Pak. J. Statist.
2022 Vol. 38(4), 431-450

MIXTURE CURE MODEL FOR RIGHT CENSORED SURVIVAL DATA WITH WEIBULL EXPONENTIATED EXPONENTIAL DISTRIBUTION

Umar Usman[1], Suleiman Shamsuddeen[1], Bello Magaji Arkilla[2] and **Aliyu Yakubu[1,3§]**

[1] Department of Mathematics, Usmanu Dan Fodiyo University Sokoto-Nigeria.
[2] Department of Public Medicine, Usmanu Dan Fodiyo University Sokoto-Nigeria.
[3] Department of Statistics, Ahmadu Bello University, Zaria-Nigeria.
Email: yakubualiyu@abu.edu.ng; aliyuyakubu40@gmail.com
[§] Corresponding author

ABSTRACT

Mixture cure rate model for right censored survival data is considered in the present paper. We consider the case where the lifetime data follow the Weibull exponentiated exponential distribution. The model contains as special case the Weibull, exponentiated exponential and exponential mixture cure rate models. The paper utilizes the maximum likelihood method to estimate the parameters of the mixture cure rate model under the assumption of right censoring. Simulations study was performed in order to ascertain the performances of the *MLE* estimators of the mixture model. We generated random samples of different sizes from the model with some arbitrary values for the parameters for 5%, 10% and 15% cure fraction values. Bias and standard error were used as discrimination criteria. Finally, the application of the proposed methodology was demonstrated considering the Eastern Cooperative Oncology Group (ECOG) phase III clinical trial e1684 and the survival time of colorectal cancer patients from Malaysia. The proposed methodology was found to fits the e1684 data better than its sub-models. Furthermore, when compared with other competing long-term mixture cure models using the survival time of colorectal cancer data, the proposed methodology was also found to fit the data better than the models compared with.

KEY WORDS

Long term survivors, Mixture lifetime model, Likelihood ratio test, Weibull exponentiated exponential distribution, Colorectal cancer.

1. INTRODUCTION

The long-term survivor mixture model assumes that the population of study is a mixture of individuals who experience the event of interest (susceptible individuals) and individuals that will never experience the event of interest (non-susceptible, cured or immune) (Maller and Zhou, 1996). That is, the study population is a mixture of susceptible individuals and non-susceptible (cured or immune) individuals.

Long-term survivor mixture model is applied in different areas which include health, reliability, criminology (see (Kannan et al., 2010) for more details). The long-term survivor mixture model is also known as standard mixture cure model. To determine the presence of immune (cured) in a given data set, Kaplan-Meier plot of the survival function is used. This is done by examining the Kaplan-Meier curve, if it shows a long and stable plateau with heavy censoring at the extreme right of the plot, then it suggest the presence of immune (cured) (Corbiere et al., 2009; Martinez et al., 2013).

Modelling survival lifetime data particularly in medical science research usually requires more sophisticated parametric models (Martinez et al., 2013). To solve this problem, new classes of parametric distributions using the Weibull distribution have been introduced by many authors in the literature. This include Weibull Exponentiated Exponential Salem and Selim, 2014; Usman et al., 2020), exponentiated Weibull (Mudholkar and Srivastava, 1993; Pal, Ali, and Woo, 2006), Weibull−Burr III (Yakubu and Doguwa, 2017), the generalized modified Weibull (Carrasco, Ortega and Cordeiro, 2008), log-beta Weibull (Ortega, Cordeiro and Kattan, 2013) and Weibull Kumaraswamy distribution (Ishaq et al., 2017).

Different parametric approaches have been used by different researchers in modelling the proportion of immune (cured), for instance, (Peng, Dear and Denham, 1998) modelled the proportion using generalized F distribution, (Cancho and Bolfarine, 2001) modelled the proportion using exponentiated Weibull distribution, (Kannan *et al.*, 2010) using exponentiated-exponential distribution, this was extended by (Mazucheli, Coelho-Barros and Achcar, 2013) to include the non-mixture cure model, (Shao and Zhou, 2004) uses Burr XII distribution and later extended by (Coelho-Barros, Achcar and Mazucheli, 2017), (Ramos, Nascimento, and Louzada, 2017) uses the Fréchet distribution and it was extended by (Kutal and Qian, 2018) and (Usman, *et al.*, 2021) uses Nadarajah-Haghighi distribution. Other distributions used to model the proportion of long-term survivors include Negative binomial distribution by (Cancho, Rodrigues and de Castro, 2011), Weibull distribution by (Achcar, Coelho-Barros, and Mazucheli, 2012), generalized modified Weibull distribution by (Martinez *et al.*, 2013) and Yule-Simon distribution by (Gallardo, Gómez and Bolfarine, 2017).

In this article, a mixture cure model for right censored survival data based on Weibull exponentiated exponential distribution was introduced. Statistical and mathematical properties of the model were studied and applications of the model were provided. The rest of the paper was organized as follows: the Weibull exponentiated exponential distribution was discussed in section 2. The Weibull exponentiated exponential mixture cure model and its sub-models were also provided in this section. The problem of estimation of the parameters of the model assuming right censoring was discussed in section 3. In section 4, statistical properties of the model such as quantile function, characteristic function, moment generating function and moments of the model were provided. In section 5, simulations studies and applications of the model were provided and we finally conclude in section 6.

2. MODEL

In this section, a class of model for survival data with long term survivors was proposed using Weibull Exponentiated Exponential distribution.

2.1 Weibull Exponentiated Exponential Distribution

An extension of the Weibull distribution called Weibull exponentiated exponential distribution was developed by (Salem and Selim, 2014; Usman, *et al.*, 2020). Mathematical and statistical properties of the distribution such as linear representation, quantile function, moment generating function, characteristic function among other properties were studied by the authors. It was shown that the pdf of the distribution can take various shapes. It was also shown that, the hazard rate function of the Weibull Exponentiated Exponential Distribution takes various shapes (for instance, the graph is constant when the shape parameters of the model takes the value one (1)), see (Usman *et al.*, 2020) for more details). This makes the distribution more flexible in modelling real life data. The distribution contains the exponentiated Weibull, Weibull, exponentiated exponential and exponential distributions as special cases. (Usman *et al.*, 2020) showed that, the distribution was better fitted by some data sets than the weibull, exponentiated exponential and exponential distributions. The probability density function (*pdf*) for a random variable t of the Weibull exponentiated exponential distribution is given by:

$$f(t) = \frac{ab\theta\phi\exp(-bt)}{1-(1-\exp(-bt))^{a}}\left[-\ell n\left\{1-(1-\exp(-bt))^{a}\right\}\right]^{\theta-1}(1-\exp(-bt))^{a-1}$$
$$\exp\left[-\phi\left(-\ell n\left\{1-(1-\exp(-bt))^{a}\right\}\right)^{\theta}\right] \quad (1)$$

where a, b, ϕ and θ are positive parameters, a and ϕ are the shape parameters, θ and b are the scale parameters. The corresponding survival function, cumulative distribution function (cdf) and hazard function are respectively given by:

$$S_s(t) = \exp\left[-\phi\left(-\ell n\left\{1-(1-\exp(-bt))^{a}\right\}\right)^{\theta}\right] \quad (2)$$

$$F_s(t) = 1-\exp\left[-\phi\left(-\ell n\left\{1-(1-\exp(-bt))^{a}\right\}\right)^{\theta}\right] \quad (3)$$

and

$$h_s(t) = \frac{ab\theta\phi\exp(-bt)(1-\exp(-bt))^{a-1}\left[-\ell n\{1-(1-\exp(-bt))^{a}\}\right]^{\theta-1}}{1-(1-\exp(-bt))^{a}}. \quad (4)$$

2.2 Weibull Exponentiated Exponential Mixture Cure Model

The mixture cure model was first developed by (Boag, 1949) and (Berkson and Gage, 1952) to modelled reallife data that consists of two types of individuals in the study population. The model is also referred to as standard cure model. The mixture cure model is a class of models that effectively modelled survival data with long term survivors. That is, it model effectively survival data that consists of susceptible subjects who experienced the event of interest and non-susceptible subjects that will never experienced the event of interest. It is important to note that, those subjects that will never experienced the event of interest are also termed as long-term survivors. The survival function of the standard mixture cure model takes the form:

$$S(t) = 1 - (1-p)F_s(t) \tag{5}$$

where $0 < p < 1$ represents the proportion of subjects that are "non-susceptible" or "long-term-survivors" regarding the event of interest. $F_s(t)$ is the baseline cumulative distribution function for the susceptible subjects. It is important to note that, the survival function of the entire population tends to the proportion of non-susceptible subjects *p* as *t* tends to infinity.

Let *t* denote the survival time of subject *i* in the study population. From expression 5, the survival function of the Weibull exponentiated exponential mixture cure model is given by:

$$S(t; a, b, \theta, \phi, p) = p + (1-p)\exp\left[-\phi\left(-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right)^\theta\right] \tag{6}$$

while the corresponding pdf and *cdf* are respectively given by:

$$f(t;\ a, b, \theta, \phi, p) = \frac{(1-p)ab\theta\phi\exp(-bt)}{1-\left(1-\exp(-bt)\right)^a}\left[-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right]^{\theta-1}$$

$$\left(1-\exp(-bt)\right)^{a-1}\exp\left[-\phi\left(-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right)^\theta\right] \tag{7}$$

and

$$F(t; a, b, \theta, \phi, p) = (1-p)\left\{1-\exp\left[-\phi\left(-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right)^\theta\right]\right\} \tag{8}$$

where $a > 0$ and $\phi > 0$ are the shape parameters, $\theta > 0$ and $b > 0$ are the scale parameters and $0 < p < 1$ is the proportion of non-susceptible.

2.3 Special Sub-Models

The Weibull Exponentiated Exponential Mixture Cure (*WEEMC*) model contains many mixture cure model as special cases. For instance, the *WEEMC* model reduces to

- $S(t; \theta, \phi, p) = p + (1-p)\exp\left(-\phi t^\theta\right)$ when $a = b = 1$ which is the survival function of the Weibull mixture cure model.
- the Weibull mixture cure model with parameters θ, p and ϕb^θ when $a = 1$.
- Exponentiated exponential mixture cure model proposed by (Kannan *et al.*, 2010)

when $\theta = \phi = 1$.

- Exponential mixture cure model with parameter $b\phi$ and p when $a = \theta = 1$. It also reduces to the exponential mixture cure model with parameters b and p when $a = \theta = \phi = 1$.

3. ESTIMATION PROCEDURE

3.1 Maximum Likelihood Estimation

Let t_i be a right censored survival time for subject i in the population of study, then $t_i = min(T_i, \delta_i)$ where T_i is the failure time of the ith subject and δ_i is a censoring indicator such that $\delta_i = 1$ if the observed lifetime t_i is not censored and $\delta_i = 0$ if the observed lifetime t_i is censored. For $i = 1,2,\cdots,n$, the likelihood for the *WEEMC* is given by:

$$L = \prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}$$

the full likelihood function for the $WEEMC$ model is obtained by considering expressions (6) and (7) as:

$$\begin{aligned} L(\Theta/t,\delta) = \prod_{i=1}^{n} & \left\{\frac{(1-p)ab\theta\phi\exp(-bt)}{1-\left(1-\exp(-bt)\right)^{a}}\right\}^{\delta_i} \left[-\ell n\{1\right. \\ & \left. -\left(1-\exp(-bt)\right)^{a}\}\right]^{\delta_i(\theta-1)} \left(1\right. \\ & \left. -\exp(-bt)\right)^{\delta_i(a-1)} \exp\left[-\delta_i\phi\left(-\ell n\{1\right.\right. \\ & \left.\left. -\left(1-\exp(-bt)\right)^{a}\}\right)^{\theta}\right]\Big[p \\ & + (1-p)\exp\left[-\phi\left(-\ell n\{1-\left(1-\exp(-bt)\right)^{a}\}\right)^{\theta}\right]\Big] \end{aligned} \tag{9}$$

where $\Theta = (a, b, \theta, \phi, p)$, the log-likelihood is obtain by taking the natural logarithm of (9), this yields:

$$\begin{aligned} \ell(\Theta/\mathrm{t},\delta) = {} & z\ell n(a) + z\ell n(b) + z\ell n(\theta) + z\ell n(\phi) + z\ell n(1-p) - b\bar{z} \\ & + a\sum \delta_i \ell n(A_i) - \sum \delta_i \ell n(B_i) \\ & + (\theta-1)\sum \delta_i \ell n\left(-\ell n(B_i)\right) - \phi\sum \delta_i\left(-\ell n(B_i)\right)^{\theta} \\ & + \sum (1-\delta_i) \\ & \ell n\left(p + (1-p)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)\right) \end{aligned} \tag{10}$$

where $\bar{z} = \sum \delta_i t_i, A_i = 1-\exp(-bt_i)$ and $B_i = 1 - A_i^a$. To obtain the maximum likelihood estimates of the parameters $\hat{a}, \hat{b}, \hat{\theta}, \hat{\phi}$ and $\hat{p}$ of a, b, θ, ϕ and p, we differentiate (10) partially with respect to the parameters $a, b, \theta\ \phi$, and p. This gives:

$$\frac{\partial \ell\left(\frac{\Theta}{\mathrm{t}},\delta\right)}{\partial a}=\frac{z}{a}+\sum \delta_i \ell n(A_i)+\sum \frac{\delta_i A_i^a \ell n(A_i)}{B_i}-(\theta-1)\sum \frac{\delta_i A_i^a \ell n(A_i)}{B_i \ell n(B_i)}-\theta\phi\sum \frac{\delta_i A_i^a \ell n(A_i)(B_i)^{\theta-1}}{B_i}-\theta\phi(1-p)\sum \frac{(1-\delta_i)A_i^a \ell n(A_i)(B_i)^{\theta-1}}{B_i[p+(1-p)\exp(-\phi(B_i)^{\theta})]} \quad (11)$$

$$\frac{\partial \ell(\Theta/t,\delta)}{\partial b}=\frac{z}{b}-\bar{z}+a\sum \frac{\delta_i t_i(1-A_i)}{A_i}+a\sum \frac{\delta_i t_i(1-A_i)A_i^{a-1}}{B_i}-a(\theta-1)\sum \frac{\delta_i t_i(1-A_i)A_i^{a-1}}{B_i \ell n(B_i)}+a\theta\phi(1-p)\sum \frac{(1-\delta_i)t_i(1-A_i)A_i^{a-1}\left(-\ell n(B_i)\right)^{\theta-1}}{B_i\left[p+(1-p)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)\right]}-a\theta\phi\sum \frac{\delta_i t_i(1-A_i)A_i^{a-1}\left(-\ell n(B_i)\right)^{\theta-1}}{B_i} \quad (12)$$

$$\frac{\partial \ell(\Theta/t,\delta)}{\partial \theta}=\frac{z}{\theta}+\sum \delta_i \ell n\left(-\ell n(B_i)\right)-\phi\sum \delta_i\left(-\ell n(B_i)\right)^{\theta}\ell n\left(-\ell n(B_i)\right)+\phi(1-p)\sum \frac{(1-\delta_i)\left(-\ell n(B_i)\right)^{\theta}\ell n\left(-\ell n(B_i)\right)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)}{p+(1-p)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)} \quad (13)$$

$$\frac{\partial \ell(\Theta/t,\delta)}{\partial \phi}=\frac{z}{\phi}-\sum \delta_i\left(-\ell n(B_i)\right)^{\theta}-(1-p)\sum \frac{(1-\delta_i)\left(-\ell n(B_i)\right)^{\theta}\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)}{p+(1-p)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)} \quad (14)$$

$$\frac{\partial \ell(\Theta/t,\delta)}{\partial p}=\frac{z}{1-p}+\sum \frac{(1-\delta_i)\left(1-\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)\right)}{p+(1-p)\exp\left(-\phi\left(-\ell n(B_i)\right)^{\theta}\right)} \quad (15)$$

equations (11) to (15) are non-linear. To solve these equations, numerical methods such as Newton-Raphson method is used in finding the *MLE*s. For interval estimation and test of hypotheses on the parameters of the *WEEMC* model, the observed Fisher Information Matrix $J(\Theta)$ is obtain. The observed Fisher Information Matrix $J(\Theta)$ of the *WEEMC* model is a 5×5 matrix defined as:

$$J(\Theta) = -\begin{pmatrix} J_{aa} & J_{ab} & J_{a\theta} & J_{a\phi} & J_{ap} \\ & J_{bb} & J_{b\theta} & J_{b\phi} & J_{bp} \\ & & J_{\theta\theta} & J_{\theta\phi} & J_{\theta p} \\ & & & J_{\phi\phi} & J_{\phi p} \\ & & & & J_{pp} \end{pmatrix} \tag{16}$$

where

$$J_{aa} = \frac{\partial^2 \ell}{\partial a^2}, J_{bb} = \frac{\partial^2 \ell}{\partial b^2}, J_{\theta\theta} = \frac{\partial^2 \ell}{\partial \theta^2}, J_{\phi\phi} = \frac{\partial^2 \ell}{\partial \phi^2}, J_{pp} = \frac{\partial^2 \ell}{\partial p^2}, J_{ab} = \frac{\partial^2 \ell}{\partial a \partial b}, J_{a\theta} = \frac{\partial^2 \ell}{\partial a \partial \theta}, J_{a\phi} = \frac{\partial^2 \ell}{\partial a \partial \phi}, J_{ap} = \frac{\partial^2 \ell}{\partial a \partial p}, J_{b\theta} = \frac{\partial^2 \ell}{\partial b \partial \theta}, J_{b\phi} = \frac{\partial^2 \ell}{\partial b \partial \phi}, J_{bp} = \frac{\partial^2 \ell}{\partial b \partial p}, J_{\theta\phi} = \frac{\partial^2 \ell}{\partial \theta \partial \phi}, J_{\theta p} = \frac{\partial^2 \ell}{\partial \theta \partial p}, J_{\phi p} = \frac{\partial^2 \ell}{\partial \phi \partial p}$$

The asymptotic distribution of $\sqrt{n}(\widehat{\Theta} - \Theta)$ is multivariate normal $N_5\left(0, J(\widehat{\Theta})^{-1}\right)$, where $J(\widehat{\Theta})^{-1}$ is the total observed information matrix computed at $\widehat{\Theta}$. The diagonal elements of $J(\Theta)^{-1}$ are the variances of the corresponding parameters while the off-diagonal elements of $J(\Theta)^{-1}$ are covariances. Hence, the asymptotic $100(1-\varepsilon)\%$ confidence interval for any of the model parameter is $parameter \pm z_{\frac{\varepsilon}{2}}\sqrt{var(parameter)}$ where $z_{\frac{\varepsilon}{2}}$ is the $100(1-\varepsilon)\%$ quantile from the standard normal distribution.

3.2 Likelihood Ratio Test (*LRT*)

In a situation where a model have sub-model(s), the *LRT* is used in testing the superiority of the fitted model over the fit(s) of its sub-model(s) for a given data set. To use the *LRT* in testing the superiority of the fitted model over the fit(s) of its sub-model(s), we construct the *LRT* statistic using the values of the restricted and unrestricted log-likelihood values. In such a situation, the hypothesis is expressed as: $H_0: \Theta = \Theta_0$ against $H_1: \Theta \neq \Theta_0$. The test statistic for the *LRT* for testing H_0 against H_1 is given as: $\Upsilon = 2(l(\widehat{\Theta}) - l(\widehat{\Theta}_0))$, where $\widehat{\Theta}$ and $\widehat{\Theta}_0$ are the log-likelihood under H_0 and H_1. The statistic Υ is asymptotically distributed as χ^2_m, where m is the degree of freedom. The null hypothesis H_0 is rejected if $\Upsilon > \Upsilon_\varepsilon$, where Υ_ε is the upper $100\varepsilon\%$ point of the χ^2_m distribution.

4. STATISTICAL PROPERTIES OF THE MODEL

In this section, the statistical properties of the *WEEMC* model such as quantile functions, median, characteristic function, moment generating function and moments were discussed.

4.1 Quantile Functions and Median

The quantile function is used in simulations. The quantile function of a random variable T with $cdfF(t)$ is defined as $Q(u) = F^{-1}(u)$. To find $F^{-1}(u)$, let $u = F(t)$, then the quantile function of the WEEMC model is obtained as:

$$Q(u) = -1/b\,\ell n\left\{1-\left[1-\exp\left[-\left(-1/\phi\log\left(\frac{1-p-u}{1-p}\right)\right)^{1/\theta}\right]\right]^{\frac{1}{a}}\right\} \quad (17)$$

where u is a random variable that follows the uniform distribution within the interval (0, 1). The first, second and third quantiles of the *WEEMC* model can easily be obtain from the expression in (17). This is done by setting $u = 0.25$, 0.50 and 0.75 for the first, second and third quantiles respectively. For example, the median of the *WEEMC* model is obtained as:

$$Q(u) = -1/b\,\ell n\left\{1-\left[1-\exp\left[-\left(-1/\phi\log\left(\frac{0.5-p}{1-p}\right)\right)^{1/\theta}\right]\right]^{\frac{1}{a}}\right\}.$$

4.2 Characteristic Function

Let T be a random variable that follows the $WEEM$ model. Then, the characteristic function of the random variable $T{\sim}WEEMC$ is defined as:

$$\begin{aligned}\vartheta_X(t) &= \int_0^\infty e^{(itx)}f(t)dt \\ &= \int_0^\infty e^{(itx)}\frac{ab\theta\phi(1-p)\exp(-bt)}{1-\left(1-\exp(-bt)\right)^a}\left[-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right]^{\theta-1} \\ &\quad \left(1-\exp(-bt)\right)^{a-1}\exp\left[-\phi\left(-\ell n\left\{1-\left(1-\exp(-bt)\right)^a\right\}\right)^\theta\right]dt \\ &= (1-p)\sum_{s=0}^{\infty}\frac{\left(\frac{it}{b}+s-1\right)!}{\left(\frac{it}{b}-1\right)!\,s!}\int_0^\infty e^{-u}\left[1-\exp\left(-\left(\frac{u}{\phi}\right)^{1/\theta}\right)\right]^{\frac{s}{a}}du \end{aligned} \quad (18)$$

and after a little simplification, we arrived at

$$\varphi_x(t) = (1-p)\sum_{s,v,w=0}^{\infty}\frac{(-1)^{v+w}\left(\frac{s}{a}\right)!\,v^w}{\left(\frac{s}{a}-v\right)!\,s!\,v!\,w!}\left(\frac{it}{b}\right)_s\int_0^\infty\left(\frac{u}{\phi}\right)^{-\frac{w}{\theta}}\exp(-u)du \quad (19)$$

$$= (1-p)\sum_{s,v,w=0}^{\infty}\frac{(-1)^{v+w}\left(\frac{s}{a}\right)!\,v^w\phi^{-\frac{w}{\theta}}}{\left(\frac{s}{a}-v\right)!\,s!\,v!\,w!}\Gamma\left(\frac{w}{\theta}+1\right)\left(\frac{it}{b}\right)_s \quad (20)$$

where

$$\left(\frac{it}{b}\right)_s = \frac{it}{b}\left(\frac{it}{b}+1\right)\left(\frac{it}{b}+2\right)\cdots\left(\frac{it}{b}+s-1\right).$$

4.3 Moment Generating Function

Assume T be a random variable from the $WEEM$ model with survival function and pdf given by equations (6) and (7) respectively. The moment generating function of the $WEEM$ model is obtained as follows:

$$E(e^{tx}) = \int_0^{\infty} e^{tx} \frac{ab\theta\phi(1-p)\exp(-bt)}{1-\left(1-\exp(-bt)\right)^{a}} \Big(1 - \exp(-bt)\Big)^{a-1} \left[-\ell n\left\{1-\left(1-\exp(-bt)\right)^{a}\right\}\right]^{\theta-1} \exp\left[-\phi\left(-\ell n\left\{1-\left(1-\exp(-bt)\right)^{a}\right\}\right)^{\theta}\right] dt \quad (21)$$

$$= (1-p)\int_0^{\infty} \left\{1-\left[1-\exp\left(-\left(\frac{u}{\phi}\right)^{1/\theta}\right)\right]^{1/a}\right\}^{-t/b} e^{-u} du \quad (22)$$

Recall that the series expansion of $(1-t)^{-n} = \sum_{i=0}^{\infty} \frac{(n+i-1)!}{(n-1)!i!} x^i$. Hence, applying to the first term in (22) gives:

$$E(e^{tx}) = (1-p)\sum_{i=0}^{\infty} \frac{(t/b+i-1)!}{(t/b-1)!\,i!} t^i \int_0^{\infty} e^{-u}\left[1-\exp\left(-\left(\frac{u}{\phi}\right)^{1/\theta}\right)\right]^{i/a} du$$

the limiting values of $\left[1-\exp\left(-\left(\frac{u}{\phi}\right)^{1/\theta}\right)\right]^{i/a}$ is between zero and one, as $u \Rightarrow 0$ and as $u \Rightarrow \infty$, then applying binomial and power series expansions to the last term, the *mgf* of the model is obtained as:

$$M_T(x) = (1-p)\sum_{i,j,k=0}^{\infty} \frac{(-1)^{j+k}(i/a)!\,j^k\phi^{-k/\theta}}{(i/a-j)!\,i!\,j!\,k!}(t/b)_i\Gamma(k/\theta+1) \quad (23)$$

where $(t/b)_i = t/b\,(t/b+1)(t/b+2)\cdots(t/b+i-1)$.

4.4 Moments

The rth moment about the origin can be obtained from the moment generating function given in equation (23). This is done by finding the rth derivative with respect to t and substituting $t=0$. Hence, the first and second moments about the origin are obtained as:

$$E(T) = \frac{d}{dt}\left[(1-p)\sum_{i,j,k=0}^{\infty} \frac{(-1)^{j+k}(i/a)!\,j^k\phi^{-k/\theta}}{(i/a-j)!\,i!\,j!\,k!}(t/b)_i\Gamma(k/\theta+1)\right]\Bigg|_{t=0}$$

$$= (1-p)\sum_{i,j,k=0}^{\infty} \frac{(-1)^{j+k}(i/a)!\,j^k\phi^{-k/\theta}}{(i/a-j)!\,i!\,j!\,k!}\Gamma(k/\theta+1)\frac{d}{dt}[(t/b)_i]\Big|_{t=0} \quad (24)$$

The last term in expression (24) can easily be expressed as: $\frac{d}{dt}[(\,t\,/\,b\,)_i]\Big|_{t=0} = \frac{(i-1)!}{b}$. Hence, the first moment is obtained as:

$$E(T) = (1-p)\sum_{i,j,k=0}^{\infty} \frac{(-1)^{j+k}(\,i\,/\,a\,)!\,j^k\phi^{-\,k\,/\,\theta}}{(\,i\,/\,a\,-j)!\,j!\,k!}\frac{\Gamma(\,k\,/\,\theta\,+1)}{ib} \tag{25}$$

Following the same procedure, the second moment of the *WEEMC* model is obtained as:

$$E(T^2) = 2(1-p)\sum_{i,j,k=0}^{\infty} \frac{(-1)^{j+k}(\,i\,/\,a\,)!\,j^k\phi^{-\,k\,/\,\theta}}{(\,i\,/\,a\,-j)!\,j!\,k!}\frac{\big(\psi(i)-\psi(1)\big)}{ib^2} \tag{26}$$

where $\psi(.)$ is digamma function, $\frac{d^2}{dt^2}[(\,t/b\,)_i]\Big|_{t=0} = \frac{2(i-1)!}{b^2}\psi(i)-\psi(1)$ since $\sum_{m=1}^{n}\frac{1}{m} = \psi(n+1)-\psi(1)$. The variance of the $WEEMC$ model is obtain using equations (25) and (26) using the relation $Var(x) = E(X^2) - \big(E(X)\big)^2$.

5. SIMULATION STUDY AND APPLICATIONS

5.1 Simulation Study

In this section, we perform simulation study so as to examine the performance of the Maximum Likelihood Estimation method. Inverse transform method was used in generating the right censored survival time from the *WEEMC* model. The following algorithm was used in generating random sample of size *n.*

Table 1
Maximum Likelihood Estimates, Bias and Standard Error (SE) Estimates

n	Parameters	a=2.5; b=3.0; θ=2.0; ϕ=4.5; p=0.05			a=2.5; b=3.0; θ=2.0; ϕ=4.5; p=0.10			a=2.5; b=3.0; θ=2.0; ϕ=4.5; p=0.15		
		Estimates	Bias	SE	Estimates	Bias	SE	Estimates	Bias	SE
30	a	3.8701	1.3701	3.5011	3.8952	1.3952	3.6784	3.9909	1.4909	4.6632
	b	3.1142	0.5142	1.5733	3.1294	0.5294	1.7627	3.2133	0.5973	2.3191
	θ	2.3194	0.3494	1.4343	2.4649	0.4849	1.5177	2.5143	0.5143	1.7417
	ϕ	4.1369	2.6369	4.6710	4.2085	2.5085	5.6040	4.1689	2.6689	6.6419
	p	0.0456	-0.0300	0.0795	0.0937	-0.0563	0.0614	0.1474	-0.0026	0.0644
50	a	3.3543	0.8543	2.9567	3.3943	0.8943	2.7766	3.6767	1.1767	3.2607
	b	2.9293	-0.4707	1.3207	2.9776	-0.4724	1.6029	3.1484	0.4784	1.6034
	θ	2.3288	0.3288	1.1499	2.4744	0.4744	1.2929	2.4056	0.4056	1.2552
	ϕ	4.3174	2.2817	4.1859	3.9746	2.4746	4.7722	3.9608	2.4608	4.6613
	p	0.0464	-0.0036	0.0411	0.1011	0.0011	0.0419	0.1483	-0.0017	0.0492
75	a	2.9452	0.4452	2.1421	3.0950	0.5950	2.1161	3.2930	0.7930	2.1491
	b	2.7062	-0.4389	1.0779	2.8492	-0.4428	1.1468	2.8654	-0.4436	1.2050
	θ	2.3763	0.3063	0.9995	2.4083	0.4083	1.0825	2.3879	0.3879	1.0119
	ϕ	4.3865	2.1865	4.0071	4.1279	2.2679	4.4189	4.3797	2.2879	4.2926
	p	0.0491	-0.0009	0.0269	0.0998	-0.0008	0.0340	0.1492	-0.0008	0.0407
100	a	2.8642	0.3642	1.9958	2.8799	0.3799	1.7678	3.0452	0.5452	2.1498
	b	2.6232	-0.3768	1.0701	2.6153	-0.3847	0.9905	2.6509	-0.3941	1.1520
	θ	2.2935	0.2935	0.8922	2.3932	0.3932	0.9608	2.3484	0.3484	1.0102
	ϕ	4.4429	1.9429	4.1556	4.5442	2.0442	3.8166	4.6991	2.1991	3.7806
	p	0.0497	-0.0003	0.0220	0.1003	0.0005	0.0297	0.1492	-0.0008	0.0353
150	a	2.6661	0.1661	1.4485	2.7321	0.2321	1.4338	2.9047	0.4047	1.8474
	b	2.5784	-0.3216	0.8592	2.5397	-0.3303	0.7982	2.6390	-0.3610	0.9105
	θ	2.2682	0.2682	0.7245	2.2884	0.2884	0.7700	2.2817	0.2817	0.7828
	ϕ	4.2826	1.9268	3.0450	4.5756	2.0756	3.3661	4.6791	2.1791	3.4982
	p	0.0503	0.0003	0.0177	0.0996	-0.0004	0.0243	0.1491	-0.0008	0.0289
200	a	2.6143	0.1143	1.2491	2.6940	0.1940	1.2925	2.7779	0.2779	1.7803
	b	2.5481	-0.3159	0.7397	2.5326	-0.3274	0.7829	2.5074	-0.3426	0.8772
	θ	2.2000	0.2000	0.6298	2.2238	0.2238	0.7080	2.2082	0.2082	0.7053
	ϕ	4.3612	1.8612	2.9036	4.5044	2.0045	3.0319	4.8582	2.1582	3.6891
	p	0.0501	0.0001	0.0154	0.1000	0.0003	0.0211	0.1493	-0.0007	0.0251
250	a	2.6333	0.1333	1.3281	2.6336	0.1336	1.1452	2.7029	0.2029	1.3368
	b	2.5141	-0.2859	0.7922	2.5345	-0.2955	0.7000	2.5399	-0.3061	0.7037
	θ	2.1212	0.1212	0.6705	2.1782	0.1782	0.5787	2.1652	0.1652	0.6447
	ϕ	4.4953	1.6953	2.7771	4.4836	1.9836	2.8387	4.6657	2.0657	3.1275
	p	0.0501	0.0001	0.0138	0.0995	-0.0003	0.0189	0.1494	-0.0006	0.0225
300	a	2.5945	0.0945	0.9994	2.5824	0.0824	1.1434	2.6703	0.1703	1.1536
	b	2.4502	-0.2498	0.5797	2.4988	-0.2501	0.6366	2.5356	-0.2764	0.6546
	θ	2.1137	0.1137	0.5163	2.1649	0.1649	0.5604	2.1489	0.1489	0.5603
	ϕ	4.6769	1.1769	2.5853	4.5026	1.5026	2.5059	4.6369	1.8369	2.6290
	p	0.0500	0.0000	0.0126	0.0997	-0.0002	0.0173	0.1496	-0.0004	0.0206

Table 2
Maximum Likelihood Estimates, Bias and Standard Error (SE) Estimates

n	Parameters	a=1.75; b=2.0;θ=1.5; ϕ=2.0; p=0.05			a=1.75; b=2.0; θ=1.5; ϕ=2.0; p=0.10			a=1.75; b=2.0; θ=1.5; ϕ=2.0; p=0.15		
		Estimates	Bias	SE	Estimates	Bias	SE	Estimates	Bias	SE
30	a	3.4800	1.7300	2.7604	3.2111	1.4611	2.9628	3.1466	1.3966	2.4781
	b	2.9236	0.9236	1.4428	2.5828	0.7828	1.7558	2.7366	0.7366	1.8989
	θ	1.5532	0.2532	0.9542	1.8219	0.3219	1.1690	1.9325	0.4325	1.2031
	ϕ	2.6111	0.6111	2.5352	3.2460	2.2461	3.8857	2.8326	0.8326	4.3133
	p	0.0464	-0.1355	0.1138	0.1035	-0.0258	0.0761	0.1424	-0.0076	0.0730
50	a	2.8035	1.0535	2.2964	2.6190	0.8690	1.9916	2.7799	1.0299	2.1816
	b	2.8138	0.8138	1.5670	2.2580	0.6258	1.2696	2.6577	0.6577	1.7777
	θ	1.6480	0.1849	0.9752	1.7909	0.2909	0.9088	1.8277	0.3277	1.0467
	ϕ	2.4360	0.4360	2.9900	3.6488	2.1488	3.2156	2.6968	0.6968	3.4058
	p	0.0292	-0.0208	0.0514	0.0977	-0.0023	0.0459	0.1473	-0.0027	0.0533
75	a	2.3809	0.6309	1.7275	2.2852	0.5352	1.4941	2.3993	0.6493	1.7609
	b	2.6902	0.6902	1.4288	2.0104	0.5104	1.0761	2.5952	0.5952	1.5587
	θ	1.6691	0.1691	0.7721	1.8539	0.2839	0.8429	1.7858	0.2858	0.9437
	ϕ	2.2151	0.2151	2.1268	3.7538	1.8738	2.8516	2.4695	0.4695	2.6107
	p	0.0475	-0.0035	0.0313	0.1003	0.0003	0.0367	0.1479	-0.0021	0.0432
100	a	2.2405	0.4905	1.5085	2.1514	0.4014	1.3750	2.2542	0.5042	1.6078
	b	2.6638	0.6638	1.3180	1.9011	-0.3989	1.0667	2.5906	0.5906	1.5005
	θ	1.6452	0.1452	0.7141	1.7671	0.2671	0.7633	1.7241	0.2241	0.8077
	ϕ	2.1327	0.1327	1.9325	3.8306	1.8306	2.0376	2.2325	0.2325	2.2003
	p	0.0483	-0.0017	0.0241	0.0997	-0.0003	0.0318	0.1493	-0.0017	0.0377
150	a	2.0692	0.3192	1.1777	2.0876	0.3376	1.1782	2.0847	0.3347	1.2967
	b	2.5533	0.5533	1.1134	1.7802	-0.3198	0.8293	2.4817	0.4817	1.2676
	θ	1.6323	0.1323	0.5682	1.6735	0.1735	0.6734	1.6673	0.1673	0.6313
	ϕ	1.9970	-0.1303	1.7781	3.9531	1.5731	1.8271	2.2148	0.2148	1.8316
	p	0.0495	-0.0005	0.0193	0.1000	0.0003	0.0261	0.1490	-0.0010	0.0307
200	a	1.9476	0.1976	1.0063	2.0300	0.2800	0.9761	1.9981	0.2481	1.0758
	b	2.4866	0.4866	1.0490	1.7292	-0.2708	0.7885	2.4030	0.4030	1.1802
	θ	1.6176	0.1176	0.5249	1.6114	0.1114	0.5438	1.6485	0.1485	0.5506
	ϕ	1.9811	-0.1298	1.5021	3.893	1.4893	1.5173	2.1551	0.1551	1.6561
	p	0.0497	-0.0003	0.0167	0.0996	-0.0004	0.0226	0.1492	-0.0008	0.0267
250	a	1.9483	0.1973	0.8837	1.9697	0.2197	0.9123	1.9369	0.1869	0.9619
	b	2.4468	0.4468	0.9017	1.7716	-0.2284	0.6919	2.3988	0.3988	1.1071
	θ	1.5940	0.0940	0.4555	1.5834	0.0834	0.5097	1.6283	0.1283	0.4981
	ϕ	1.9411	-0.1258	1.3010	3.6242	0.6242	1.4522	2.0696	0.0696	1.6845
	p	0.0498	-0.0002	0.0150	0.0995	-0.0005	0.0202	0.1492	-0.0008	0.0239
300	a	1.8797	0.1297	0.8348	1.9188	0.1688	0.9206	1.9027	0.1527	0.8484
	b	2.4834	0.4834	0.9557	1.7864	-0.2136	0.8188	2.3258	0.3258	0.9494
	θ	1.5973	0.0937	0.4287	1.5933	0.0933	0.5172	1.6212	0.1212	0.4754
	ϕ	1.8146	-0.1054	1.1996	3.4246	0.4246	1.4189	2.1150s	0.0115	1.4890
	p	0.0501	0.0001	0.0137	0.1001	0.0001	0.0186	0.1498	-0.0002	0.0219

i. Generate a random sample u_i of size n from $U(0,1)$.

ii. For a cure fraction p, calculate the random survival time using the equation $t_i = -1/b\ \ell n\left\{1-\left[1-\exp\left[-\left(-1/\phi\ \ell n\left(\frac{1-p-u}{1-p}\right)\right)^{1/\theta}\right]\right]^{\frac{1}{a}}\right\}$ when $u_i \leq 1-p$ else t_i is infinity.

iii. Generate a random sample of the censoring times $c_1, c_2, \cdots, c_n$ from the *WEE* distribution.

iv. Calculate the minimum between the survival time and the censoring time. That is, $z_i = min(t_i, c_i), \delta_i = I(t_i \leq c_i), i = 1,2,\cdots,n$

v. The observed right censored survival data is $Z = (z_i, \delta_i)$, for $i = 1,2,\cdots,n$.

This algorithm was used in generating right censored survival times. Samples of different sizes with different proportion of cure fraction values were generated for some arbitrary parameter values. On the other hand, the censoring variable was assumed to follow the Weibull exponentiated exponential distribution.

In assessing the performance of the estimates, bias and standard error were used. Bias is defined as the difference between the expected value of an estimator and the true parameter value. That is, $Bias = E(\widehat{\Theta} - \Theta)$. While standard error is defined as the ratio of standard deviation to the square root of sample size. Mathematically, $S.E = \frac{\sigma}{\sqrt{n}}$. The standard error is used in measuring the accuracy of an estimator.

All simulation results were based on 1000 replications for the different sample sizes considered for each parameter setting. The simulation results are shown in table 1 and 2. The tables shows the the mean estimates of the parameters together with bias and standard error (SE). Careful examination of the estimates in these tables reveals that, the average estimate are close to the respective true parameter values especially when the sample sizes gets larger. The bias and standard error of all examined parameters are small and gets smaller as sample size increases for all the different settings. Hence, the simulation results reveals that the proposed method has a good performance overall.

5.2 Applications

Two data sets were used in illustrating the methodology of the mixture model studied in this article. The first data is the melanoma data from the Eastern Cooperative Oncology Group (ECOG) phase III clinical trial e1684 available in the smcure package in R software. The e1684 trial consists of patients who were treated with wide local excision and complete regional lymph node dissection and then randomized to adjuvant high dose IFN (20 MU/m IV 5 days per week for 4 weeks, followed by 10 MU/m 3 days per week SC for 48 weeks) or observation group. Hence, analysis of treatment effects versus observation group was based on data from 285 patients randomized to IFN or observation group in the trial e1684.

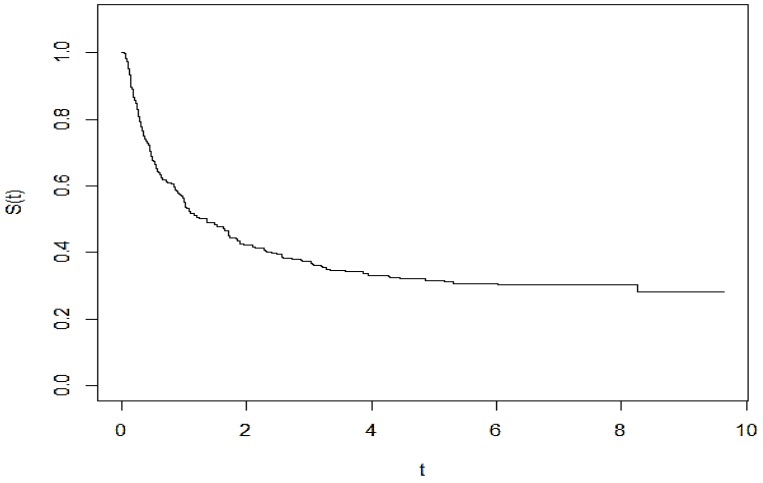

Figure 1: Kaplan-Meier Relapse-Free Survival Curve

The Kaplan-Meier survival curve of this data is given in figure 1. From the Kaplan-Meier survival curve, it can easily be seen after about 8-years follow-up, some patients have not experienced any recurrence after treatments. That is, the curve level off at a value greater than 0.2 but less than 0.3. Hence, we say that, there exist cured patients in this data. The data was fitted to the *WEEMC* model with its sub-models: *WMC*, *EEMC* and *EMC* models. The maximum likelihood estimates of the parameters, their standard errors (SE) and 95% confidence interval (*CI*) of the fitted models are given in Table 3. *AIC*, *BIC* and *CAIC* values of the fitted models are also given in this table. From the statistics in this table, we can conclude that the *WEEMC* model is more efficient compared to *WMC*, *EEMC* and *EMC* models, since *AIC*, *BIC* and *CAIC* values of the *WEEMC* model are smaller than that of the *WMC*, *EEMC* and *EMC* models.

Table 3
Maximum Likelihood Estimate of e1684 Data

Model	Parameter	Estimate	SE	95% CI	AIC	BIC	CAIC
WEEMC	a	4.8789	0.0080	(4.8632,4.8947)	757.7328	775.9952	747.9479
	b	3.1414	0.0034	(3.1347,3.148)			
	θ	0.4021	0.0311	(0.3412,0.4631)			
	ϕ	0.7821	0.0907	(0.6044,0.9598)			
	p	0.2443	0.0401	(0.1658,0.3229)			
WMC	θ	0.9608	0.0524	(0.8581,1.0635)	812.0700	823.0275	806.1554
	ϕ	0.8063	0.0666	(0.6757,0.9368)			
	p	0.4235	0.0253	(0.3739,0.4731)			
EEMC	a	0.9321	0.0897	(0.7563,1.108)	781.2920	792.2495	775.3774
	b	0.8377	0.0903	(0.6607,1.0147)			
	p	0.299	0.0277	(0.2447,0.3533)			
EMC	b	0.8863	0.0676	(0.7538,1.0189)	779.8358	787.1408	775.8784
	p	0.2991	0.0276	(0.245,0.3532)			

The asymptotic variance covariance matrix for the fitted *WEEMC* model based on the e1684 data is given as:

$$\begin{pmatrix} & a & b & \theta & \phi & p \\ a & 0.0000648 & -0.0000005 & -0.0000041 & 0.0000018 & -0.0000016 \\ b & -0.0000005 & 0.0000116 & 0.0000006 & -0.0000021 & 0.0000001 \\ \theta & -0.0000041 & 0.0000006 & 0.0009658 & 0.0008022 & 0.0006196 \\ \phi & 0.0000018 & -0.0000021 & 0.0008022 & 0.0082223 & 0.0020902 \\ p & -0.0000016 & 0.0000001 & 0.0006196 & 0.0020902 & 0.0016078 \end{pmatrix}$$

Likelihood ratio test was carried out to test for the significance of the proposed model with its related sub-models at 5% significance level. The computed test statistic for the comparison between *WEEMC* model with *EMC* model, *WEEMC* model with *WMC* model and *WEEMC* model with *EEMC* model are respectively evaluated as:

$$\tau_E = 2(-373.866 - (-387.918)) = 28.1030$$
$$\tau_W = 2(-373.866 - (-403.035)) = 58.3372$$

and

$$\tau_{EE} = 2(-373.866 - (-387.646)) = 27.5592$$

Table 4
Likelihood ratio Test Statistic for e1684 Data

Model Compared with	Hypothesis	τ	p value
EMC Model	$H_0: a = \theta = \phi = 1$ Vs $H_1: H_0$ is not true	28.1030	0.000003
WMC Model	$H_0: a = b = 1$ Vs $H_1: H_0$ is not true	58.3372	0.000000
EEMC Model	$H_0: \theta = \phi = 1$ Vs $H_1: H_0$ is not true	27.5592	0.000001

Table 4 gives the summary statistics for the likelihood ratio test between the *WEEMC* model and its sub-models. We can easily observe from this table that the p-values of *EMC*, *WMC* and *EEMC* models are all significant. Hence, the likelihood ratio test shows that the *WEEMC* model is more efficient than its sub-models.

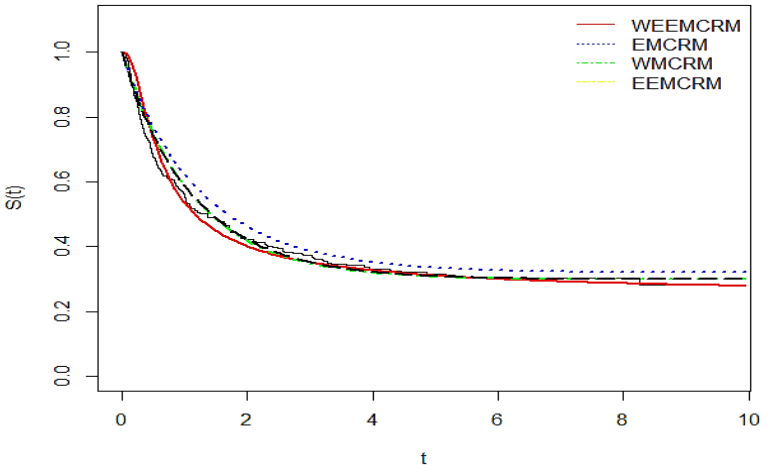

Figure 2: Fitted Survival Curves of the e1684 Data

We also compared the Kaplan-Meier survival curve with the parametric curves. Figure 2 is the Kaplan-Meier survival curve overlaid with the *WEEMC*, *WMC*, *EEMC* and *EMC* survival curves. The curve of the *WEEMC* model is closer to the Kaplan-Meier survival curve in comparison with the survival curves of its sub-models.

The second data consist of the time from the day of diagnosis to death of one hundred and seven patients suffering from Colorectal Cancer in Malaysia. The patients were followed for a long period of time between January, 2001 and December, 2010. The patients were either treated by surgery only or by chemotherapy, surgery and radiotherapy in the University of Malaya Medical Center (UMMC). Patients treated with surgery only are termed as treatment group 1 while those treated with chemotherapy, surgery and radiotherapy are termed treatment group 2. There were a total of Eighty (80) patients in treatment group 1 and twenty seven (27) in treatment group 2. The survival time is the time from the date of commencement of treatment to death, loss to follow-up or end of the study. The Kaplan-Meier survival curve for the data is given in Figure 3.

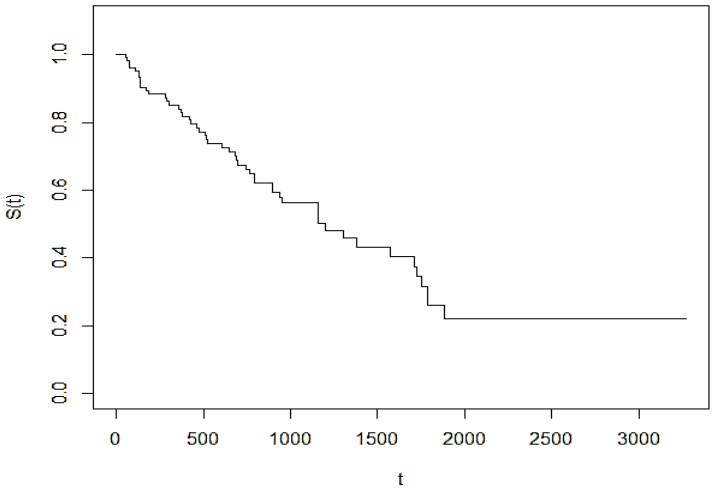

Figure 3: Fitted survival Curves of the Malaysian Colorectal Cancer Data

The Kaplan-Meier curve for the colorectal cancer patients is given in figure 3. From this graph, we observed that after about 1850 days, the curves level off at a value close to 0.2. This suggests that there is presence of long-term survivors in the data sets.

Table 5
Maximum Likelihood Estimates of the Colorectal Cancer Data

Model	Parameter	Estimate	SE	95% CI	AIC	BIC	CAIC
WEEMCR Model	a	1.0333	0.5158	(0.0224,2.0442)	873.0896	886.4537	863.3866
	b	0.0011	0.0001	(0.001,0.0013)			
	θ	1.3324	0.4436	(0.463,2.2018)			
	ϕ	0.8726	0.0559	(0.7629,0.9822)			
	p	0.2442	0.0805	(0.0864,0.402)			
GGMCR Model	β	0.0009	0.0005	(-0.0007,0.0009)	893.5502	904.2415	885.7463
	θ	0.0001	0.0004	(0.0000,0.0018)			
	ϕ	2.0900	1.0756	(-0.018,4.1984)			
	p	0.2200	0.1305	(-0.0357,0.4757)			
MWMCR Model	β	0.1113	0.0465	(-0.0667,0.1154)	900.273	910.9643	892.4691
	θ	0.0244	0.0677	(-0.0215,0.2440)			
	ϕ	0.0018	0.0002	(0.0013,0.0022)			
	p	0.3538	0.0733	(0.2101,0.4974)			
GMWMCR Model	β	0.0684	0.0313	(1.1000,1.2226)	902.6352	915.9993	892.9322
	θ	1.1613	0.2042	(-0.3317,0.4686)			
	ϕ	0.2933	0.0783	(0.1399,0.4467)			
	α	0.0004	0.0002	(0.0001,0.0008)			
	p	0.2646	0.1036	(0.0616,0.4676)			

Table 5 gives the summary statistics of the Malaysian colorectal cancer data. The data was fitted to the *WEEMC*, generalized Gompertz mixture cure (*GGMC*), modified Weibull mixture cure (*MWMC*) and generalized modified Weibull mixture cure (*GMWMC*) models. The maximum likelihood estimates of the parameters, their standard errors (SE) and 95% confidence interval (*CI*) of the fitted models together with the *AIC*, *BIC* and *CAIC* values of the fitted models are given. From the table, the proposed *WEEMC* model is seen to have the least *AIC, BIC* and *CAIC* values. Hence, we conclude that the *WEEMC* model is more efficient compared to *GGMC, MWMC* and *GMWMC* models based on the colorectal cancer data.

6. CONCLUSION

In this article, a new cure model for right censored survival data was proposed based on *WEE* distribution. The model contains the exponential, Weibull and exponentiated exponential mixture cure models as special case. Simulation study was developed under various parameter settings. The simulation study showed that the maximum likelihood method of estimation is a suitable method in estimating the parameters of the proposed mixture model. Furthermore, we compared the performance of the models with the sub-models of the *WEEMC* model using real data sets. It was found that the proposed model is the best model in comparison to its sub-models for modelling the ECOG phase III clinical trial e1684. Finally, we fitted the survival times of one hundred and seven (107) colorectal cancer data and compared its performance with that of *GGMC, MWMC* and *GMWMC* models. It was found that the proposed WEE long-term survivor model is more efficient than the *GGMC, MWMC* and *GMWMC* models.

REFERENCES

1. Achcar, J.A., Coelho-Barros, E.A. and Mazucheli, J. (2012). Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications*, 51(1), 1-9.
2. Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.
3. Alzaghal, A., Famoye, F. and Lee, C. (2013). Exponentiated T-x family of distributions with some applications. *International Journal of Statistics and Probability*, 2(3), 31-49.
4. Berkson, J. and Gage, R.P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47(259), 501-515.
5. Boag, J.W. (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.
6. Cancho, V.G. and Bolfarine, H. (2001). Modeling the presence of immunes by using the exponentiated Weibull model. *Journal of Applied Statistics*, 28(6), 659-671.
7. Cancho, V.G., Rodrigues, J. and de Castro, M. (2011). A flexible model for survival data with a cure rate: A Bayesian approach. *Journal of Applied Statistics*, 38(1), 57-70.

8. Carrasco, J.M., Ortega, E.M. and Cordeiro, G.M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics and Data Analysis*, 53(2), 450-462.
9. Coelho-Barros, E.A., Achcar, J.A. and Mazucheli, J. (2017). Cure rate models considering the Burr XII distribution in presence of covariate. *Journal of Statistical Theory and Applications*, 16(2), 150-164.
10. Corbiere, F., Commenges, D., Taylor, J.M. and Joly, P. (2009). A penalized likelihood approach for mixture cure models. *Statistics in medicine*, 28(3), 510{524.
11. Gallardo, D.I., Gómez, H.W. and Bolfarine, H. (2017). A new cure rate model based on the Yule-Simon distribution with application to a melanoma data set. *Journal of Applied Statistics*, 44(7), 1153-1164.
12. Gamel, J.W., McLean, I.W. and Rosenberg, S.H. (1990). Proportion cured and mean log survival time as functions of tumour size. *Statistics in Medicine*, 9(8), 999-1006.
13. Ishaq, A., Usman, A., Tasi'u, M., Aliyu, Y. and Idris, F. (2017). A new weibull-kumaraswamy distribution: Theory and applications. *Nigerian Journal of Scientific Research*, 16(2), 158-166.
14. Kannan, N., Kundu, D., Nair, P. and Tripathi, R.C. (2010). The generalized exponential cure rate model with covariates. *Journal of Applied Statistics*, 37(10), 1625-1636.
15. Kutal, D. and Qian, L. (2018). A non-mixture cure model for right-censored data with Fréchet distribution. *Stats*, 1(1), 176-188.
16. Maller, R.A. and Zhou, X. (1996). *Survival analysis with long-term survivors*. Wiley New York.
17. Martinez, E.Z., Achcar, J.A., Jácome, A.A. and Santos, J.S. (2013). Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. *Computer methods and programs in biomedicine*, 112(3), 343-355.
18. Mazucheli, J., Coelho-Barros, E.A. and Achcar, J.A. (2013). The exponentiated exponential mixture and non-mixture cure rate model in the presence of covariates. *Computer Methods and Programs in Biomedicine*, 112(1), 114-124.
19. Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2), 299-302.
20. Ng, S. and McLachlan, G. (1998). On modifications to the long-term survival mixture model in the presence of competing risks. *Journal of Statistical Computation and Simulation*, 61(1-2), 77-96.
21. Ortega, E.M., Cordeiro, G.M. and Kattan, M.W. (2013). The log-beta Weibull regression model with application to predict recurrence of prostate cancer. *Statistical Papers*, 54(1), 113-132.
22. Pal, M., Ali, M.M. and Woo, J. (2006). Exponentiated Weibull distribution. *Statistica*, 66(2), 139-147.
23. Peng, Y., Dear, K.B. and Denham, J. (1998). A generalized F mixture model for cure rate estimation. *Statistics in Medicine*, 17(8), 813-830.
24. Ramos, P.L., Nascimento, D. and Louzada, F. (2017). The long term Fréchet distribution: Estimation, properties and its application. *arXiv preprint arXiv:1709.07593*.

25. Salem, H.M. and Selim, M. (2014). The generalized Weibull-exponential distribution: properties and applications. *International Journal of Statistics and Applications*, 4(2), 102-112.
26. Shao, Q. and Zhou, X. (2004). A new parametric model for survival data with long-term survivors. *Statistics in medicine*, 23(22), 3525-3543.
27. Sy, J.P. and Taylor, J.M. (2000). Estimation in a cox proportional hazards cure model. *Biometrics*, 56(1), 227-236.
28. Usman, U., Shamsuddeen, S., Arkilla, B.M. and Aliyu, Y. (2020). Inferences on the Weibull exponentiated exponential distribution and applications. *International Journal of Statistical Distributions and Applications*, 6(1), 1-10.
29. Usman, U., Suleiman, S., Arkilla, B.M. and Aliyu, Y. (2021). Nadarajah-Haghighi Model for Survival Data with Long Term Survivors in the Presence of Right Censored Data. *Pakistan Journal of Statistics and Operation Research*, 695-709.
30. Yakubu, A. and Doguwa, S.I. (2017). On the properties of the Weibull-Burr III distribution and its application to uncensored and censored survival data. *CBN Journal of Applied Statistics*, 8(2), 91-116.