognormal distribution is given by

$$S(t_j) = 1 - \Phi\left[\frac{\log(t_j) - \mu}{\sigma} \right]. \quad (17)$$

Using the log link function, the lognormal regression is defined such that

$$E(t_i) = \exp(W\varphi),$$

where W' is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$e^{(\mu + \sigma^2/2)} = \exp(W\varphi).$$

Thus,

$$\mu = \mathbf{W}\boldsymbol{\varphi} - \frac{\sigma^2}{2}$$

This therefore gives the likelihood function as

$$L = \prod_{j=1}^n \left[(2\pi)^{-1/2} (t_j \sigma)^{-1} \exp \left\{ \frac{-1}{2\sigma^2} [\log(t_j) + \sigma^2/2 - \mathbf{W}\boldsymbol{\varphi}]^2 \right\} \right]^{\delta_j} \times \left[1 - \Phi \left[\frac{\log(t_j) + \sigma^2/2 - \mathbf{W}\boldsymbol{\varphi}}{\sigma} \right] \right]^{(1-\delta_j)} \quad (18)$$

By choosing a Gaussian prior for the regression coefficients such that $\varphi_i \sim \mathcal{N}(0, \sigma_0^2)$ and a half Cauchy prior for the scale parameter such that $\sigma \sim HC(0, k_0)$, the joint posterior distribution is obtained by applying the Bayes theorem. This gives

$$P(\sigma, \boldsymbol{\varphi} | W, t_j) \propto L(t_j | \sigma, \boldsymbol{\varphi}, W) P(\sigma) P(\boldsymbol{\varphi}). \quad (19)$$

That is,

$$P(\sigma, \boldsymbol{\varphi} | W, t_j) \propto \prod_{j=1}^n \left[(2\pi)^{-1/2} (t_j \sigma)^{-1} \exp \left\{ \frac{-1}{2\sigma^2} [\log(t_j) + \sigma^2/2 - W\boldsymbol{\varphi}]^2 \right\} \right]^{\delta_j} \times \left[1 - \Phi \left[\frac{\log(t_j) + \sigma^2/2 - W\boldsymbol{\varphi}}{\sigma} \right] \right]^{(1-\delta_j)} \times \frac{2 \times k_0}{P(\sigma^2 + k_0^2)} \times \prod_{i=1}^I \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left(-\frac{1}{2 \times \sigma_0^2} \varphi_i^2 \right). \quad (20)$$

2.8.5 Bayesian Analysis Using Skew Normal Model

Consider that the distribution of the random sample of independently and identically distributed survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the skew normal with parameters μ , σ and γ . The PDF of the skew normal distribution due to Azzalini (1985) as cited in Wei et al. (2025) is given by

$$f(t_j) = \frac{2}{\sigma} \phi \left(\frac{t_j - \mu}{\sigma} \right) \Phi \left[\frac{\gamma(t_j - \mu)}{\sigma} \right], \mu, \gamma \in \mathbb{R}, \sigma > 0, \quad (21)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ represent the CDF and PDF of the standard normal distribution respectively. The CDF is and

$$F(t_j) = \Phi \left(\frac{t_j - \mu}{\sigma} \right) - 2T \left(\frac{t_j - \mu}{\sigma}, \gamma \right), \quad (22)$$

where $T(u, \gamma)$ is Owen's T function.

The survival function of the skew normal distribution is given by

$$S(t_j) = 1 - \Phi\left(\frac{t_j - \mu}{\sigma}\right) - 2T\left(\frac{t_j - \mu}{\sigma}, \gamma\right). \quad (23)$$

Using the log link function, the skew normal regression is defined such that

$$E(t_i) = \exp(W\varphi),$$

where W is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$\mu + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}} = \exp(W\varphi).$$

Thus,

$$\mu = \exp(W\varphi) - \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}.$$

This therefore gives the likelihood function as

$$\begin{aligned} L = \prod_{j=1}^n & \left[\frac{2}{\sigma} \phi\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}\right) \Phi\left[\frac{\gamma\left(t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}\right)}{\sigma}\right] \right]^{\delta_j} \\ & \times \left[1 - \Phi\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}\right) - 2T\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}, \gamma\right) \right]^{(1-\delta_j)} \end{aligned} \quad (24)$$

By choosing a Gaussian prior for the regression coefficients such that $\varphi_i \sim N(0, \sigma_0^2)$, a half Cauchy prior for the parameters (σ, γ) such that $\sigma \sim HC(0, k_0)$ and $\gamma \sim HC(0, k_1)$, the joint posterior distribution is obtained by applying the Bayes theorem. This gives

$$P(\sigma, \gamma, \varphi | W, t_j) \propto L(t_j | \sigma, \gamma, \varphi, W) P(\sigma) P(\gamma) P(\varphi). \quad (25)$$

That is,

$$\begin{aligned} P(\sigma, \varphi | W, t_j) & \propto \prod_{j=1}^n \left[\frac{2}{\sigma} \phi\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}\right) \Phi\left[\frac{\gamma\left(t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}\right)}{\sigma}\right] \right]^{\delta_j} \\ & \times \left[1 - \Phi\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}\right) - 2T\left(\frac{t_j - \exp(W\varphi) + \frac{\gamma\sigma}{\sqrt{1+\gamma^2}} \sqrt{\frac{2}{\pi}}}{\sigma}, \gamma\right) \right]^{(1-\delta_j)} \\ & \times \frac{2 \times k_0}{P(\sigma^2 + k_0^2)} \times \frac{2 \times k_1}{P(\gamma^2 + k_1^2)} \times \prod_{i=1}^I \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{1}{2 \times \sigma_0^2} \varphi_i^2\right). \end{aligned} \quad (26)$$

2.8.6 Bayesian Analysis Using Shifted Lognormal Model

Consider that the distribution of the random sample of independently and identically distributed survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the shifted lognormal with parameters μ , σ and λ . The PDF of the shifted lognormal distribution due to Aitchison and Brown (1969) as cited in Dabet et al. (2018) is given by

$$f(t_j) = \frac{1}{\sigma(t_j - \lambda)\sqrt{2\pi}} e^{[\ln(t_j - \lambda) - \mu]^2 / 2\sigma^2}, \mu, \sigma > 0, t_j \geq \lambda. \quad (27)$$

The CDF of the shifted lognormal distribution is given by

$$F(t_j) = \frac{1}{2} \left[1 - \operatorname{sgn}\{[-\ln(t_j - \lambda) + \mu]/\sigma\} \operatorname{erf}\left(\left|[-\ln(t_j - \lambda) + \mu]/\sigma\right|/\sqrt{2}\right) \right]. \quad (28)$$

The survival function of the shifted lognormal distribution is given by

$$S(t_j) = 1 - \frac{1}{2} \left[1 - \operatorname{sgn}\{[-\ln(t_j - \lambda) + \mu]/\sigma\} \operatorname{erf}\left(\left|[-\ln(t_j - \lambda) + \mu]/\sigma\right|/\sqrt{2}\right) \right]. \quad (29)$$

Using the log link function, the shifted lognormal regression is defined such that

$$E(t_j) = \exp(W\varphi),$$

where W' is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$\exp(\mu) = \exp(W\varphi) / \left[\left\{ 1 + \sqrt{[\exp(\sigma^2) - 1]} \right\} \sqrt{\exp(\sigma^2)} \right].$$

Thus,

$$\mu = W\varphi - \sigma^2.$$

This therefore gives the likelihood function as

$$L = \prod_{j=1}^n \left[\frac{1}{\sigma(t_j - \lambda)\sqrt{2\pi}} e^{[\ln(t_j - \lambda) - W\varphi + \sigma^2]^2 / 2\sigma^2} \right]^{\delta_j} \times \left[1 - \frac{1}{2} \left[1 - \operatorname{sgn}\{[-\ln(t_j - \lambda) + W\varphi - \sigma^2]/\sigma\} \operatorname{erf}\left(\left|[-\ln(t_j - \lambda) + W\varphi - \sigma^2]/\sigma\right|/\sqrt{2}\right) \right] \right]^{(1-\delta_j)} \quad (30)$$

By choosing a Gaussian prior for the regression coefficients such that $\varphi_i \sim N(0, \sigma_0^2)$, a half Cauchy prior for the parameters (σ, λ) such that $\sigma \sim HC(0, k_0)$ and $\lambda \sim HC(0, k_1)$, the joint posterior distribution is obtained by applying the Bayes theorem. This gives

$$P(\sigma, \lambda, \varphi | W, t_j) \propto L(t_j | \sigma, \lambda, \varphi, W) P(\sigma) P(\lambda) P(\varphi). \quad (31)$$

That is,

$$\begin{aligned} P(\sigma, \varphi | W, t_j) &\propto \prod_{j=1}^n \left[\frac{1}{\sigma(t_j - \lambda)\sqrt{2\pi}} e^{[\ln(t_j - \lambda) - W\varphi + \sigma^2]^2 / 2\sigma^2} \right]^{\delta_j} \\ &\times \left[1 - \frac{1}{2} \left[1 - \operatorname{sgn}\{[-\ln(t_j - \lambda) + W\varphi - \sigma^2]/\sigma\} \operatorname{erf}\left(\left| \frac{[-\ln(t_j - \lambda) + W\varphi - \sigma^2]/\sigma}{\sqrt{2}} \right| \right) \right] \right]^{(1-\delta_j)} \\ &\times \frac{2 \times k_0}{P(\sigma^2 + k_0^2)} \times \frac{2 \times k_1}{P(\lambda^2 + k_1^2)} \times \prod_{i=1}^I \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left(-\frac{1}{2 \times \sigma_o^2} \varphi_i^2\right). \end{aligned} \quad (32)$$

2.8.7 Bayesian Analysis Using Fréchet Model

Consider that the distribution of the random sample of survival times $\mathbf{t} = (t_1, t_2, \dots, t_n)'$, is that of the Fréchet distribution with parameters $\alpha > 0$ and $\lambda > 0$. The PDF and CDF of the Fréchet distribution due to Fréchet (1927) as cited in Zamanah et al. (2022) are given respectively by

$$f(t_j) = \alpha \lambda (t_j)^{-\lambda-1} e^{-\alpha (t_j)^{-\lambda}}, t_j > 0. \quad (33)$$

and

$$F(t_j) = e^{-\alpha (t_j)^{-\lambda}}. \quad (34)$$

The survival function of the Fréchet distribution is given by

$$S(t_j) = 1 - e^{-\alpha (t_j)^{-\lambda}}. \quad (35)$$

Using the log link function, the Fréchet regression is defined such that

$$E(t_i) = \exp(W\varphi),$$

where W' is a $k \times 1$ vector of explanatory variables and φ is a $k \times 1$ vector of regression coefficients. That is,

$$E(t_j) = \exp(W\varphi).$$

Further,

$$\alpha \Gamma(1 - 1/\lambda) = \exp(W\varphi).$$

Thus,

$$\alpha = \frac{\exp(W\varphi)}{\Gamma(1-1/\lambda)}.$$

This therefore gives the likelihood function as

$$L = \prod_{j=1}^n \left[\frac{\lambda (t_j)^{-\lambda-1} e^{\{w_j \varphi - [(t_j)^{-\lambda} \exp(w_j \varphi)] / \Gamma(1-1/\lambda)\}}}{\Gamma(1-1/\lambda)} \right]^{\delta_j} \times \left[e^{-[(t_j)^{-\lambda} \exp(w_j \varphi)] / \Gamma(1-1/\lambda)} \right]^{(1-\delta_j)}. \quad (36)$$

By choosing a Gaussian prior for the regression coefficients such that $\varphi_i \sim N(0, \sigma_0^2)$ and a half Cauchy prior for the shape parameter such that $\lambda \sim HC(0, k_0)$, the joint posterior distribution is obtained by applying the Bayes theorem. This gives

$$P(\lambda, \varphi | W, t_j) \propto L(t_j | \lambda, \varphi, W) P(\lambda) P(\varphi). \quad (37)$$

That is,

$$\begin{aligned} P(\lambda, \varphi | W, t_j) &\propto \prod_{j=1}^n \left[\frac{\lambda (t_j)^{-\lambda-1} e^{\{w_j \varphi - [(t_j)^{-\lambda} \exp(w_j \varphi)] / \Gamma(1-1/\lambda)\}}}{\Gamma(1-1/\lambda)} \right]^{\delta_j} \\ &\times \left[e^{-[(t_j)^{-\lambda} \exp(w_j \varphi)] / \Gamma(1-1/\lambda)} \right]^{(1-\delta_j)} \\ &\times \frac{2 \times k_0}{P(\lambda^2 + k_0^2)} \times \prod_{i=1}^I \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{1}{2 \times \sigma_0^2} \varphi_i^2\right). \end{aligned} \quad (38)$$

2.9 Model Selection Criteria

To determine the model of best fit to the TB data, leave one out cross-validation (LOO-CV) criterion and the widely applicable information criterion (WAIC) were used in the study. The LOO-CV due to Vehtari et al. (2017) is a method that utilizes the log-likelihood evaluated at the posterior simulations of the parameter values to estimate pointwise out-of-sample prediction accuracy from a fitted Bayesian model. The LOO-CV is given as

$$\text{elpd}_{\text{loo}} = \sum_{j=1}^n \log p(w_j | w_{-j}), \quad (39)$$

where

$$p(w_j | w_{-j}) = \int p(w_j | \varphi) p(\varphi | w_{-j}) d\varphi$$

is the LOO predictive density for data point w_j given the remaining data points w_{-j} and ϕ represents the parameter space of the model.

As an information-based model selection criterion, the WAIC is given by

$$\text{WAIC} = -2 \sum_{j=1}^n \log p_{\text{post}}(w_j) + 2pD, \tag{40}$$

where $\sum_{j=1}^n \log p_{\text{post}}(w_j)$ is the sum of predictive density for each data point and pD is the effective number of parameters (Umata et al. 2022). The model with least values of LOO-CV and WAIC represents the best model (Farhin and Khan 2022).

3 Results and Discussions

3.1 Summary Statistics

Figure 1 presents the histogram and kernel density plots of the time until onset of TB. The results show that the ages of the patients until onset of TB are skewed to the right. That is, majority of the TB patients experienced the onset of TB at a younger age on average.

Table 1 shows the descriptive statistics illustrating the distribution of the patients across various categories of the predictor variables. Table 1 shows that among the different sexes, TB infections were high among males accounting for 72.3 % of all the infections. Also, majority of the patients representing 64.1 % were married followed by 30.3 % of the patients being single. Further, the section of the patients who were unemployed represented a majority of 70.0 % compared to the employed group. Across the various religions, majority of the TB infections representing 56.6 % were shown to be among Christians. Across various educational levels also, the level of TB infection was higher among those with no form of formal education accounting for 43.1 % of the infections. Finally, the descriptive statistics showed diabetes as a co-morbidity among 27.4 % of the patients infected with TB.

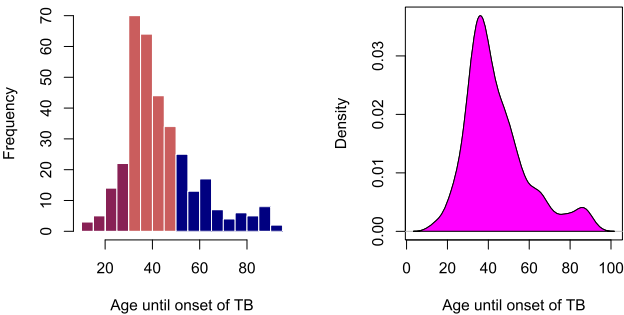


Figure 1: Histogram and kernel density plots of the tuberculosis data.

Table 1: Descriptive statistics of patients based on some predictor variables.

| Variable | Category | Frequency (%) |
|-------------------|------------------|---------------|
| Sex | Female | 95 (27.7 %) |
| | Male | 248 (72.3 %) |
| Marital status | Divorced/Widowed | 19 (5.5 %) |
| | Married | 220 (64.1 %) |
| | Single | 104 (30.3 %) |
| Employment status | Unemployed | 103 (30.0 %) |
| | Employed | 240 (70.0 %) |
| Religion | Christian | 194 (56.6 %) |
| | Muslim | 30 (8.7 %) |
| | Traditional | 119 (34.7 %) |
| Educational level | No education | 148 (43.1 %) |
| | Primary | 58 (16.9 %) |
| | Secondary | 108 (31.5 %) |
| | Tertiary | 29 (8.5 %) |
| Diabetes status | Yes | 94 (27.4 %) |
| | No | 249 (72.6 %) |

3.2 Model Selection Results

To evaluate the effects of the predictor variables on the time until onset of TB in the Bolgatanga Municipality, Bayesian parametric models were fitted to the TB data and their performances compared using the LOO-CV criterion and WAIC. Table 2 shows the accuracy measures of the fitted Bayesian regression models. The results showed the Weibull model as the best fitted model with least values of LOO-CV and WAIC.

Table 3 shows the parameter estimates of the posterior distribution for the Weibull regression model. Among the predictors, the results showed sex, religion, educational level and diabetes status to be statistically significant at 95 % confidence level.

Table 2: Model accuracy results on the tuberculosis data.

| Model | LOO estimates | | WAIC |
|-------------------|---------------|---------------|-------------------|
| | Elpd-diff | se-diff | |
| Weibull | 0.0000 | 0.0000 | 2,643.6000 |
| Lognormal | −10.4000 | 7.5000 | 2,664.4000 |
| Shifted lognormal | −11.7000 | 8.0000 | 2,666.9000 |
| Skew normal | −31.0000 | 9.0000 | 2,705.5000 |
| Fréchet | −56.3000 | 14.4000 | 2,754.5000 |
| Exponential | −317.1000 | 12.9000 | 3,277.8000 |

Bolded means best based on selection criteria.

Table 3: Parameter estimates of Weibull Bayesian regression on the tuberculosis data.

| Parameter | Estimates | Standard error | 95 % Credible interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|------------------------|------|----------|----------|
| Intercept | 3.86 | 0.07 | (3.74, 4.00) | 1.00 | 5,513 | 4,212 |
| Sex | -0.26 | 0.04 | (-0.33, -0.19) | 1.00 | 6,392 | 4,596 |
| Marital status | | | | | | |
| Female (Ref) | | | | | | |
| Married | 0.01 | 0.06 | (-0.11, 0.12) | 1.00 | 5,582 | 4,343 |
| Single | -0.09 | 0.06 | (-0.22, 0.02) | 1.00 | 5,361 | 3,992 |
| Divorced (Ref) | | | | | | |
| Employment status | | | | | | |
| Unemployed | -0.04 | 0.03 | (-0.10, 0.02) | 1.00 | 6,760 | 4,359 |
| Employed (Ref) | | | | | | |
| Religion | | | | | | |
| Muslim | 0.02 | 0.05 | (-0.01, 0.18) | 1.00 | 7,320 | 4,139 |
| Traditional | 0.17 | 0.03 | (0.10, 0.21) | 1.00 | 7,416 | 4,609 |
| Christian (Ref) | | | | | | |
| Educational level | | | | | | |
| Primary | 0.22 | 0.04 | (0.14, 0.29) | 1.00 | 6,925 | 4,878 |
| Secondary | 0.22 | 0.03 | (0.15, 0.28) | 1.00 | 6,880 | 5,101 |
| Tertiary | 0.28 | 0.05 | (0.18, 0.38) | 1.00 | 6,976 | 4,966 |
| None (Ref) | | | | | | |
| Diabetes status | | | | | | |
| Yes | -0.15 | 0.03 | (-0.21, -0.09) | 1.00 | 7,099 | 4,315 |
| No (Ref) | | | | | | |

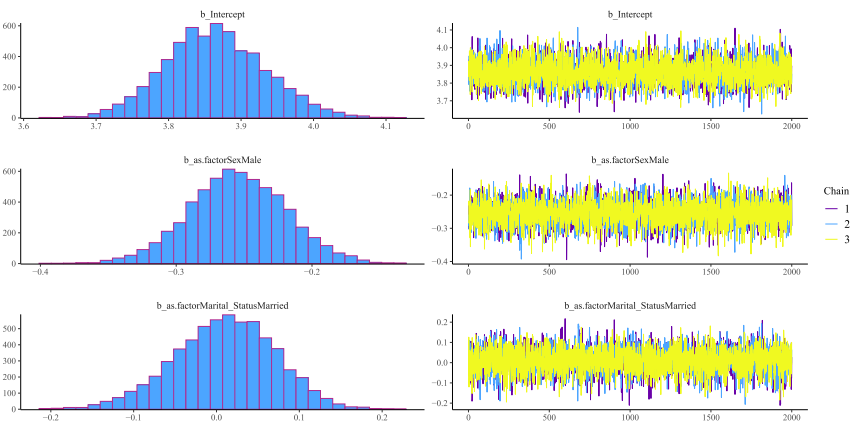


Figure 2: Density and Trace plot of Weibull Bayesian regression parameters.

3.3 Model Diagnostic Checks

This section investigates the adequacy of the fitted Weibull Bayesian regression model to determine whether there was satisfactory model convergence. Figures 2–5 show the trace and density plots of the parameters of the fitted Weibull Bayesian regression model. The results showed an overlay of the chains illustrating a good mixing of the three chains for all the parameters. The results confirm that the Monte Carlo Markov Chain simulation converged satisfactorily to the target posterior distribution in the estimation of the parameters. The results in

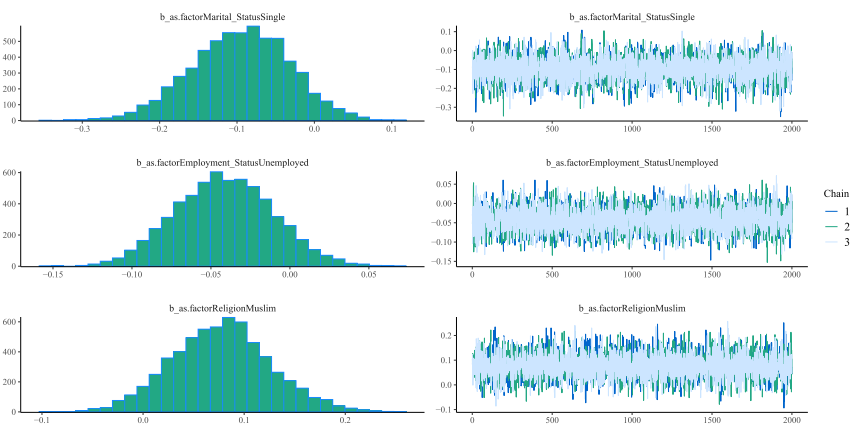


Figure 3: Density and Trace plot of Weibull Bayesian regression parameters.

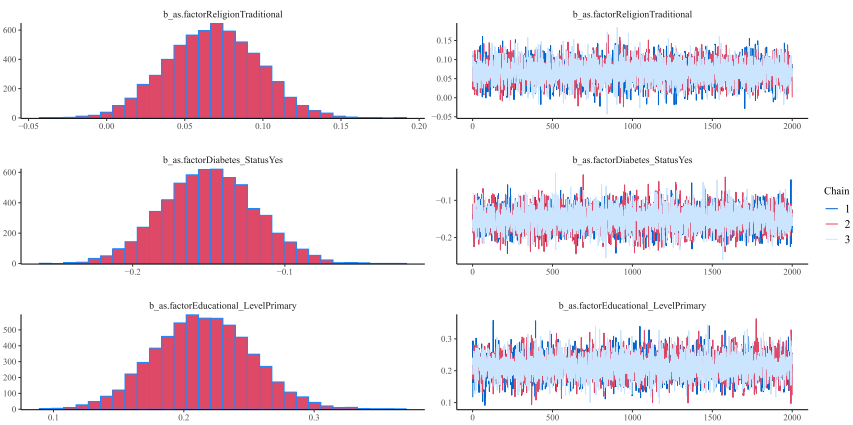


Figure 4: Density and Trace plots of some parameters for Weibull Bayesian regression parameters.

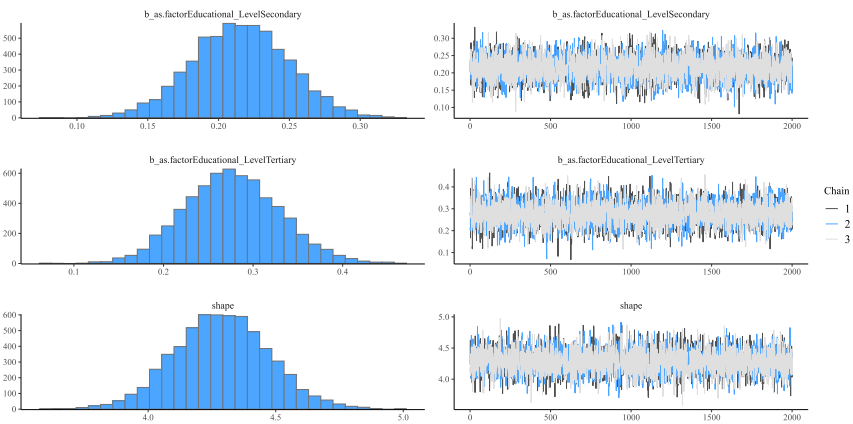


Figure 5: Density and Trace plot of Weibull Bayesian regression parameters.

Table 3 also support this finding with the Rhat values being 1 for all the parameters of the model.

Further, the posterior predictive plot in Figure 6 shows an overlap of the posterior and replicated distributions having the same trend. The results confirm that the posterior predictive capability of the Weibull regression model was adequate in modelling the time until onset of TB in the Bolgatanga Municipality.

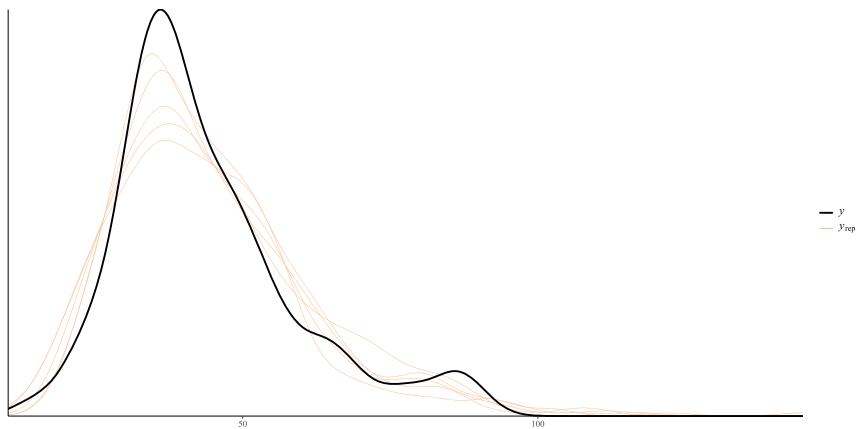


Figure 6: Posterior predictive plot of Weibull Bayesian regression.

3.4 Discussions

The time frame for onset of TB in the Bolgatanga Municipality was shown to be skewed to the right from the summary results of the study thereby making classical regression analysis inappropriate for application. Also, the summary results showed that 27.4 % of the patients infected with TB had diabetes as a comorbidity.

As a robust modelling approach, the Bayesian framework was utilized in modelling time until onset of TB as a response variable in this study. Varied Bayesian parametric models that were empirically evaluated included: skew normal, lognormal, shifted-lognormal, exponential, Weibull and Fréchet Bayesian regression models. The Weibull Bayesian regression model was the best performing model with least values of LOO-CV and WAIC.

The posterior estimation results of the Weibull Bayesian regression identified sex, religion, educational level and diabetes status as statistically significant prognostic factors at 95 % confidence level. However, marital status and employment status showed no statistical significance.

Thus, at the 95 % confidence level, the results illustrated a statistically significant difference in the time frame for the onset of TB between males and females. Specifically, it was empirically established that the time frame for onset of TB is shorter in males than females. This shows that men get TB at a younger age than women. Thus, males have a higher risk of getting infected with TB at a younger age than that of females in the Bolgatanga Municipality. This observation is consistent with Humayun et al. (2022) that males have a higher risk of getting infected with TB at a younger age than that of females.

Again, the findings showed a statistically significant difference in the time frame for the onset of TB between Christians and Traditionalists. The findings specifically confirmed that Traditionalists have a longer time frame for the onset of TB compared to Christians. Between Christians and Muslims however, the results showed no significant difference in the time frame for the onset of TB. The difference could be as a result of the mass gatherings associated with both Christian and Muslim worship compared to Traditional worship. This finding is consistent with the observations of other studies such as Hoang and Gautret (2018) and Yezli et al. (2023) that mass gathering, as practiced by Muslims and Christians, is associated with an increased risk of TB infection.

Across different educational levels, the findings indicated a statistically significant difference in the time frame for the onset of TB between individuals with no form of formal education and those with higher educational levels. Specifically, the results confirmed that individuals with higher educational levels (primary, secondary and tertiary) have a longer time frame for the onset of TB compared to those with no formal education. That is, individuals with lower levels of education get infected with TB at a younger age compared to individuals with higher levels of education. This finding is consistent with Medhi and Mahanta (2008) that individuals with lower levels of education get infected with TB at a younger age compared to individuals with higher levels of education.

The findings further established a statistically significant difference in the time frame for the onset of TB between individuals infected with diabetes and individuals without diabetes. Specifically, the findings confirmed that individuals infected with diabetes have a shorter time frame for the onset of TB compared to individuals without diabetes. The finding is consistent with Boadu et al. (2024) and Buasroung et al. (2022) that individuals with diabetes have a higher risk of TB infection.

4 Conclusions and Recommendations

4.1 Conclusions

The study was designed to model the time until onset of TB in the Bolgatanga Municipality of the Upper East Region of Ghana and identify possible disease associations of some risk factors. To this end, skew normal, lognormal, shifted lognormal, exponential, Weibull and Fréchet models were fitted to the TB data using the Bayesian procedure. The Weibull model was established as the best performing model in comparison with the other models. Based on the posterior estimation results of the Weibull model, sex, religion, educational level and diabetes status were

established as observed characteristics that significantly influence the time frame for the onset of TB in the Bolgatanga Municipality. Specifically, it was established that males have a shorter time frame for the onset of TB compared to females. It was also established that Traditionalists have a longer time frame for the onset of TB compared to Christians. Between Christians and Muslims, no significant difference was however detected in the time frame for the onset of TB in the Bolgatanga Municipality. Across different educational levels, it was established that individuals with higher educational levels have a longer time frame for the onset of TB compared to individuals with no education. Again, it was established that individuals with diabetes have a shorter time frame for the onset of TB compared to individuals without diabetes.

4.2 Recommendations

Based on the findings of this study, it is recommended for the Ghana Health Service and other healthcare agencies within the Municipality to institute advocacy programmes on; the need for lifestyle modifications, raising awareness about TB infections and the need for early screening and treatment. This will get the public well informed and help in the prevention of TB infections. Education on the need for lifestyle modifications can help prevent conditions such as diabetes and contribute in reducing the risk of TB infection in the Municipality. As a high risk group, it is also recommended for males to consider frequent screening for early detection and treatment of TB in the Bolgatanga Municipality.

4.3 Study Limitations

The use of secondary data limited the choice of variables in the study. Consequently, important variables such as HIV, hypertension and alcohol usage which could influence TB infection could not be captured in the current study. Nonetheless, the findings of this study are useful and have policy implications in the Bolgatanga Municipality.

Appendix

See Appendix Tables 1–5.

Table 1: Parameter estimates of Bayesian skew normal regression on the tuberculosis data.

| Parameter | Estimates | Standard error | 95 % Credible interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|------------------------|------|----------|----------|
| Intercept | 48.26 | 3.18 | (41.82, 54.36) | 1.00 | 5,926 | 4,413 |
| Sex | -4.86 | 2.08 | (-9.20, -1.10) | 1.00 | 5,104 | 4,127 |
| Marital status | -3.18 | 2.95 | (-8.86, 2.76) | 1.00 | 4,583 | 4,431 |
| | -9.24 | 23.11 | (-15.17, -3.03) | 1.00 | 4,775 | 4,254 |
| Employment status | -1.97 | 1.51 | (-4.92, 1.06) | 1.00 | 8,459 | 3,633 |
| Religion | 0.91 | 2.31 | (-3.73, 5.50) | 1.00 | 7,303 | 4,457 |
| | 5.21 | 1.48 | (2.31, 8.13) | 1.00 | 7,299 | 4,561 |
| Educational level | 8.00 | 1.98 | (4.09, 11.86) | 1.00 | 7,062 | 5,058 |
| | 6.57 | 1.64 | (3.46, 9.87) | 1.00 | 5,717 | 5,034 |
| | 11.29 | 2.53 | (6.35, 16.22) | 1.00 | 6,601 | 4,805 |
| Diabetes status | -5.76 | 1.54 | (-8.87, -2.78) | 1.00 | 8,111 | 4,629 |
| | No (Ref) | | | | | |

Table 2: Parameter estimates of Bayesian lognormal regression on the tuberculosis data.

| Parameter | Estimates | Standard error | 95 % Credible interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|------------------------|------|----------|----------|
| Intercept | 3.84 | 0.08 | (3.69, 4.00) | 1.00 | 5,883 | 4,308 |
| Sex | −0.16 | 0.04 | (−0.23, −0.09) | 1.00 | 6,930 | 4,579 |
| Marital status | −0.05 | 0.07 | (−0.19, 0.09) | 1.00 | 5,292 | 4,698 |
| | −0.21 | 0.07 | (−0.36, −0.07) | 1.00 | 5,170 | 4,516 |
| Employment status | −0.06 | 0.04 | (−0.13, 0.01) | 1.00 | 7,516 | 4,548 |
| Religion | 0.03 | 0.05 | (−0.08, 0.13) | 1.00 | 9,238 | 4,763 |
| | 0.11 | 0.03 | (0.04, 0.18) | 1.00 | 7,812 | 4,288 |
| Educational level | 0.20 | 0.04 | (0.12, 0.29) | 1.00 | 7,861 | 5,012 |
| | 0.18 | 0.04 | (0.10, 0.25) | 1.00 | 7,332 | 4,405 |
| | 0.26 | 0.06 | (0.14, 0.28) | 1.00 | 7,766 | 4,266 |
| Diabetes status | −0.15 | 0.04 | (−0.22, −0.07) | 1.00 | 7,914 | 3,966 |
| | No (Ref) | | | | | |

Table 3: Parameter estimates of Bayesian shifted-lognormal regression on the tuberculosis data.

| Parameter | Estimates | Standard Error | 95 % Credible Interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|------------------------|------|----------|----------|
| Intercept | 3.80 | 0.09 | (3.62, 3.97) | 1.00 | 2,832 | 2,517 |
| Sex | 0.17 | 0.04 | (-0.25, -0.09) | 1.00 | 4,457 | 3,423 |
| Marital status | | | | | | |
| Female (Ref) | | | | | | |
| Married | -0.05 | 0.07 | (-0.20, 0.09) | 1.00 | 3,314 | 3,185 |
| Single | -0.22 | 0.08 | (-0.38, -0.07) | 1.00 | 3,235 | 3,191 |
| Employment status | | | | | | |
| Divorced (Ref) | | | | | | |
| Unemployed | -0.07 | 0.04 | (-0.15, 0.00) | 1.00 | 4,303 | 3,780 |
| Employed (Ref) | | | | | | |
| Religion | | | | | | |
| Muslim | 0.03 | 0.06 | (-0.09, 0.14) | 1.00 | 5,434 | 3,524 |
| Traditional | 0.12 | 0.04 | (0.05, 0.19) | 1.00 | 4,423 | 3,822 |
| Educational level | | | | | | |
| Christian (Ref) | | | | | | |
| Primary | 0.21 | 0.05 | (0.12, 0.31) | 1.00 | 3,577 | 3,488 |
| Secondary | 0.18 | 0.04 | (0.11, 0.26) | 1.00 | 4,159 | 4,039 |
| Tertiary | 0.27 | 0.06 | (0.14, 0.39) | 1.00 | 4,359 | 3,051 |
| Diabetes status | | | | | | |
| None (Ref) | | | | | | |
| Yes | -0.15 | 0.04 | (-0.23, -0.08) | 1.00 | 4,166 | 3,820 |
| No (Ref) | | | | | | |

Table 4: Parameter estimates of Bayesian exponential regression on the tuberculosis data.

| Parameter | Estimates | Standard error | 95% Credible interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|-----------------------|------|----------|----------|
| Intercept | 3.92 | 0.28 | (3.39, 4.48) | 1.00 | 5,451 | 4,711 |
| Sex | -0.20 | 0.14 | (-0.47, 0.07) | 1.00 | 7,490 | 4,901 |
| Marital status | -0.05 | 0.25 | (-0.57, 0.43) | 1.00 | 4,842 | 4,504 |
| | -0.20 | 0.27 | (-0.74, 0.31) | 1.00 | 4,788 | 4,323 |
| Employment status | -0.06 | 0.13 | (-0.30, 0.19) | 1.00 | 6,833 | 4,775 |
| Religion | 0.06 | 0.20 | (-0.32, 0.47) | 1.00 | 7,489 | 4,084 |
| | 0.10 | 0.12 | (-0.14, 0.33) | 1.00 | 7,684 | 4,736 |
| Educational level | 0.21 | 0.16 | (-0.10, 0.52) | 1.00 | 6,639 | 3,989 |
| | 0.19 | 0.13 | (-0.07, 0.45) | 1.00 | 7,253 | 4,681 |
| | 0.27 | 0.21 | (-0.13, 0.69) | 1.00 | 7,585 | 4,936 |
| Diabetes status | -0.15 | 0.13 | (-0.39, 0.11) | 1.00 | 7,835 | 4,484 |
| | No (Ref) | | | | | |

Table 5: Parameter estimates of Bayesian Fréchet regression on the tuberculosis data.

| Parameter | Estimates | Standard error | 95 % Credible interval | Rhat | Bulk_ESS | Tail_ESS |
|-------------------|-----------|----------------|------------------------|------|----------|----------|
| Intercept | 3.90 | 0.08 | (3.74, 4.06) | 1.00 | 5,663 | 4,698 |
| Sex | -0.02 | 0.04 | (-0.09, 0.05) | 1.00 | 7,263 | 4,795 |
| | | | | | | |
| Marital status | -0.09 | 0.07 | (-0.23, 0.06) | 1.00 | 4,887 | 4,239 |
| | -0.29 | 0.08 | (-0.43, -0.13) | 1.00 | 5,006 | 4,626 |
| | | | | | | |
| Employment status | -0.11 | 0.04 | (-0.19, -0.04) | 1.00 | 7,734 | 4,437 |
| | | | | | | |
| Religion | -0.03 | 0.06 | (-0.15, 0.08) | 1.00 | 7,707 | 4,204 |
| | 0.15 | 0.04 | (0.08, 0.23) | 1.00 | 6,925 | 4,831 |
| | | | | | | |
| Educational level | 0.23 | 0.05 | (0.13, 0.32) | 1.00 | 6,868 | 4,636 |
| | 0.12 | 0.04 | (0.05, 0.20) | 1.00 | 6,933 | 4,606 |
| | 0.29 | 0.06 | (0.16, 0.41) | 1.00 | 6,654 | 5,012 |
| | | | | | | |
| Diabetes status | -0.11 | 0.04 | (-0.19, -0.04) | 1.00 | 7,041 | 4,219 |
| | | | | | | |
| | No (Ref) | | | | | |

References

- Aitchison, J., and J. A. C. Brown. 1969. *The Lognormal Distribution. With Special Reference to Its Uses in Economics*. Cambridge: Cambridge University Press.
- Azzalini, A. 1985. "A Class of Distributions Which Includes the Normal Ones." *Scandinavian Journal of Statistics* 12: 171–8.
- Berida, T., and C. W. Lindsley. 2024. "Move over COVID, Tuberculosis Is Once Again the Leading Cause of Death from a Single Infectious Disease." *Journal of Medicinal Chemistry* 67 (24): 21633–40.
- Blangiardo, M., and M. Cameletti. 2015. *Spatial and Spatio-temporal Bayesian Models with R-INLA*. New York: John Wiley & Sons Inc.
- Boadu, A. A., M. Yeboah-Manu, S. Osei-Wusu, and D. Yeboah-Manu. 2024. "Tuberculosis and Diabetes Mellitus: The Complexity of the Comorbid Interactions." *International Journal of Infectious Diseases* 146: 1–8.
- Buasroung, P., T. Petnak, P. Liwtanakitpipat, and S. Kiertiburanakul. 2022. "Prevalence of Diabetes Mellitus in Patients with Tuberculosis: A Prospective Cohort Study." *International Journal of Infectious Diseases* 116: 374–9.
- Dabet, A., H. Homma, H. Homma, and F. E. Gunawan. 2018. "Effectiveness of Pre-screening Test for Safe Design of Components Fabricated with Abaca Natural Fiber." *International Journal of Mechanical Engineering and Technology* 9 (13): 1207–15.
- Deku, J. G., E. Aninagyei, I. Bedzina, G. Nudo, E. Ativi, and P. Mensah, et al. 2024. "Trends of Mycobacterium tuberculosis and Rifampicin Resistance at the Ho Teaching Hospital in Ghana." *PLoS One* 19 (6): 1–13.
- Farhin, S., and A. A. Khan. 2022. "Bayesian Survival Analysis of Acute Encephalitis Syndrome with Censoring Mechanism Using Brms Package." *Journal of Statistics Applications & Probability* 11 (3): 963–80.
- Fréchet, M. 1927. "Sur la loi de Probabilité de l'écart Maximum." *Annales de la Société Polonaise de Mathématique* 6: 93–116.
- Hahn, U. 2014. "The Bayesian Boom: Good Thing or Bad?" *Frontiers in Psychology* 5 (765): 1–12.
- Hoang, V. T., and P. Gautret. 2018. "Infectious Diseases and Mass Gatherings." *Current Infectious Disease Reports* 20 (11): 1–12.
- Humayun, M., J. Chirenda, W. Ye, I. Mukeredzi, H. A. Mujuru, and Z. Yang. 2022. "Effect of Gender on Clinical Presentation of Tuberculosis (TB) and Age-specific Risk of TB, and TB-Human Immunodeficiency Virus Coinfection." *Open Forum Infectious Diseases* 9 (10): 1–9.
- Iddrisu, A. K., E. Amikiya, and F. K. Bukari. 2024. "Spatio-Temporal Characteristics of Tuberculosis in Ghana." *F1000Research* 11 (20): 1–30.
- Kiani, B., A. R. Rahmati, R. Bergquist, S. Hashtarkhani, N. Firouraghi, and N. Bagheri, et al. 2021. "Spatio-Temporal Epidemiology of the Tuberculosis Incidence Rate in Iran 2008 to 2018." *BMC Public Health* 21 (1093): 1–20.
- Medhi, G. K., and J. Mahanta. 2008. "Impact of Education and Income on Pulmonary Tuberculosis (PTB) Among Adult Tea Garden Workers of Assam." *International of Infectious Diseases* 12: e343.
- Rastegar, M., E. Nazar, M. Nasehi, S. Sharafi, V. Fakoor, and M. T. Shakeri. 2024. "Bayesian Estimation of the time-varying Reproduction Number for Pulmonary Tuberculosis in Iran: A Registry-based Study from 2018 to 2022 Using New Smear-positive Cases." *Infectious Disease Modelling* 9 (3): 963–74.
- Rather, N. A., and T. A. Rather. 2017. "New Generalizations of Exponential Distribution with Applications." *Journal of Probability and Statistics* 2017: 1–9.
- Robert, C. P., and J. M. Rousseau. 2023. "A Special Issue on Bayesian Inference: Challenges, Perspectives and Prospects." *Philosophical Transactions of the Royal Society A* 381 (20220155): 1–4.

- Umeta, A. K., S. F. Yermosa, and A. G. Dufera. 2022. "Bayesian Parametric Modeling of Time to Tuberculosis Co-infection of HIV/AIDS Patients at Jimma Medical Center, Ethiopia." *Scientific Reports* 12 (16475): 1–18.
- Vehtari, A., A. Gelman, and J. Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-one-out Cross-validation and WAIC." *Statistics and Computing* 27: 1413–32.
- Wei, Z., S. T. B. Choy, T. Wang, and X. Zhu. 2025. "Bayesian Stochastic Frontier Models Under the Skew-normal Half-normal Settings." *Journal of Productivity Analysis* 64: 81–91.
- Weibull, W. 1951. "A Statistical Distribution Function of Wide Applicability." *Journal of Applied Mechanics* 18: 293–7.
- World Health Organization. 2020. *Global Tuberculosis Report 2020*. Geneva: World Health Organization.
- World Health Organization. 2023. *Global Tuberculosis Report 2023*. Geneva: World Health Organization.
- World Health Organization. 2024. *Global Tuberculosis Report 2024*. Geneva: World Health Organization.
- Yezli, S., Y. Yassin, A. Mushi, F. Maashi, N. M. Abdelmalek, A. H. Awam, et al. 2023. "Undiagnosed and Missed Active Pulmonary Tuberculosis During Mass Gatherings: A Prospective cross-sectional Study from the Hajj Pilgrimage." *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology* 42 (6): 727–40.
- Zamanah, E., S. Nasiru, and A. Luguterah. 2022. "Harmonic Mixture Weibull-G Family of Distributions: Properties, Regression and Applications to Medical Data." *Computational and Mathematical Methods*. <https://doi.org/10.1155/2022/2836545>.