EAJOS Vol. 2,No. 1, June 2026 ISSN (Online): 3088-5426 | ISSN (Print): 2953-2728

Eastern Africa Journal of Official Statistics

Website: https://journal.eastc.ac.tz/index.php/eajos/index

Bayesian and Frequentist Approach for the Mixture Cure Models with Generalized Log-logistic Baseline: An Application to Cancer Data

[1]Beatrice Amin Njau* and [2]Magreth Anga Kimaro

[1,2] Eastern Africa Statistical Training Centre. P.O. Box 35103, Dar es salaam.
*Corresponding Author's Email: beatrice.njau@eastc.ac.tz

KEYWORDS

Bayesian approach; censored; frequentist approach; mixture cure model; susceptible individuals.

ABSTRACT

Most event-data studies assume that everybody involved in the study will eventually encounter an instance of interest; nevertheless, it is anticipated that some of these participants will be exposed to the event. Cure models are frequently employed in time-to-event analysis to handle this data type. In mixed cure models, the target population is considered a mix of susceptible and non-susceptible individuals. The statistical analysis of these models provides the probability of cure (incidence model) and time-to-event in the vulnerable sub- population (delay model). This research presented a maximum likelihood estimate (MLE) and a Bayesian analysis for the six-parameter Generalized Log-Logistic (GLL) mixture cure model with cured, censored, and covariate variables. A mixture cure model with a GLL baseline is proposed to account for the rate of cured participants in the analysis. To apply various parametric hazard-based regression models, we recommend the GLL baseline distribution, which is based on a reasonable baseline hazard. The recommended model performs well on real-world data sets, highlighting the need for adaptive parametric regression formulations with time-dependent and time-independent covariates to assess hazard function and rate across time. Bayesian survival analysis with a non-informative prior outperformed the MLE method in terms of simulation parameter bias. The Generalized Log-Logistic model fits real cancer data better than the other survival models.

1 Introduction

It is well known that cure rate models, also known as survival models with a cure fraction, are becoming more prevalent in data analysis across multiple fields. Cure rate models, for instance, have been used to simulate time-to-event data from cancers such as melanoma, leukemia, breast, prostate, non-Hodgkin lymphoma, head, and neck cancer (Omer at el., 2020). A significant proportion of patients with these diseases are regularly cured. Cure rate models are a type of model for censored survival data in which some subjects do not develop the event of interest despite being followed for an extended period of time (Omer at el., 2020).

According to de Castro and Gomez (2020), some who did not witness the event of interest are often referred to as" cured subjects" or" long-term survivors." There is a strong case for cured subjects in cancer clinical trials because if the therapies are successful, the cancer is removed and the subject will not experience the recurrence of the disease. This is particularly true for cancer patients in the beginning phases. Similar incidents can be found in other disciplines, including economics and social studies de Castro and Gomez (2020), Cooner et al. (2007), and (Omer at el., 2020). The literature on cure rate models is vast and rapidly expanding, (Khan and Khosa,2016); (Gallardo et al.2018). The mixture cure model (Ibrahim et al.,2001), (Berkson and Gage, 2012), and the proportion time cure model are two of the most popular cure rate models (Hannah and Quinn,1979). For the latter, Bayesian's formulation is a game changer (Adnan and Arasan,2018). Bayesian inference in cure rate models is discussed in detail by (Teo et al.2012), (Tsodikov et al.,2003), and de Castro and Gomez (2020), examined advancements in cure rate models.

Tsodikov et al. (2003), proposed the mixture cure model (MCM) first, and (Ibrahim et al.,2001), and (Berkson and Gage.,2012) expanded on it three years later. Numerous researchers, for example, (Adnan and Arasan,2018), (Muse et al., 2021), (Omer at el., 2020), and (Tsodikov et al., 2003), suggested that the mixture cure model (MCM) does not support the proportional hazard property for the entire population.

Cure rate models, which are viewed as a broader variation of classic survival models, play an important role in assessing long-term survival data. There models have been the field of study since the 1990's (Omer at el., 2020), (Tsodikov et al., 2003). These are two types of cure fraction models, mixture cure models and non-mixture cure models. According to the mixture cure models, the data set is split into two clusters, cured and those who are uncured individuals (Omer at el., 2020), (Tsodikov et al., 2003).

The non-mixture cure model (Tsodikov et al., 2003), was proposed to validate the cured and uncured individuals, all findings allow the significant consequences of indicators of the probability of being cured. Tsodikov et al. (1996), stated in their earlier findings on the enclosed cumulative hazard model utilizing statistical inference that such a confined cumulative hazard model is simple in terms of computation, has a better perception, and its framework is versatile for the survival function, which could provide some competitive advantage when going to construct procedures for maximum likelihood estimation (Tsodikov et al., 2003). A variety of approaches to modeling the non-mixed cure model have been proposed in the literature, (Omer at el., 2020), (Tsodikov et al., 2003), (Suga et al., 2003) and (Muse et al., 2022). They defined the marginal distribution of T and offered a generalization for the confined cumulative hazard model (1), formula as;

$$S(t) = E_N[P(T \geq t|N)] = m\big[\log\big(S_N(t)\big)\big] \quad (1)$$

Whereby, E_{N} is the likeliness determined over N, $S_{\mathrm{N}}(t)$ ia a survival function of $X_1, \ldots, X_{\mathrm{N}}$ underlying event times, N stands for the number of cancer patients, it can have any limited integer-value distribution, m (.) is the corresponding moment generating function (mgf), $m(t) = E(exp)^{tN}$ well as a cure rate described by $P_r(N = 0) = m(\inf)$ (Omer at el., 2020) and (Muse et al., 2022). Due to the work in (Kass et al., 1998), additional distribution of N further than the Poisson distribution, namely the geometric and negative binomial distributions, can be implied (Tsodikov et al., 2003), (Muse et al., 2022). Scientists worry more about the standard of the dispersion of susceptible respondents' survival times, as well as the impact of explanatory variables influencing cure rates and survival times when modeling survival data with cured possibility. Such two factors are essential when analyzing lifetime data. The cure fraction model, in equation (2) could be used to investigate lifetime data in the appearance of cured subjects, as well as factors that influence curing likelihood, must display an appropriate survival time distribution (Omer at el., 2020), and (Tsodikov et al., 2003). Ignoring this specification may lead to untrusted parameter estimations. Cure rate models are preferable to classical

survival models for survival data with cured likelihood and the capacity to spot a suitable distribution for survival time (Omer at el., 2020), (Tsodikov et al., 2003), and (Muse et al., 2022).

$$P(v = 1|X) = \frac{\exp(\gamma' X)}{1 - \exp(\gamma' X)} \tag{2}$$

where $P\,(v = 1 \mid X)$ is the cure fraction model for a susceptible subject with q-dimensional covariates X and γ, an unidentified regression parameter vector of significant interest, v specifies, the value 0 or 1, whether the sampled individual is susceptible or not (Omer at el., 2020), (Muse et al., 2022). The literature provides a deep investigation of cure fraction models, including the selection of best model that fits the criteria and the effects of covariates on the likelihood of becoming cured and the survival times of susceptible individuals. We refer the reader to (Omer at el., 2020), (Tsodikov et al., 2003), and (Muse et al., 2022).

for further information on the consequences of covariates on the response variable. Numerous studies measured the effectiveness of cure models in terms of estimator effects on cure fraction and survival times using the mixture cure model and the confined cumulative hazard model), (Tsodikov et al., 2003), and (Muse et al., 2022).

Nevertheless, a few other compositions of non-mixture cure models, have been used. Therefore, in this study, we suggested a mixture cure model with generalized log-logistic, which is a flexible survival time distribution. In this paper, we look at the performance of the Bayesian and frequentist mixture cure models, both of which are sensitive to various aspects of the rate of failure function. For instance, consider the presence of the cure rate, censored findings, and predictors in the cancer set of data bmt, which is available in the R package survival (Lawless, 2011), and (Tsodikov et al., 2003). This data is statistically analyzed using the maximum likelihood estimation method. As a result, the primary goal of this study is to develop a generalized log-logistic mixed cure model. The new proposal's distributional parameters and covariates are estimated using the Bayesian and maximum likelihood estimation strategies for assessing model parameters (Muse et al.,2022).

The model's generalized log-logistic mixed cure performance is compared with the most widely used lifetime mixed cure models, log-logistic BX-II and Weibull (Boag,1949)

Convergence diagnostic methods focused on the Markov chain Monte Carlo (McMC) strategy were also used.

The entirety of the paper is organized as follows. Section 2 describes the Cure models, the GLL distribution functions, and its sub-model distributions, as well as some of its properties. Section 3 derives the basic hazard and some mathematical properties of the GLL distribution. Section 4 discusses the motivated data and GLL simulations using Bayesian and frequentist approaches, as well as specific cases. Section 5 presents the inferential processed log-likelihood

functions of such six proposed cure models. The model selection procedure is described in Section 6. Section 7 summarizes the findings.

2 Cure Model

Two general classes of cure models have been presented in the literature to accommodate lifetime data in clinical studies, are the mixture cure models and non-mixture cure models (Omer at el., 2020). However, apart from therapies, these models can be used to investigate real-world data in a variety of fields, including economics, reliability, criminology, sociology, education, and marketing, to mention a few (Omer at el., 2020), (Tsodikov et al., 2003), and (Muse et al., 2022).

The proposed model varies depending upon the situation of interest to the researcher; the overall theme is to examine time until the event occurs, however the event never occurs in some participants (Omer at el., 2020). This part of the study will go over the mixture cure model and its six different compositions.

2.1 Mixture Cure Model (MCM)

A mixture cure model (MCM) is a popular method for modeling long-term survivor data. The MCM has the advantage of allowing predictor variables to have varying effects on cured patients and vulnerable individual's survival times (Omer at el., 2020), and (Muse et al., 2022). On the other side, this model cannot prove the property of proportional hazard functions (Omer at el., 2020). Besides that, the MCM lacks a biological interpretation, especially in cancer relapse studies (Omer at el., 2020), and (Muse et al., 2022).

Let T denote the time of occurrence and $\eta\epsilon\,(0,1)$ the probability of cure. Assume that $S^*(t)$ and $f^*(t)$ are the survival and density functions for susceptible subjects, respectively (Omer at el., 2020), (Tsodikov et al., 2003), and (Muse et al., 2022). As a result, the overall population survival function is negatively affected.

$$S(t|\eta) = \eta + (1-\eta)S^*(t) \tag{3}$$

The associated probability density function(pdf) is $f(t|\eta) = (1-\eta)f^*(t)$. For $S^*(t)$, various parameter distributions, such as exponential, Weibull, and log-logistic, can be selected (Hurvich andTsai,1989)

Assume we get a random sample size p and examine the pair (t_j, κ_j), in which κ_j is the censoring indicator variable, with zero and one features for censored and uncensored findings, respectively, and $j = 1, \ldots, m$ (Omer at el., 2020), and (Tsodikov et al., 2003).

. Equation 3 is the likelihood function for the MCM model (4)

$$L_{mcm} = \prod_{j=1}^{p} [f(t|\eta)]^{\kappa j} \times [S(t|\eta)]^{1-\kappa j}$$

$$L_{mcm} = \prod_{j=1}^{p} \left[(1-\eta) f^*(t_j)\right]^{\kappa j} \times \left[\eta + (1-\eta) S^*(t_j)\right]^{1-\kappa j}$$

Let $\psi = (\eta, \kappa, \zeta, \alpha, \varpi, \beta)^r$ where α, β and $\eta, \varpi, \kappa, \zeta,$ are two shape parameters and four scale parameters; and suppose the mixture cure model 3, and the log-likelihood function for ψ is then used, as showing in the equation 5 and 6.

$$L_{mcm}(\psi) = \ln(1-\eta) \sum_{j=1}^{p} \kappa_j + [ln(\alpha) + ln(\varpi) + ln(\beta) + ln(\varsigma)] + (-\zeta\varpi) + \sum_{j=1}^{p} (\kappa_j) + (\beta - 1) \sum_{j=1}^{p} \kappa_j ln\left(1 + (1 - exp(-\zeta\varpi)^{\alpha-1})\right) + \tag{6}$$

$$\sum_{j=1}^{p} \boldsymbol{\kappa}_j \, exp\{-\alpha[1+(1-exp(-(\zeta\varpi)^{\alpha})^{\beta}-1]\} + \sum_{j=1}^{p}(1-\kappa_j) ln\left\{\eta + (1-\eta)\left[1+\left(1-exp(-(-\zeta\varpi)^{\alpha})\right)^{\beta}-1\right]\right\}$$

3 Baseline Hazard

The log-logistic distribution is often utilized in oncology research in survival analysis because its hazard rate function is changeable and its parameters are easy to estimate. Nonetheless, more advanced parametric models are frequently required in medical studies. To accomplish this, the log-logistic distribution has also been extended to different classes of parametric distributions, namely the generalized log-logistic distribution (Omer at el., 2020), (Noor et al.,2021), (Tsodikov et al., 2003), and (Muse et al., 2022).

odd log- logistic (Chen et ai., 1999), alpha power log-logistic and exponentiated log-logistic geometric (Omer at el., 2020), (Muse et al., 2021) and (Muse et al., 2022).

Muse et al. (2021), distributions are more details about the established extensions of the classical log-logistic and are obtainable in (Omer at el., 2020). Furthermore, right-censored survival data are typically seen in cancer clinical trials with periodic follow-ups, where the survival times are made up of some exact observed and some right-censored observations (Omer at el., 2020), (Muse et al., 2022).

The log-logistic distribution is sealed in the proportional odds (PO) and mixture cure models. So, under the proportional hazard (PH) model, it is not closed. At first, when generalized,

it could be used as a starting point for all parametric hazard-based regression models (Omer at el., 2020), (Muse et al., 2022). Klein et al. (2014), due to its mathematical versatility and adaptability, the study proposed a generalized Weibull model. As a result, we concentrate on assessing right-censored data in various hazard-based

regression models employing a generalized log-logistic (GLL) baseline in this paper. There are three parameters in the GLL distribution (a scale parameter, and two shape parameters). The distribution can be modified, and the two shape parameters allow for a variety of hazard shapes (Omer at el., 2020), (Muse et al., 2022). Assume that the GLL is a probability distribution of the susceptible subject's survival time. The hazard rate function (7) for this distribution is as follows:

$$h(t;\psi) = \frac{\alpha\varpi(\varpi t)^{\alpha-1}}{1-(\zeta t)^{\alpha}}, t \geq 0, \ \ \alpha, \varpi, \zeta > 0 \tag{7}$$

where $\varpi > 0, \alpha > 0, \zeta > 0$ are the distribution's unknown parameters, and $\boldsymbol{\psi} = (\varpi, \alpha, \zeta)^{r}$ (Noor et al., 2021), (Muse et al., 2022). To accomplish this, the log-logistic distribution has also been extended to different classes of parametric distributions, namely the generalized log-logistic distribution (Omer at el., 2020), and (Muse et al., 2022).

$$H(t;\psi) = \frac{\varpi^{\alpha}}{\zeta^{\alpha}}, \log[[1+(\zeta t)^{\alpha}]] \quad t \geq 0, \ \ \varpi, \alpha, \zeta > 0 \tag{8}$$

The GLL distribution is represented by the symbol $T \sim GLL$ $(, \alpha, \zeta)$ (8)and Its survival function (9) is manifested as:

$$S(t;\psi) = [1+(\zeta t)^{\alpha}]^{\frac{\varpi^{\alpha}}{\zeta^{\alpha},}} \quad t \geq 0, \ \ \varpi, \alpha, \zeta > 0 \tag{9}$$

3.1 Special Cases of the GLL Baseline Hazard

Equation 7 contains different special cases of GLL distribution (Muse et al.,2021), (Tsodikov et al., 2003), and (Muse et al., 2022).

. The following are the distributions: - Log-logistic (LL): when $\varpi = \zeta$, equation 7 reduces to the hrf of

LL, which $h(t;\psi) = \frac{\alpha\varpi(\varpi t)^{\alpha-1}}{1-(\varpi t)^{\alpha}}, t \geq 0, \ \ \alpha, \varpi, > 0$ - 3- parameter Burr-XII (BXII-3): when

$\zeta = \varpi p^{\frac{-1}{\alpha}}$, where $\zeta > 0$, and $p = {}^{1}/_{\varsigma}$, equation 7 gives us to the hazard rate function of a BXII-3, which

$h(t, \varphi) = \frac{\alpha\varpi(\varpi t)^{\alpha-1}}{[1+p(\varpi t)^{\alpha}]}, t \geq 0,$ $\varpi, \alpha, p > 0,$ 2- parameter Burr-XII (BXII-2) (Omer et al.,2020), (Tsodikov et al., 2003), and (Muse et al., 2022).

: when $\zeta = 1$, equation 7 decreases to a hazard rate function BXII-2 (10), which

$$h(t, \varpi) = \frac{\alpha\varpi(\alpha\varpi)^{\alpha-1}}{[1+t^{\alpha}]}, t \geq 0, \quad \varpi, \alpha > 0. \tag{10}$$

- Weibull (W): when $\zeta \to 0$, equation 7decreases to the hazard rate function of the Weibull (Tsodikov et al.,2003), which

$$h(t; \boldsymbol{\psi}) = \alpha\, \varpi\, (\varpi\, t)^{\alpha-1}, \quad t \geq 0, \quad \varpi, \alpha > 0 \tag{11}$$

- Exponential (E) when $\alpha = 1$, and $\zeta \to 0$, equation 7 decreases to the hazard rate function of an Exponential (Muse et al., 2021), which $h(t; \boldsymbol{\psi}) = \varpi, \ \varpi > 0$

-Exponentiated Exponential (EE)

$$h(t; \boldsymbol{\psi}) = (\varpi)^{a}, \quad t \geq 0, \quad \varpi, a > 0 \tag{12}$$

Standard Log-logistic (SLL) (Omer at el., 2020), and (Muse et al., 2022).

- : when $k = \zeta = 1$, equation 7 diminishes to an SLL's hazard rate function (Refley and Lewis,1992), which;

$$h(t, \varpi) = \frac{\alpha t}{[1+t^{\alpha}]}, t \geq 0, \quad \alpha > 0. \tag{13}$$

4 Motivated Data

Bone marrow cancer is a catch-all term for cancers like multiple myeloma that begin when cells in the bone marrow grow and spread uncontrollably. It is one of the major causes of cancer-related deaths in men and women in the United States, claiming the lives of 400 people in 2002. In the United States, this disease was expected to kill 12,640 people by 2020, (7,090 men and 5,550 women). In 2020, bone marrow was expected to kill 117,077 people worldwide (Tsodikov et al., 2003).

The cumulative overall risk for bone marrow cancer is roughly 5% to 6%, with inheritable and lifestyle factors influencing this risk. According to (Maller and Zhou.,1996), 20% to 30% of all bone marrow cancer cases have a possible clearly defined inherited cause, and 3% to 5% of bone marrow cancers happen in genetically defined high-risk bone marrow cancer family clinical

manifestations. Even though the genes responsible for reasonable bone marrow cancer cases are still being identified, many of the genes responsible for high-risk bone marrow cancer cases have already been identified (Maller and Zhou.,1996), These biological discoveries have been translated into medical practice, resulting in improved risk assessment via the use of genetic screening. Although most of bone marrow patients are cured by their initial treatment, distinguishing them from uncured patients is impossible. As a result, it is critical to precisely estimate the likelihood of cure in order to be prepared to provide further treatment to improve the survival of uncured bone marrow patients (Omer at el., 2020), and (Muse et al. 2022). In this study, we take into account bone marrow data from R bmt in the Survival package on cancer, obtained from a clinical trial https://CRAN.R-project.org/package=smcure. A total of 91 patients were included in this trial, with 49% of them receiving the treatment. The remaining people were right-censored.

Various studies on the analysis of bone marrow data can be found in the literature. Kass et al. (1998), assumed that indicators only affected the cure fraction, whereas (Suga et al.2003), did not include covariates in their analysis. We use a Bayesian and a frequentist approach, as well as the GLL distribution, to examine the impact of variables on the cure rate and survival time of susceptible individuals.

5 Procedures for Making Inferences

In this section, we will look at the frequentist (via maximum likelihood estimation) and Bayesian inference procedures for the generalized log-logistic mixture cure model (Muse et al., 2022). The simulation aspect of the study employed non-informative / subjective priors' algorithms. In addition, a bone marrow cancer data set related to cured and uncured events was used in a real data set application.

5.1 The likelihood of the right-censored data

In this subsection, we look at predicting the parameters of the proposed model using right-censored testing data (Boag. 1949), and (Muse et al., 2022). Because of its appealing properties such as asymptotic unbiasedness, consistency, asymptotic normality, and asymptotic efficiency, maximum likelihood estimation is one of the most popular methods for estimating the parameters of modified linear regression models and continuous probability distributions. (Peng and Tailor, 2017), and (Muse et al., 2022).

The probability that certain individuals are unable to see the event of interest in its totality is among the reasons why time-to-event analysis requires the application of distinct techniques (Boag.

1949), and (Muse et al., 2022). Inadequate(non-informative) observations are those that have been censored or truncated and cannot be removed from the analysis. Furthermore, they must be successfully identified and handled in the statistical model. Further information on the estimation of censoring data, whether in distribution or regression, can be found in (Noor et al.2021), (Klein et al., 2014). and (Muse et al., 2022). Suppose there are n individuals concerned with survival time devoted by $T_1, T_2, \cdots, T_n$. Let that
the data are subject to right censoring, we observe $t_i = \min(T_i, C_i)$, where $C_i > 0$ refers to a potential censoring time for individual i. Letting $\kappa_i = \mathrm{I}(T_i, C_i)$ equal to 1 if $T_i \leq C_i$ and 0 otherwise. Let that a right censored random sample consisting of data (t_i, κ_i, x_i), $i = 1, 2 \cdots, n$, is available, where t_i is a censoring time or a life according to whether $\kappa_i = 0$ or 1, respectively and $\boldsymbol{x_i} = x_1, x_2, \cdots, x_n$ is an $n \times 1$ External covariates in a column vector for the i^{th} participant. We make the assumption of non-informative censoring, which means that the survival time distribution reveals nothing about the censoring, time distribution, and vice versa (Boag. 1949), and (Muse et al., 2022)

The presumption of non-informative censoring is justifiable once censoring is independent, (such as censoring is assumed random within any sub-population of concern), and/or random (such as, hazard rates for censored and uncensored findings that persist in the hazard collection are the same). For more information on uninformative censorship, t_i, and κ_i, under
non-informative censoring, are random variables with $P(t_i, c_i, \kappa_i = 0) = S(T_i < C_i) = S(t_i)$ and $P(t_i, \kappa_i = 1) = f(t_i) \times [f(t_i)]^{\kappa_i} [S(t_i)]^{(1-\kappa_i)}$ is the joint probability density function of t_i and κ_i, which is the i^{th} individual's contribution to the likelihood function (Kutal and Qian, 2017), and (Muse et al., 2022). Individuals $\boldsymbol{i}$ add value $\boldsymbol{f}(t_i)$ a probability function if an event happens at t_i, and $S(t_i)$ if individuals are censored at t_i. The likelihood function (14) is obtained by combining data from censored findings (Boag. 1949).

$$L(\xi) = \prod_{i=1}^{n} [f(t_i)]^{\kappa_i} [S(t_i)]^{(1-\kappa_i)} \tag{14}$$

in which $\boldsymbol{\xi}$ is a parameter vector explaining the T_i distribution. In terms of hazard rate and survival functions, the following probability density function explanation is used:

$f(t_i) = h(t_i)S(t_i)$.

The likelihood function (15)is also known as

$$\prod_{i=1}^{n} [h(t_i)S(t_i)]^{\kappa_i} [S(t_i)]^{(1-\kappa_i)} \tag{15}$$

Through simplifying we obtain equation (16);

$$L(\xi) = \prod_{i=1}^{n} [h(t_i)]^{\kappa_i} S(t_i) \tag{16}$$

In case of a parametric MCM Model, the log-likelihood function (4)can be expressed as follows;

$$L(\psi) = \sum_{i=1}^{n} \kappa_i log[h(t_i; x_i)] + \sum_{i=1}^{n} log[S(t_i; x_i)] \tag{17}$$

where; $\psi = (\xi, \beta)$, ξ is a vector distribution parameter, and $\boldsymbol{\beta}$ is a vector of Bayesian regression coefficient. The Newton-Raphson optimization algorithm can be used to maximize the equation (17) directly, therefore testing of hypothesis and estimating the intervals of model parameters can be accomplished using the MLE estimators' roughly normality (Omer at el., 2020), (Boag. 1949), (Kass et al.,1998), and (Muse et al., 2022).

In GLL MCM, the likelihood function (4) is given by

$$L(\psi) = \prod_{i=1}^{n} \left[\frac{\alpha {t_i}^{\alpha-1} \left(\varpi e^{x^{\alpha\prime} \beta^*} \right)^{\alpha}}{\left(1 + \left(t_i \zeta e^{x^{\alpha\prime} \beta} \right)^{\alpha} \right)} \right]^{\varpi_i} \prod_{i=1}^{n} \left[\left[1 - \left(t_i \zeta e^{x' \beta^*} \right)^{\alpha} \right]^{\frac{\kappa_i^{\alpha}}{-\zeta^{\alpha}}} \right] \tag{18}$$

In our case, let us imagine $d = \sum_{i=1}^{n} \kappa_i$ (Omer at el., 2020). The full log-likelihood function is expressed using the likelihood for the generalized log-logistic MCM model (18) as:

$$L(\psi) = dln\sigma + (\alpha - 1) dln(t_i) + \alpha dln\left(\varpi e^{x' \beta^*} \right) - dln\left(1 + \left(t_i \zeta e^{x' \beta^*} \right)^{\alpha} \right)$$

$$- \left(\frac{\varpi}{\alpha} \right)^{\alpha} \sum_{i=1}^{n} ln\left(1 + \left(t_i \kappa_i e^{x' \beta^*} \right)^{\alpha} \right) \tag{19}$$

where; $\boldsymbol{\psi}$ indicates the aggregate of all model parameters. The distributional parameters are included in our model. $\xi = (\alpha, \kappa, \varpi)$ that define the basic hazard and survival functions $h_0(t)$ and $S_0(t)$, the regression coefficients $\boldsymbol{\beta}$, leading to $\psi = (\xi, \boldsymbol{\beta})$. Obtaining the MLE's of $\psi' = (\alpha, \kappa, \varpi)$ and $\boldsymbol{\beta}$, we maximize equation (19) effectively in regards to (α, κ, ϖ) and $\boldsymbol{\beta}$. or we can get $(\hat{\alpha}, \hat{\kappa}, \hat{\varpi}, \hat{\beta})$ by solving the first derivative of the log likelihood function with regard to $(\alpha, \kappa, \varpi, \beta)$ and equating to zero. However, it is significant to solve the equation $\psi' = (\alpha, \kappa, \varpi,$ and $\boldsymbol{\beta})$ analytically, such as numerical iteration method (Newton-Raphson algorithm) (Boag. 1949), and (Muse et al., 2022)

6 Simulation

In all simulation studies, the inferential properties of the proposed models were analyzed, and the MCM model with GLL baseline hazard (Generalized log-logistic MCM) was incorporated and aligned to the corresponding true-creating model from the GLL MCM (Boag,1949), and (Muse et al., 2022). Besides that, for each model's regression coefficients, the bias, root mean square error (RMSE), and probability coverage were used to measure the reliability of the estimators (Boag,1949). In addition, we used the Bayesian approach to diagnose the convergence scenario.

6.1 Monitoring the performance of the estimators

The difference between the expected value of a statistic and the true value of the associated parameter is characterized as bias (20) in statistics. As a result, bias is a measure of an estimator's systematic mistake. The bias indicates the estimator's distance from the true value of the parameter. The bias of the estimators is calculated by using the following:

$$Bias(\hat{\psi}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\psi} - \psi) \qquad (20)$$

A positive bias indicates overestimation, while a negative bias indicates underestimation (Geweke et al.,1991) The average difference between a statistical model's estimated and actual values is measured by RMSE. It is the standard deviation of the residuals in mathematics. The precision of the estimators, measured as root mean square error (RMSE) (21), makes the overall accurateness and is calculated:

$$RMSE(\hat{\psi}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\psi} - \psi)^2} \quad (21)$$

The precision of the estimates is determined by this metric. Influence estimates are more precise when the RMSE is low.

Another accuracy statistic is the Naive standard error (22), which is computed by dividing the posterior (*SD*) standard deviation by the square root of sample size. As a result, the larger the sample size, the smaller the standard error. Instead of assessing posterior uncertainty, the Naive SE utilizes simulation error (Glady,2003).

$$NaiveSE = \frac{Posterior\ SD}{\sqrt{n}} \quad (22)$$

6.2 Coverage probability (CP)

The coverage probability (23) of 95% is the proportional of N simulated sets of data for which real estimates were used in the 95% confidence interval (CI). The far more precise the estimates, the closer the results are to a coverage probability of 95%, (Glady,2003), and (Njau et al.,2022) describes the CP definition,

$$CP = \hat{\psi} \pm 1.96 \times SE(\hat{\psi}) \quad (23)$$

6.3 Model Choice

The Akaike Information Criteria (AIC) presented by (Suga et al.,2003), (Cai et al.,2012) was used to evaluate the similarity of hazard function based on various distributions. The Bayesian Information Criterion (BIC), introduced by (Rubio et al., 2019) and Hannan-Quinn information Criteria (HQIC) suggested by Edward J Hannan and Barry G Quinn (Hannan and Quinn,1979), and (Wang et al.,2018). Consistent Akaike's Information Criterion (CAIC), introduced by (Purvis et al.,1995) and Bias Consistent Akaike Information Criterion (BCAIC) introduced by (Hurvich and Tsai,1989), and (Peng and Yu, 2021). A lower information criterion value shows an excellent model fit. Also, it is critical to use multiple information criterion because they are used to measure different angles of the model for justification. The following are the definitions of the AIC, BIC, HQIC, BCAIC, and CAIC:

$$AIC = -2\ln\left[L\left(\psi\right)\right] + 2h,$$

$$BIC = -2\ln\left[L\left(\psi\right)\right] + h\ln(p)$$

$$HQIC = -2\ln\left[L\left(\psi\right)\right] + 2h$$
$$\ln[\ln(p)],\ BCAIC = -2\ln\left[L\left(\psi\right)\right] +$$
$$h[\ln(p) + 1] \text{ and}$$
$$CAIC = -2\ln\left[L\left(\psi\right)\right] + 2h + [2h * (h+1)] / (n - \ln\left[L\left(\psi\right)\right])$$

where $L(\psi)$ denotes the likelihood function, h denotes the free parameter's number in the model, and p the number of observational data according to Table 1, GLL has the smallest AIC and thus is the best fit model for the data.

6.4 Limitations of the study.

Some risk factors, such as older age, a family history of cancer, tobacco use, being overweight, alcoholic beverages certain viral infections (e.g., HPV), specific chemicals, and contact with radiation (for example, ultraviolet radiation), may have been overlooked or incompletely collected, potentially contributing to the situation.

6.5 Significance of the Model

The proposed model has several advantages for various disciplines. The model's good estimation can provide a framework for informing the formation or design of public health policies. In addition, it provides a framework for future research; and insight into the relationship between risk factors and hazards to the onset of cancer.

6.6 An illustration based on actual data from CRAN R packages The utility of the GLL distribution is demonstrated in this section by using real datasets. The shape of Figure 1, reveals the GLL distribution's suitability, Figure 2 shows the overall survival curve, figure 3 presents the survival probability for the two treatment groups for our

datasets. The bone marrow data shows a decrease in failure rate followed by increasing in failure rate, and then again, the decrease in failure rate as described in Figure 1

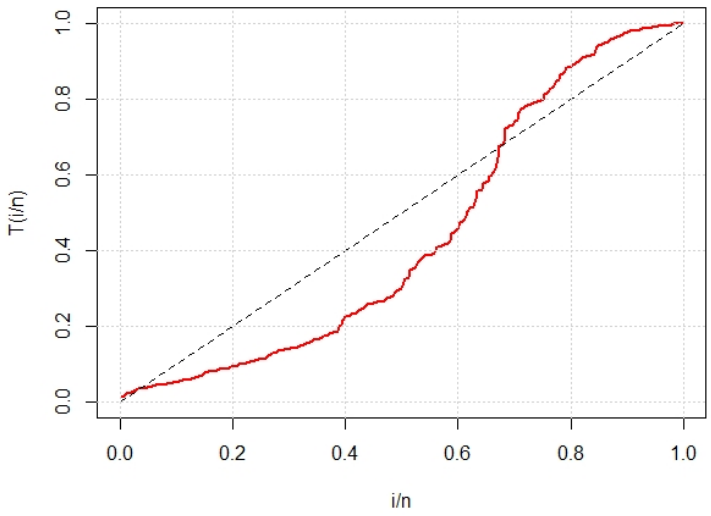

Figure 1: The total time on the test (TTT) plot for the data set

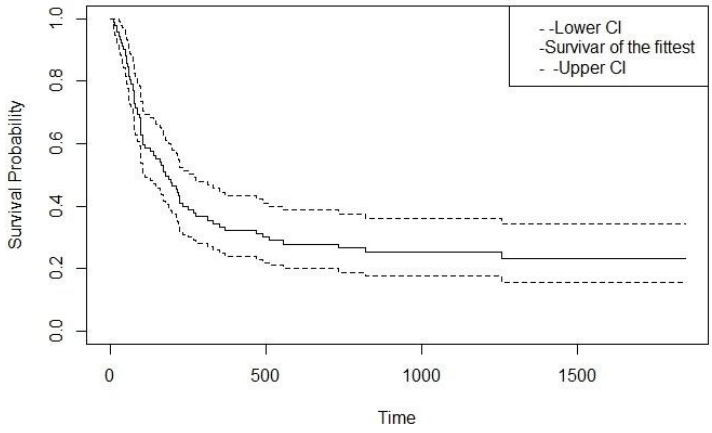

Figure 2: The overall survival curve

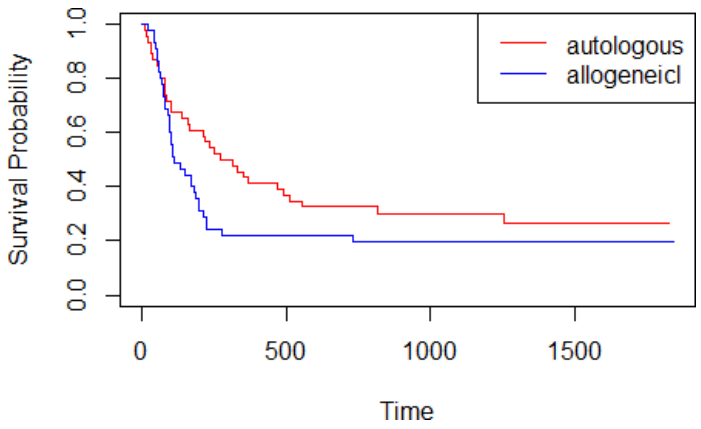

Figure 3: Survival curve for the two treatment groups

6.7 Sub-distributions of the GLL distribution

Weibull (W), exponential (E), Exponentiated Exponential (EE), BXII, Exponentiated Weibul (EW), Log-logistic (LL) distributions are special or sub-models of the GLL distribution. Table 1 shows sub-models of the GLL distribution for various parameter values.

Table 1: Maximum likelihood results for the GLL distribution and its sub models, using a Frequentist approach

Models	Parameters and True values	Estimates	AIC	BIC	CAIC	BCAIC	HQIC
GLL	Constant=0.5	1.13	754.12	765.73	753.66	772.73	759.98
	ϖ=1	2.46					
	α=0.5	6.45					
	ζ=1	4.09e-05					
LL	Constant=0.5	1.28	762.15	773.11	761.89	776.11	766.54
	ϖ=1	1.1					
	α=0.5	0.27					
W	Constant=0.5	1.608	772.41	783.37	772.16	786.37	776.80

	ϖ =1	1.005					
	α=0.5	0.291					
BXII	Constant=0.5	1.285	761.54	761.28	761.28	775.50	765.94
	ϖ=1	0.864					
	α=0.5	0.259					
EE	Constant=0.5	1.074	771.76	782.72	771.51	785.72	776.15
	ϖ=1	0.67					
	a=0.5	0.291					

Table 1 show the maximum likelihood estimates (MLE) and information criterion value systems for AIC, BIC, CAIC, BCAIC, and HQIC, acquired by the mixed cure model based on estimated distribution of case involved the bone marrow data. We can see from the findings in Table 1 that the GLL distribution based on the mixed cure models also offers a better fit to the data than other distribution-specific cases. We discover that the fifth cure models fit the data evenly well when we assume a GLL distribution and compare LL, the W, EE model and BXII.

Table 2: Numerical overviews of the GLLMCM model's posterior properties.

Characteristic	Parameters					
	alpha	beta C1	beta C2	beta U	zeta	varpi
Mean	1.9568	-29.8013	-12.5681	0.1001	0.1617	0.0496
Median	1.9471	-26.5964	-10.7766	0.1067	0.0331	0.0931
Mode	1.9	-15	-5	0.15	0.025	0.075
Variance	0.1404	380.8947	636.9609	0.0578	0.0010	0.0164
Skewness	0.3773	-0.8322	-0.3055	-0.0164	2.1528	2.416
Kurtosis	0.2435	0.5061	0.2277	0.3343	5.6533	7.225
Minimum	0.7445	-146.5047	-138.6164	-1.0849	0.0072	0.0159
Maximum	4.1347	-0.0237	86.9673	1.1880	0.2757	1.1230
SD	0.37482	19.51652	25.23808	0.24042	0.03181	0.0374
Naive SE	6.843e-04	3.563e-02	4.608e-02	4.389e-04	2.344e-04	5.808e-05
Time series SE	0.003097	0.069654	0.073356	0.001263	0.006505	0.001579

2.5 percentile	1.30212	-75.14970	-66.17133	-0.36787	0.03525	0.01483
Q_1	1.70643	-41.66272	-28.62653	-0.05445	0.0643	0.02358
Median / Q_2	1.94710	-26.59640	-10.77661	0.10676	0.09316	0.360
Q_3	2.20777	-14.33568	4.48561	0.26807	0.16792	0.0613
97 percentiles	2.7686	-2.9028	33.4440	0.5756	0.5302	0.1381

6.8 Methods for McMC convergence diagnostics

This paper employed GLLMCM models inside a Bayesian framework to assess the efficacy of Bayesian estimators as presented in Table 2. The McMC used the Metropolis-Hastings and Gibbs sampler algorithms to determine the convergence of the posterior parameters as shown in Table 3, to determine stationarity. A variety of graphical methods can be used to test McMC convergence. Four popular graphical tools for McMC convergence diagnostics can be used to investigate a chain's convergence and mixing characteristics plots such as trace plots Figure 4, Geweke diagnostic plots Figure 6, Gelman diagnostic plots Figure 7, density plots Figure 5, and Paired plots Figure 8 (Muse et al.,2022), and (Purvis et al., 1995). Time series plots (as well referred to as essential trace plots) show how quickly the chain mixes. A time series plot depicts the chain's realization rather than the iteration number. Such tools examine the value chains obtained for each parameter or regression coefficient to determine whether or not the simulation process has reached a stable state (Muse et al.,2022).

Table 3: Some commonly used statistical test for convergence diagnostics are outlined

Parameters	Geweke	Reflex-Lewis	Heidelberger	Welch				
	diagnostic Pr> |z| values	sample(N) True	P-values	Test of stationarity	Test for half-width	RMSE	BIAS	Coverage probability
alpha	-0.5571	17188	0.287	passed	0.0254	0.0660	-0.0621	95
beta C1	0.6934	7819	0.4543	passed	0.7652	0.5106	-0.0305	96
beta C2	-0.4273	7329	0.7860	passed	0.73056	0.7533	0.6312	96
beta U	-0.8658	7819	0.06813	passed	0.0126	0.1501	-0.1498	95.5
varpi	-0.4309	38192	0.4463	passed	0.0114	0.0076	-0.0045	94
zeta	-0.5521	39174	0.3792	passed	0.046	0.025	-0.0216	95

Andrew Gelman (Kass et al.,1998), proposed that the maximum permissible limit of potential scale reduction factor (PSRF) and multivariate potential scale reduction factor (MPSRF)

be near 1, $\hat{R}$ <1.1, and the effective number of sample size simulation draws be greater than or equal to 100, for checking the convergence of McMC simulations. The PSRF in Table 4 for each parameter is less than 1.1, the number of sample size simulation draws is greater than 100, and the Naive SE for all distributional parameters and regression coefficients is less than the standard deviations (SD), as predicted, indicating that the McMC algorithm has converged to the posterior distribution (Omer at el., 2020).

Table 4: Estimates, Potential scale reduction factors(psrf) and Multivariate psrf for distribution parameters and regression coefficients, respectively.

Parameters	Estimate	$\hat{R}$	Upper C.I.
alpha=1.9	1.9622	1.00	1.00
beta C1=-30	-29.9695	1.00	1.00
beta C2=-12	-12.6313	1.00	1.00
beta U=0.09	0.09982	1.00	1.00
varpi=0.04	0.0445	1.01	1.04
zeta=0.14	0.1421	1.02	1.04
Multivariate PSRF	1.01		

6.8.1 Paired plots

Cross-correlation generates a tile plot with the correlations between all parameters to diagnose potential convergence problems caused by highly correlated parameters. The absolute scale argument specifies whether the scale must be between -1 and +1. The whole scale is used by default, which puts the cross-correlation issues in context. However, in cases where cross-correlation between parameters is not a significant issue, using relative scales to identify the most problematic parameters may be helpful.

Other well-known formal test statistics for McMC technique convergence, such as, Heidelberger- Welch test statistic (Heidelberger and Welch,1983), (Elgarhy et al.,2017), Geweke diagnostic (Kass et al.,1998), and Raftery-Lewis's diagnostic (Naslina et al., 2020) were presented in Table 3 above.

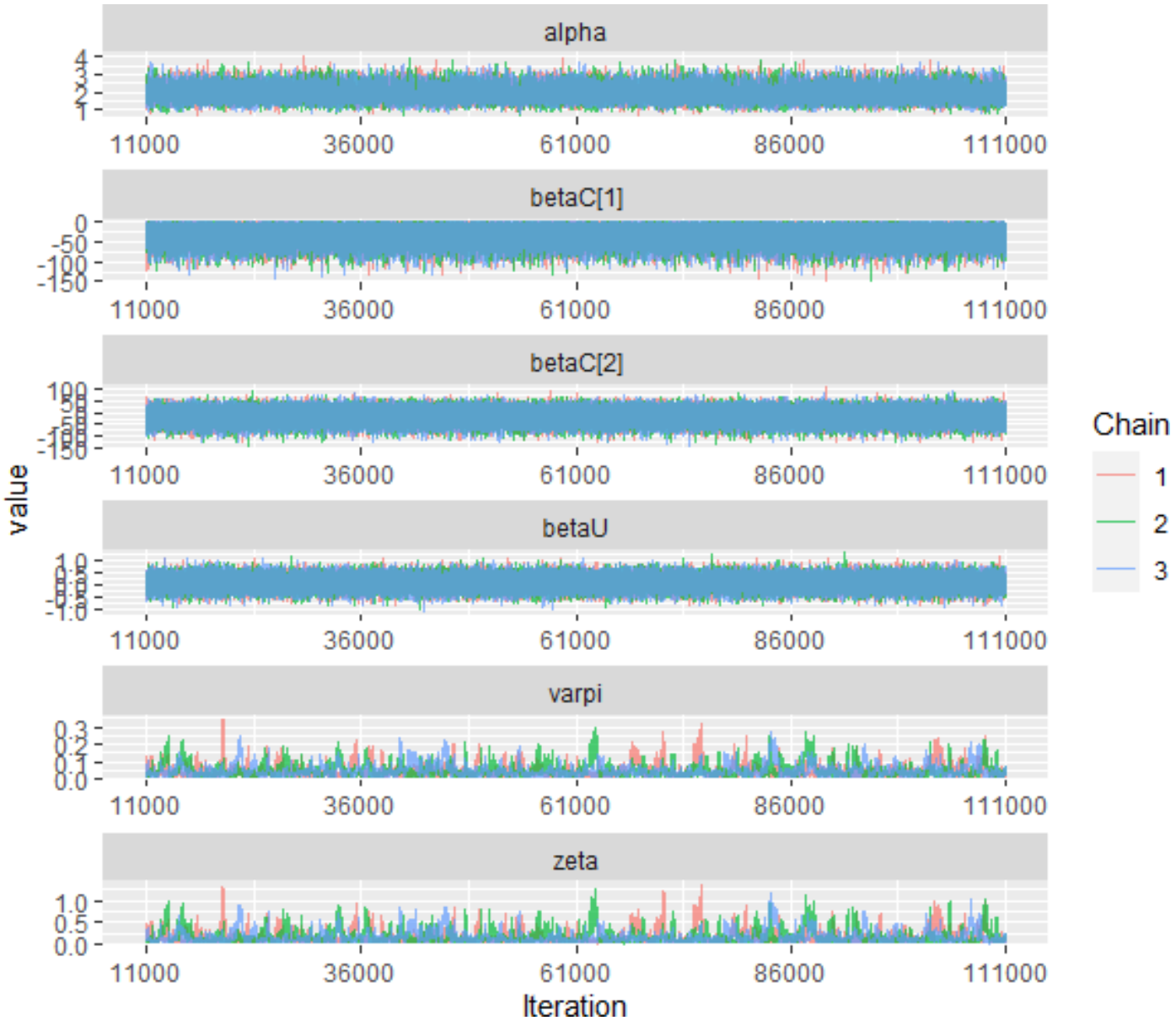

Figure 4: Trace plots of three parallel chains of regression coefficients and distributional parameters for GLLMCM.

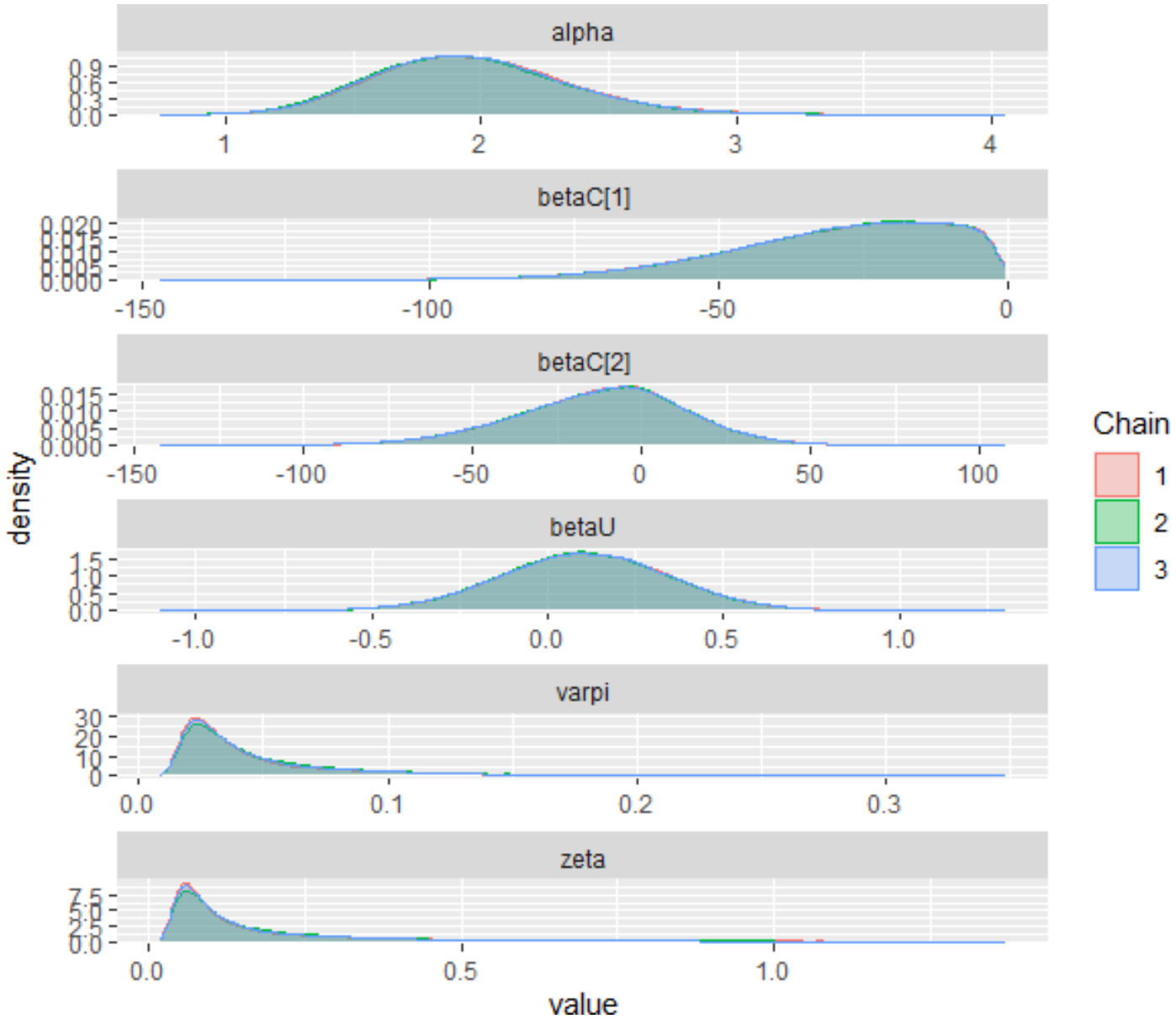

Figure 5: Density plots of three concurrent chains of the GLLMCM model's regression coefficients and distributional parameters.

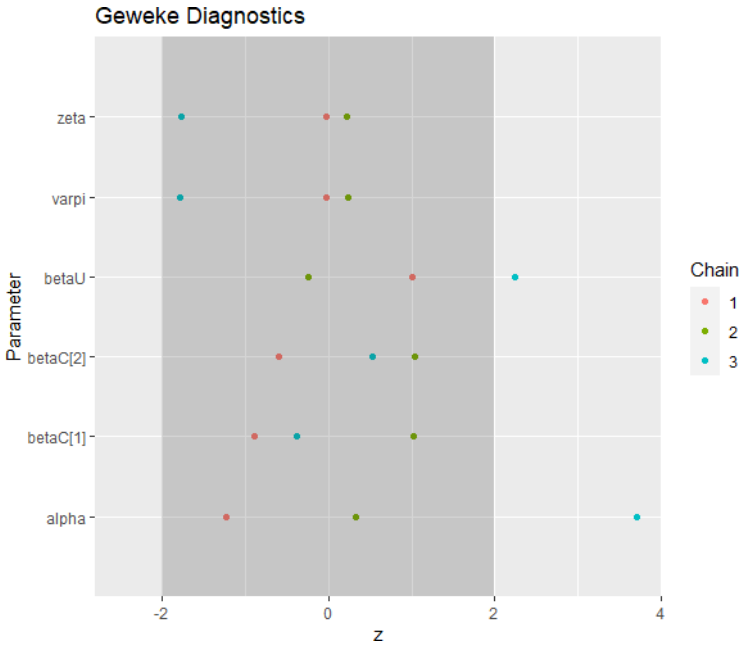

Figure 6: Geweke plots of three concurrent chains for convergence diagnostics.

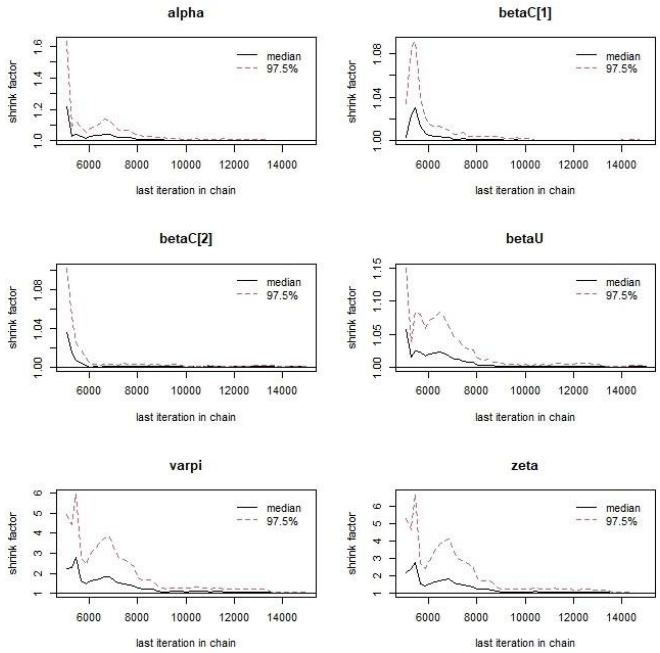

Figure 7: Gelman plots for convergence diagnostics.

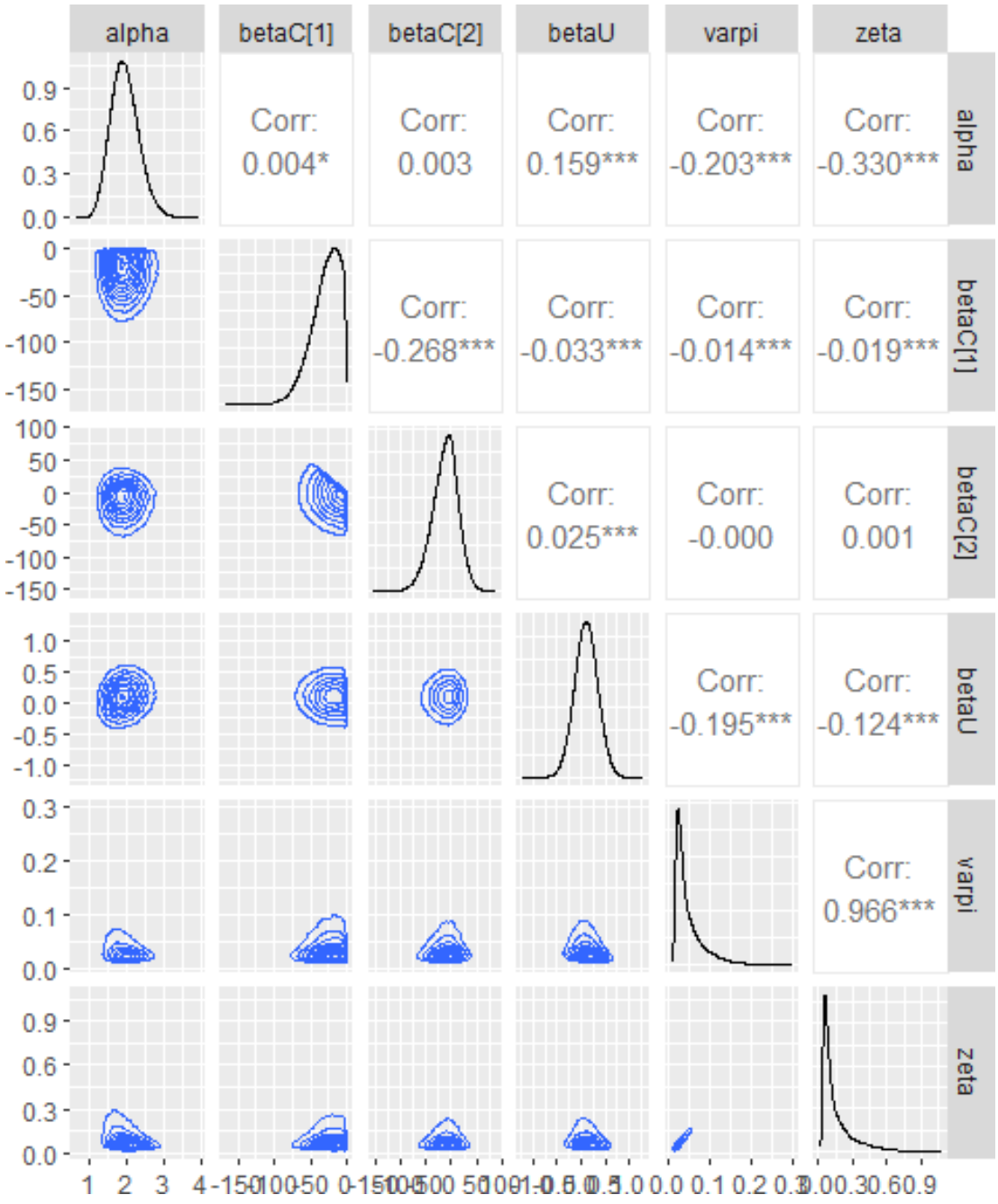

Figure 8: Paired plot displaying contour/scatter plots at the lower quadrant (—), Densities (—) and Cross correlations (upper quadrant)

7 Conclusion.

This paper introduces the generalized Log-Logistic MCM model with the new six-parameter. There are numerous well-known sub-models of this distribution, including the LL, EE, W, and BXII distributions. The illustrations based on actual data were presented through the TTT plot and survival curves. Also, the convergence diagnostic such as Geweke, Gelman, Density, Trace plots and paired plots were well covered. The statistical properties of the new model, including the hazard rate function, quantile function, moments, and mean deviations, are determined. The maximum likelihood method uses both the Bayesian and Frequentist approaches to estimate model parameters. Real-world applications have demonstrated that the new distribution is extremely useful in dealing with lifetime data and outperforms other distributions commonly used for adapting this type of data.

Conflicts of interest

The author declares that, there are no potential conflicts of interest in the publication, public's reaction or pattern of the article

Acknowledgments

I would like to convey my gratefulness to my co-author Magreth Anga Kimaro for her participation to this study.

Data Availability

The real datasets utilized in this work are publicly available online, particularly in R, where (Bone marrow transplant) bmt-cancer data are obtained using the package smcure; and the relevant documentation can be freely downloaded from CRAN webpage https://CRAN.R-project.org/package=smcure.

References

Adnan, W.N.A.W.M and Arasan, J. (2018). Hypothesis testing for the parameters of log-logistic regression model with left-truncated and right-censored survival data, Jurnal Teknologi, 80(3), 2018. ISSN2180\OT1\textendash3722.

Berkson, J., & Gage, R. P. (1952). Survival Curve for Cancer Patients Following Treatment. Journal of the American Statistical Association, 47(259), 501–515. https://doi.org/10.1080/01621459.1952.10501187

Cai, C., Zou, Y., Peng, Y., & Zhang, J. (2012). smcure: an R-package for estimating semiparametric mixture cure models. Computer methods and programs in biomedicine, 108(3), 1255–1260. https://doi.org/10.1016/j.cmpb.2012.08.013

Chen, M.-H., Ibrahim, J. G., & Sinha, D. (1999). A New Bayesian Model for Survival Data with a Surviving Fraction. Journal of the American Statistical Association, 94(447), 909–919. https://doi.org/10.2307/2670006

Cooner, F., Banerjee, S., Carlin, B. P., & Sinha, D. (2007). Flexible Cure Rate Modeling Under Latent Activation Schemes. Journal of the American Statistical Association, 102(478), 560–572. https://doi.org/10.1198/016214507000000112

De Castro, M., Gómez, Y.M. (2020). A Bayesian Cure Rate Model Based on the Power Piecewise Exponential Distribution. Methodology Computing in Applied Probability 22, 677–692. https://doi.org/10.1007/s11009-019-09728-2

Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., . . . Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. Social Psychology, 45*(3)*, 142–152. https://doi.org/10.1027/1864-9335/a000178

Elgarhy, M.; Shakil, M.; and Golam Kibria, B. M. (2017). Exponentiated Weibull-Exponential Distribution with Applications, Applications and Applied Mathematics: An International Journal (AAM), Vol. 12, Iss. 2, Article 5.

Federal Reserve Bank of Minneapolis. Research Department, & Geweke, John. (1991). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments (Number 148) [Reports]. Federal Reserve Bank of Minneapolis. Staff Reports. Federal Reserve Bank of Minneapolis. https://jstor.org/stable/community.28109706

Gallardo, D. I., Gómez, Y. M., & de Castro, M. (2018). A flexible cure rate model based on the polylogarithm distribution. Journal of Statistical Computation and Simulation, 88(11), 2137–2149. https://doi.org/10.1080/00949655.2018.1451850

Grady W. M. (2003). Genetic testing for high-risk colon cancer patients. Gastroenterology, 124(6), 1574–1594. https://doi.org/10.1016/s0016-5085(03)00376-7

Hannan, E. J., & Quinn, B. G. (1979). The Determination of the Order of an autoregression. Journal of the Royal Statistical Society. Series B (Methodological), 41(2), 190–195. http://www.jstor.org/stable/2985032

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrical*, *76*(2), 297–307. https://doi.org/10.2307/2336663

Ibrahim, J. G, Chen, M.H, and Sinha, D. (2001) . Bayesian survival analysis, volume 2. springer. https://doi.org/10.1007/978- 1-4757-3447-8.

Jerald F Lawless. (2011) Statistical models and methods for lifetime data. John Wiley and Sons, ISBN1118031253,9781118031254

John W Boag, (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society. Series B (Methodological), 11(1):15 53. https://www.jstor.org/stable/2983694.

Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. M. (1998). Markov Chain Monte Carlo in Practice: A Roundtable Discussion. The American Statistician, 52(2), 93–100. https://doi.org/10.2307/2685466

Khan, S., Khosa, S. (2016). Generalized log-logistic proportional hazard model with applications in survival analysis. *J Stat Distrib App* **3**, 16. https://doi.org/10.1186/s40488-016-0054-z

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W. E., Kutal, D. H., & Qian, L. (2018). A Non-Mixture Cure Model for Right-Censored Data with Fréchet Distribution. *Stats*, *1*(1), 176-188. https://doi.org/10.3390/stats1010013

Maller, R.A, and Xian Z, (1996). *Survival Analysis with Long-Term Survivors*. Chichester; volume 525. Wiley New York. https://lccn.loc.gov/96026979

Muse AH, Mwalili S, Ngesa O, Almalki SJ, Abd-Elmougod GA. (2021). Bayesian and Classical Inference for the Generalized Log-Logistic Distribution with Applications to Survival Data. Comput Intell Neurosci. Doi: 10.1155/2021/5820435. PMID: 34671390; PMCID: PMC8523281.

Muse, A. H., Mwalili, S., Ngesa, O., Alshanbari, H. M., Khosa, S. K., & Hussam, E. (2022). Bayesian and frequentist approach for the generalized log-logistic accelerated failure time model with applications to larynx-cancer patients. *Alexandria Engineering Journal*, *61*(10), 7953–7978. https://doi.org/10.1016/j.aej.2022.01.033

Muse, A. H., Mwalili, S., Ngesa, O., Chesneau, C., Al-Bossly, A., & El-Morshedy, M. (2022). Bayesian and Frequentist Approaches for a Tractable Parametric General Class of Hazard-Based Regression Models: An Application to Oncology Data. *Mathematics*, *10*(20), 3813. https://doi.org/10.3390/math10203813

Muse AH, Tolba AH, Fayad E, Abu Ali OA, Nagy M, Yusuf M. (2021). Modelling the COVID-19 Mortality Rate with a New Versatile Modification of the Log-Logistic Distribution. Comput Intell Neurosci. Doi: 10.1155/2021/8640794. PMID: 34782836; PMCID: PMC8590594.

Naslina.A.M.N. N, Jayanthi. A, Syahida.Z.H, and Bakri.A.M. (2020). Assessing the goodness of fit of the gompertz model in the presence of right and interval censored data with covariate. Austrian Journal of Statistics, 49(3):57-71.

https://doi.org/10.17713/ajs.v49i3.1085

Njau.B. A, Mwalili.M.S., Orwa.G.O. (2022). Bayesian Estimation of the Colon Cancer Frailty Hazards Functions for Mixture Cure Model.GSJ: Volume 10, ISSN2320-9186

Noor, Farzana, Masood, Saadia, Zaman, Mehwish, Siddiqa, Maryam, Wagan, Raja Asif, Khan, Imran Ullah, Sajid, Ahthasham, (2021). Bayesian Analysis of Inverted Kumaraswamy Mixture Model with Application to Burning Velocity of Chemicals, Mathematical Problems in Engineering, 2021, 5569652, 18 pages, 2021. https://doi.org/10.1155/2021/5569652

Omer, M. E. A. M. E., Bakar, M. R. A., Adam, M. B., & Mustafa, M. S. (2020). Cure Models with Exponentiated Weibull Exponential Distribution for the Analysis of Melanoma Patients. *Mathematics*, *8*(11), 1926. https://doi.org/10.3390/math8111926

Peng, Y., & Taylor, J. M. G. (2017). Residual-Based Model Diagnosis Methods for Mixture Cure Models. *Biometrics*, *73*(2), 495–505. http://www.jstor.org/stable/44695174

Peng, Y., & Yu, B. (2021). Cure Models: Methods, Applications, and Implementation (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429032301

Purvis, A., Rambaut, A., & Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. (1995). Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analyzing comparative data. *CABIOS*, *11*(3), 247–251.

Raftery, A. E., & Lewis, S. M. (1992). [Practical Markov Chain Monte Carlo]: Comment: One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo. *Statistical Science*, *7*(4), 493–497. http://www.jstor.org/stable/2246100

Rubio, F. J., Remontet, L., Jewell, N. P., & Belot, A. (2019). On a general structure for hazard-based regression models: An application to population-based cancer research. *Statistical methods in medical research*, *28*(8), 2404–2417. https://doi.org/10.1177/0962280218782293

Ryosuke Suga, Kiyoshi Yonekawa, Mutsuto Kawahara. (2003) - Estimation of tidal current using Kalman filter finite element method with AIC, Editor(s): K.J. Bathe, Computational Fluid and Solid Mechanics.

Teo BW, Xu H, Koh YY, Li J, Sinha AK, Shuter B, Sethi S, Lee EJ. (2012). Estimating kidney function in a multiethnic Asian population with multiple filtration markers. Am J Kidney. Doi: 10.1053/j.ajkd.05. 008.. PMID: 22721932

Tsodikov AD, Ibrahim JG, Yakovlev AY. (2003). Estimating Cure Rates from Survival Data: An Alternative to Two-Component Mixture Models. J Am Stat Assoc. Doi: 10.1198/0162214503000001007. PMID: 21151838; PMCID: PMC2998771.

Wang, X., Ryan Yue, Y., & Faraway, J.J. (2018). Bayesian Regression Modeling with INLA (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781351165761

Yakovlev, A., Tsodikov, A.D., & Asselain, B. (1996). Stochastic Models of Tumor Latency and Their Biostatistical Applications. Volume 1. World Scientific. ISBN9814501840,9789814501842.