Kwaghe International
Journal of Arts, Humanities and Religious Studies
ISSN : 0000-0000
Index: Harvard, Boston, Sydney University, Dimensions, Lens, ResearchGet Scilit, Semantic, Google Scholar, Base etc
https://doi.org/10.58578/KIJAHRS.v1i1.3821

The Weibull-Exponentiated Exponential Cure Fraction Model for Right Censored Survival Data with Applications to Cancer Data

Aliyu Yakubu[1*], Naziru Isah Muhammad[2], Jamilu Yunusa Falgore[3], Adam Rabiu[4]
[1,2,3]Ahmadu Bello University, Zaria, Nigeria; [4]National Open University of Nigeria
yakubualiyu@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Aug 20, 2024	Sep 3, 2024	Sep 6, 2024	Sep 9, 2024

Abstract

The cure fraction model also known as the long-term survival model is used in fitting data from a population with two different types of individuals: individuals who experienced the event of interest (susceptible) and individuals who will never experience the event of interest (non-susceptible). The present paper introduced a cure fraction model considering the Weibull exponentiated exponential distribution that will be used in modeling such type of information. The parameters of the model were estimated via the maximum likelihood procedure (MLE) under the assumption of right censoring. Furthermore, the statistical properties of the model were studied comprehensively. Simulation studies and medical data sets were used to demonstrate the applicability of the proposed methodology. Bias and standard error were used as discrimination criteria in the simulation study while Akaike Information criteria (AIC), Bayesian Information Criteria (BIC), and Consistent Akaike Information criteria (CAIC) were used as discrimination criteria in real-life applications. Results from the applications showed that the Weibull exponentiated exponential non-mixture cure fraction model is a strong competitor.

Keywords: Survival Analysis, Mixture Cure Fraction Model, Non-Mixture Cure Fraction Model, Weibull Cure Fraction Model, Right censoring

KIJAHRS Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

Introduction

Survival analyses are statistical methods for analyzing time-to-event data such as death, heart attack, device failure and so on. Different researchers have applied different techniques non-parametric methods (such as Kaplan-Meier estimator or log-rank test), semi-parametric methods (cox proportional hazard model), or parametric methods (using statistical distribution) to analyze this data (Martinez, Achcar, J´acome, & Santos, 2013). The Weibull distribution is one of the distributions that is widely used in this area because of the flexibility of its hazard function and the facility to estimate its parameters (Peto, Lee, & Paige, 1972). However, data sets in medical research require more sophisticated parametric models (Martinez et al., 2013). To solve this problem, new extensions of the Weibull distribution have been proposed by different researchers. For instance, we have the exponentiated Weibull by (Mudholkar & Srivastava, 1993; Pal, Ali, & Woo, 2006), the generalized modified Weibull by (Carrasco, Ortega, & Cordeiro, 2008), log−beta Weibull by (Ortega, Cordeiro, & Kattan, 2013), Weibull exponentiated exponential by (Salem & Selim, 2014; Usman, Shamsuddeen, Arkilla, & Aliyu, 2020), Weibull−Burr III by (Yakubu & Doguwa, 2017), Weibull Kumaraswamy distribution by (Ishaq, Usman, Tasiˆau, Aliyu, & Idris, 2017) among others. One important assumption in survival analysis is that every subject in the study population will eventually experience the event of interest if the follow-up time is large. However, due to the recent advancements in the field of medicine especially in the areas of new drugs and treatment regimens, many subjects have lived longer with diseases such as cancer and heart disease. For instance, it may be observed that cohorts of patients with certain types of cancer have been permanently cured, that is, they show no recurrence of the disease. Those patients who are cured and are not censored are referred to as long-term survivors or non-susceptible, while those who do not develop a recurrence of the disease are termed susceptible. Hence, the data is said to be a mixture of these two types of subjects: susceptible subjects who experienced the event of interest and non-susceptible subjects who will never experience the event of interest (Maller & Zhou, 1996). Cure fraction models are used in modeling such types of data. There are two types of cure fraction models: the mixture and the non-mixture cure fraction models.

The mixture cure rate model also referred to as the standard cure rate model is the most popular type of cure fraction model and was first developed by (Boag, 1949) and further developed by (Berkson & Gage, 1952). The model was later studied extensively by

different researchers including (Achcar, Coelho-Barros, & Mazucheli, 2012; Farewell, 1986; Gamel, McLean, & Rosenberg, 1990; Kannan, Kundu, Nair, & Tripathi, 2010; Mazucheli, Coelho-Barros, & Achcar, 2013; Meeker, 1987; Ng & McLachlan, 1998; Peng, Dear, & Denham, 1998; Shao & Zhou, 2004; Sy & Taylor, 2000; Usman, Suleiman, Arkilla, & Aliyu, 2021; Usman, Shamsuddeen, Arkilla, & Yakubu, 2022) among many others. The study population in the mixture cure rate model assumed that a certain fraction of the population are long-term survivors or non-susceptible while the remaining fraction is susceptible to the event of interest. The non-mixture cure rate model also referred to as the bounded cumulative hazard model or the promotion time cure rate model was first introduced by (Yakovlev et al., 1993) and further discussed by (Andrei, Asselain, et al., 1996; Chen, Ibrahim, & Sinha, 1999; Tsodikov, Ibrahim, & Yakovlev, 2003). The model was motivated by the underlying biological mechanism and was developed under the assumption that the number of cancer cells that remain active after treatment follows poison distribution. Although the mixture cure rate model appears to be attractive and is widely used, Chen et al. (1999); and Uddin, Islam, and Ibrahim (2006), have identified some drawbacks of the model: the model cannot have a proportional hazard structure in the presence of covariates, the model yields an improper posterior distribution for any type of non-informative priors when covariates are included through the cure fraction parameter and the model does not appear to describe the underlying biological process generating the failure time in the context of relapse where cure rate model is frequently used.

To model the proportion of immune, different parametric and non-parametric models have been used by different researchers. For example, the survival function of the failure time of uncured patients was modeled by a product of a log-normal survival function and the survival function of some background distribution for the normal population (Boag, 1949). Jones, Powles, Machin, and Sylvester (1981) applied the exponential distribution for uncured patients in their mixture model while a simulation study for this model was carried out (Goldman, 1984). This method was further studied by (Ghitany & Maller, 1992). The Weibull distribution was used in modeling the failure time of uncured patients by different researchers such as (Farewell, 1982, 1986; Ghitany & Maller, 1992). Peng et al. (1998) modeled the proportion of large-scale clinical trials with long follow-up of lymphoma patients using the generalized F distribution. Other distributions used to model the proportion of non-susceptible are exponentiated-weibull by (Cancho & Bolfarine,

2001), Burr XII distribution by Shao and Zhou (2004); Coelho-Barros, Achcar, and Mazucheli (2017), exponentiated exponential distribution by (Kannan et al., 2010; Mazucheli et al., 2013), negative binomial distribution by Cancho, Rodrigues, and Castro (2011), Weibull distribution by (Achcar et al., 2012), generalized modified Weibull distribution by (Martinez et al., 2013), Frechet distribution by (Ramos, Nascimento, & Louzada, 2017; Kutal & Qian, 2018), Nadarajah-Haghighi distribution by (Usman et al., 2021). The parameters of the non-mixture cure fraction model were estimated via a maximum likelihood estimation procedure considering the data to be uncensored (Uddin, Islam, & Ibrahim, 2006). While (Uddin, Sen, Noor, Islam, & Chowdhury, 2006) estimate the parameters of the non-mixture cure fraction model assuming uncensored data using a non-parametric maximum likelihood method of estimation. A semi-parametric maximum likelihood estimation procedure for the non-mixture model for interval-censored time-to-event data was developed by (Liu & Shen, 2009). Classical and non-classical methods of estimation were used to estimate the parameters of the non-mixture cure fraction model (Lopes & Bolfarine, 2012). Herring and Ibrahim (2002) introduced a parametric method for estimating the parameters of the non-mixture model for non-ignorable missing covariates.

In the present article, we introduced a non-mixture cure fraction model for survival data considering the Weibull exponentiated exponential distribution in the presence of cure fraction and censoring.

Properties of the model were studied and applications of the model to some real-life data were provided. The rest of the paper is organized as follows: In section 2, we introduced the Weibull exponentiated exponential distribution to the non-mixture cure fraction model. Section 3 introduces the Maximum Likelihood Function method of estimation in estimating the parameters of the model assuming right censoring. While statistical properties of the model were provided in section 4. Simulations study and applications of the model were respectively provided in sections 5 and 6. We finally conclude in section 7.

The Non-Mixture Cure Rate Model

In this section, an alternative to the Weibull exponentiated exponential mixture cure rate model introduced by (Usman *et al.*, 2022) called the Weibull exponentiated exponential

non-mixture cure rate model was introduced considering the Weibull exponentiated exponential distribution in the presence of censoring.

Weibull exponentiated exponential distribution was introduced (Salem & Selim, 2014) and later studied extensively by (Usman *et al.*, 2020). The distribution is an extension of the well-known Weibull distribution. The probability density function (*pdf*) of the distribution was shown to take several shapes. The graph of the hazard rate function of the distribution was shown to take various shapes which makes it to be more flexible in modeling real-life data. The *pdf*, cumulative distribution function (*CDF*), survival function, and hazard rate function of the Weibull exponentiated exponential distribution are respectively given as:

$$f(t)=\frac{\alpha\beta\theta\phi\exp(-\beta t)}{1-(1-\exp(-\beta t))^{\alpha}}\left[-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right]^{\theta-1} (1-\exp(-\beta t))^{\alpha-1}\exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right)^{\theta}\right] \quad (1)$$

$$F(t)=1-\exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right)^{\theta}\right] \quad (2)$$

$$S(t)=\exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right)^{\theta}\right] \quad (3)$$

and

$$h(t)=\frac{\alpha\beta\theta\phi\exp(-\beta t)(1-\exp(-\beta t))^{\alpha-1}\left[-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right]^{\theta-1}}{1-(1-\exp(-\beta t))^{\alpha}} \quad (4)$$

where $a,\beta,\phi,$ and θ are positive parameters, a and φ are the shape parameters, θ and β are the scale parameters. As earlier mentioned, the non-mixture cure rate model was motivated by the underlying biological mechanism, and the survival function following (Chen *et al.*, 1999) is developed as follows:

Assume N is the number of cancer cells for a subject after treatment. Assume that the number of cancer cells is Poisson distributed with parameter μ since the number of cancer cells may grow rapidly and produce a detectable cancer disease. Also, let Z_k denote the random time for the *kth* cancer cell to produce a detectable cancer mass. Assuming Z_k are independently and identically distributed (iid) with a common distribution function and

survival function ($F(T)$ and $S(T)$). Assume further, that Z_k is independent of N. Then, the time to relapse of cancer is defined by the random variable $T \in T = \min\{Z_k, 0 \leq k \leq N\}$, where $P(Z_0 = \infty) = 1$. Hence, the survival function of T is given by:

$$\begin{aligned} S(t) &= P(number\ \ of\ \ cancer\ \ by\ \ time\ \ t) \\ &= P(N=0) + P(Z_1 > t, Z_2 > t, \ldots, Z_N > t, N \geq 1) \\ &= \exp(-\mu) + \sum_{N=1}^{\infty} S^N(t) \frac{\mu^N}{N!} \exp(-\mu) \\ &= \exp(-\mu + \mu S_0(t)) \\ &= \exp(-\mu F_0(t)) \\ S(t) &= p^{F_0(t)} \end{aligned} \tag{5}$$

where $p = \exp(-\mu)$ is the proportion of non-susceptible that lies in the interval [0,1]. The corresponding *cdf*, *pdf,* and hazard rate function of the non-mixture cure rate model are respectively given as:

$$F(t) = 1 - p^{F_0(t)} \tag{6}$$

$$f(t) = -\ln p f_n(t) p^{F_0(t)} \tag{7}$$

and

$$h(t) = -\ln p f_0(t) \tag{8}$$

where $S_0(t)$ is the survival function for the susceptible group. Considering the *WEE* distribution, the survival function, *pdf* and hazard rate function for the Weibull exponentiated exponential non-mixture cure rate model (*WEENMCR*) are respectively:

$$S(t/\alpha,\beta,\theta,\phi) = p^{1-\exp\left[-\phi\left(-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right)^{\theta}\right]} \tag{9}$$

$$\begin{aligned} f(t/\alpha,\beta,\theta,\phi) = {} & \frac{-\ln(p)\alpha\beta\theta\phi\exp(-\beta t)}{1-\left(1-\exp(-\beta t)\right)^{\alpha}} \left[-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right]^{\theta-1} \\ & \exp\left[-\phi\left(-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right)^{\theta}\right] \\ & \left(1-\exp(-\beta t)\right)^{\alpha-1} p^{\exp\left[-\phi\left(-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right)^{\theta}\right]} \end{aligned} \tag{10}$$

and

$$h(t/\alpha,\beta,\theta,\phi) = \frac{-\ln(p)\alpha\beta\theta\phi\exp(-\beta t)}{1-(1-\exp(-\beta t))^{\alpha}}(1-\exp(-\beta t))^{\alpha-1} \left[-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right]^{\theta-1} \exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right)^{\theta}\right] \quad (11)$$

where $a,\theta > 0$ are the shape parameters, $\beta,\phi > 0$ are scale parameters and p is the proportion of non-susceptible and it lies between zero and one. The *WEENMCR* model as its mixture counterpart also contains some well-known non-mixture cure rate models in the literature as special cases.

Sub-models of the WEENMCR Model

- when $\theta = \phi = 1$, the *WEENMCR* model reduces to the exponentiated exponential non-mixture model proposed by (Mazucheli *et al.*, 2013).
- when $a = \beta = 1$, it reduces to the Weibull non-mixture cure rate model investigated by (Achcar *et al.*, 2012).
- when $\theta = a = \beta = 1$, it reduces to the exponential non-mixture cure rate model.

Maximum Likelihood Estimation and Likelihood Ratio Test

Let t_i for $i = 1,2,\cdots,n$ be right-censored survival time for the *ith* subject in the study population. Assume δ_i be 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. Then, $t_i = min(T_i,\delta_i)$. The likelihood function for the *WEENMCR* model is obtained as:

$$
\begin{aligned}
L(\Phi) &= \prod_{i=1}^{n} h(t_i)^{\delta_i} S(t_i) \\
&= \prod_{i=1}^{n} \left[\frac{-\ln(p)\alpha\beta\theta\phi \exp(-\beta t)}{1-(1-\exp(-\beta t))^{\alpha}} (1-\exp(-\beta t))^{\alpha-1} \left[-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\} \right]^{\theta-1} \right. \\
&\left. \exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\} \right)^{\theta} \right] \right]^{\delta_i} p^{1-\exp\left[-\phi\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\} \right)^{\theta} \right]}
\end{aligned}
\tag{12}
$$

The log-likelihood function is obtained by taking the natural logarithm of (12) which yields:

$$
\begin{aligned}
\ell(\Phi) &= q\ln(-\ln(p)) + q\ln(\alpha) + q\ln(\beta) + q\ln(\theta) + q\ln(\phi) - b\sum \delta_i t_i + \\
&(\alpha-1)\sum \delta_i \ln\left(1-e^{-\beta t_i}\right) - \sum \delta_i \ln\left(1-\left(1-e^{-\beta t_i}\right)^{\alpha}\right) + \\
&(\theta-1)\sum \delta_i \ln\left(-\ln\left(1-\left(1-e^{-\beta t_i}\right)^{\alpha}\right)\right) - \phi\sum \delta_i \left(-\ln\left(1-\left(1-e^{-\beta t_i}\right)^{\alpha}\right)\right)^{\theta} + \\
&\ln(p)\sum\left[1-\exp\left[-\phi\left(-\ln\left(1-(1-\exp(-\beta t))^{\alpha}\right)\right)^{\theta}\right]\right]
\end{aligned}
\tag{13}
$$

differentiating (13) partially with respect to *a, β, θ, ϕ,* and p and equating to zero gives the score function as:

$$
\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{q}{\alpha} - \frac{1}{\alpha}\sum \frac{\ln(b_i)}{1-b_i} + \frac{(\theta-1)}{\alpha}\sum \frac{(1-b_i)\ln(b_i)\delta_i}{c_i} + \\
&\frac{(\theta-1)}{\alpha}\sum b_i(1-b_i)\ln(b_i)(-c_i)^{\theta-1}(\delta_i + d_i)
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
\frac{\partial \ell}{\partial \beta} &= \frac{q}{\beta} - \sum \delta_i t_i + (\alpha-1)\sum \frac{\delta_i t_i a_i}{1-a_i} + \alpha\sum \frac{\delta_i t_i a_i (1-a_i)^{\alpha-1}}{1-b_i} + \\
&\alpha(\theta-1)\sum \frac{\delta_i t_i a_i (1-a_i)^{\alpha-1}(1-b_i)}{c_i} + \alpha\theta\phi\sum t_i a_i (1-a_i)^{\alpha-1}(1-b_i)(c_i)^{\theta-1}(\delta_i - d_i)
\end{aligned}
\tag{15}
$$

$$
\frac{\partial \ell}{\partial \theta} = \frac{q}{\theta} + \sum \delta_i \ln(-c_i) - \phi\sum (-c_i)^{\theta}\ln(-c_i)(\delta_i - d_i \ln(p))
\tag{16}
$$

$$\frac{\partial \ell}{\partial \phi} = \frac{q}{\phi} + \sum (-c_i)^{\theta} \left(d_i \ln(p) - \delta_i \right) \tag{17}$$

$$\frac{\partial \ell}{\partial p} = \frac{q}{p} + \frac{1}{p} \sum (1 - d_i) \tag{18}$$

where $a_i = e^{\beta t_i}$, $b_i = (1 - a_i)^{\alpha}$, $c_i = ln(1 - b_i)$ and $d_i = e^{(-\varphi(-\alpha_i))\theta}$. However, equating (18) to zero and solving for p yields

$$\hat{p}(\alpha, \beta, \theta, \phi) = \exp\left(\frac{r}{\sum d_i - n} \right) \tag{19}$$

Thus, the *MLE* of p can be obtained algebraically using (19) while equations (14) to (17) can easily be solved using numerical methods. Interval estimation and hypothesis testing on the parameters of the *WEENMCR* model can be studied using the observed fisher information matrix $I(\Phi)$ given as:

$$I(\Theta) = -\begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\theta} & I_{\alpha\phi} & I_{\alpha p} \\ & I_{\beta\beta} & I_{\beta\theta} & I_{\beta\phi} & I_{\beta p} \\ & & I_{\theta\theta} & I_{\theta\phi} & I_{\theta p} \\ & & & I_{\phi\phi} & I_{\phi p} \\ & & & & I_{pp} \end{pmatrix} \tag{20}$$

Where $I_{aa} = \frac{\partial^2 \ell}{\partial a^2}$ and $I_{ab} = \frac{\partial^2 \ell}{\partial a \partial b}$. The asymptotic distribution of $\sqrt{n}\left(\hat{\Theta} - \Theta\right)$ is multivariate normal $N_5\left(0, I\left(\hat{\Theta}\right)^{-1}\right)$, where $I\left(\hat{\Theta}\right)^{-1}$ is the total observed information matrix computed at $\hat{\Theta}$. The diagonal elements $I\left(\hat{\Theta}\right)^{-1}$ are variances of the corresponding parameters while the off-diagonal elements of $I\left(\hat{\Theta}\right)^{-1}$ are covariances. Hence, the asymptotic $100(1-\varepsilon)\%$ confidence interval for any of the parameters of the *WEENMCR* model is respectively

$$\alpha \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\alpha)},\ \beta \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\beta)},\ \theta \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\theta)},\ \phi \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\phi)} \text{ and } p \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(p)},$$

where $z_{a/2}$ is the $100(1 - \varepsilon)\%$ quantile from the standard normal distribution.

Statistical properties of the WEENMCR Model

Statistical properties such as quantile function, median, simulation, moment generating function, and moments of the *WEENMCR* model were discussed in this section.

Quantile Function and Simulation

The quantile function of the *WEENMCR* model is obtained as:

$$Q(u) = -\frac{1}{\beta}\ln\left\{1-\left[1-\exp\left\{-\left(-\frac{1}{\phi}\ln\left[\frac{\ln(1-u)}{\ln(p)}\right]\right)^{1/\theta}\right\}\right]^{1/\alpha}\right\} \tag{21}$$

by letting $u = 1 - p^{1-\exp\left\{-\phi\left(-\ln\left[1-\{1-\exp(-\beta t)\}^{\alpha}\right]\right)^{\theta}\right\}}$, where u follows the uniform distribution with parameters zero and one. The first, second, and third quantiles of the *WEENMCR* model are obtained by substituting u = 0.25, 0.50, and 0.75 respectively in equation (21). Hence, the median of the *WEENMCR* model is given by:

$$\text{median} = -\frac{1}{\beta}\ln\left\{1-\left[1-\exp\left\{-\left(-\frac{1}{\phi}\ln\left[\frac{\ln(0.5)}{\ln(p)}\right]\right)^{1/\theta}\right\}\right]^{1/\alpha}\right\} \tag{22}$$

However, the median of the WEENMCR model is zero when p takes the value of 0.5. Galton's coefficient of skewness and Moor's coefficient of kurtosis can easily be obtained using the quantile function in (21) by making the appropriate substitutions.

Simulation

A simulation study could be used to examine the performance of the Maximum Likelihood Estimation method discussed in section 3. To generate right-censored survival time data from the *WEENMCR* model inverse transform method could be employed since the quantile function of the *WEENMCR* model is in closed form. The following algorithm could be adapted to generate a random sample of size n from the introduced methodology.

1. Generate a random sample of size n u_i for i = 1,2,⋯n from the uniform distribution. That is, $u_i \sim U(0,1)$.
2. Assume p is a cure fraction parameter, and return the random survival time

$$t_i = -\frac{1}{\beta}\ln\left\{1-\left[1-\exp\left\{-\left(-\frac{1}{\phi}\ln\left[\frac{\ln(1-u)}{\ln(p)}\right]\right)^{1/\theta}\right\}\right]^{1/\alpha}\right\}$$ when $u_i \leq 1 - p$ else t_i is infinity.

3. Generate the censoring times c_i for $i = 1,\cdots,n$ from the *WEE* distribution.
4. $z_i = min(t_i, c_i)$ is the obtained right censored survival time.
5. The observed right censored survival data is $Z = (z_i, \delta_i)$ for $i = 1,2,\cdots,n$, where δ_i is a censoring indicator.

Characteristic Function

Assume T is a random variable that follows the *WEENMCR* model with survival and pdf given by equations (8) and (10) respectively. This pdf however, can be written in the form $f(t) = -ln(p)\, f_u(t)\exp(-(-ln(p))F_u(t))$ following Ibrahim *et al.* (2001). Applying power series expansion gives:

$$f(t) = -ln(p)\, f_u(t)\sum_{j=1}^{\infty}\frac{(-1)^j}{j!}\left(-ln(p)\right)^j F_u(t)^j \tag{23}$$

substituting equations (1) and (2) in (23) yields:

$$f(t) = \frac{\alpha\beta\theta\phi\exp(-\beta t)}{1-\left(1-\exp(-\beta t)\right)^{\alpha}}\left(1-\exp(-\beta t)\right)^{\alpha-1}\left[-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right]^{\theta-1}$$
$$\sum_{j,k=1}^{\infty}\frac{(-1)^{j+k}\left(-ln(p)\right)^{j+1}}{(j-k)!k!}\exp\left[-\phi(1+k)\left(-\ln\left\{1-\left(1-\exp(-\beta t)\right)^{\alpha}\right\}\right)^{\theta}\right]$$

Let $\gamma = \phi(j+1)$ and $\delta_{j,k} = \frac{(-1)^{j+k}\left(-ln(p)\right)^{j+1}}{(j-k)!(j+1)!}$, then

$$f(t) = \sum_{j,k=0}^{\infty}\delta_{jk}\, f(t|\alpha,\beta,\theta,\gamma) \tag{24}$$

where $f(t/a, \beta, \theta, \gamma)$ is the pdf of WEE distribution with parameters a, β, θ, γ. The characteristic function of the *WEENMCR* model denoted by $\varphi_X(t)$ is evaluated as:

$$\begin{aligned}\varphi_X(t) &= \int_{-\infty}^{\infty} e^{itx} f(x)dx \\ &= \int_{-\infty}^{\infty} e^{itx} \sum_{j,k=0}^{\infty} \delta_{jk} f(x|\alpha,\beta,\theta,\gamma)dx \\ &= \sum_{j,k=0}^{\infty} \delta_{jk} \int_0^{\infty} e^{itx} f(x|\alpha,\beta,\theta,\gamma)dx\end{aligned} \tag{25}$$

and $i = \sqrt{-1}$. Let

$$\begin{aligned}I = \int_0^{\infty} e^{itx} \frac{\alpha\beta\theta\phi \exp(-\beta t)}{1-(1-\exp(-\beta t))^{\alpha}} \left[-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right]^{\theta-1} \\ (1-\exp(-\beta t))^{\alpha-1} \exp\left[-\gamma\left(-\ln\left\{1-(1-\exp(-\beta t))^{\alpha}\right\}\right)^{\theta}\right] dx\end{aligned} \tag{25}$$

let $u = \gamma(-ln\{1-(1-\exp(-\beta x))^{a}\})^{\theta}$ then

$$I = \int_0^{\infty} \left[1-\left(1-\exp\left(-\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)\right)^{\frac{1}{\alpha}}\right]^{-\frac{it}{\beta}} e^{-u} du \tag{26}$$

the series expansion of $(1-x)^{-q} = \sum_{l=0}^{\infty} \frac{(q+l-1)!}{(q-1)!l!} x^l$. Therefore,

$$\left[1-\left(1-\exp\left(-\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)\right)^{\frac{1}{\alpha}}\right]^{-\frac{it}{\beta}} =$$

$$\sum_{l=0}^{\infty} \frac{\left(\frac{it}{\beta}+l-1\right)!}{\left(\frac{it}{\beta}-1\right)!l!} \left(1-\exp\left(-\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)\right)^{\frac{1}{\alpha}}$$ as u tends to zero and as u tends to infinity, the limiting values $\left(1-\exp\left(-\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)\right)^{\frac{1}{\alpha}}$ are between zero and one, hence applying binomial series expansion yields:

$$\left[1-\left(1-\exp\left(-\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)\right)^{\frac{1}{\alpha}}\right]^{-\frac{it}{\beta}} = \sum_{l,m}^{\infty}\frac{(-1)^m\left(\frac{1}{\alpha}\right)!\left(\frac{it}{\beta}+l-1\right)!}{\left(\frac{it}{\beta}-1\right)!\left(\frac{1}{\alpha}-m\right)!l!m!}\exp\left(-m\left(\frac{u}{\gamma}\right)^{\frac{1}{\theta}}\right)$$
$$= \sum_{l,m,n}^{\infty}\frac{(-1)^{m+n}\left(\frac{1}{\alpha}\right)!\left(\frac{it}{\beta}+l-1\right)!m^n\gamma^{\frac{n}{\theta}}}{\left(\frac{it}{\beta}-1\right)!\left(\frac{1}{\alpha}-m\right)!l!m!n!}u^{\frac{n}{\theta}} \quad (28)$$

substituting in (26), the integral becomes:

$$I = \sum_{l,m,n}^{\infty}\frac{(-1)^{m+n}\left(\frac{1}{\alpha}\right)!\left(\frac{it}{\beta}+l-1\right)!m^n\gamma^{\frac{n}{\theta}}}{\left(\frac{it}{\beta}-1\right)!\left(\frac{1}{\alpha}-m\right)!l!m!n!}\int_0^{\infty}u^{\frac{n}{\theta}}e^{-u}du \quad (29)$$

and substituting (29) in (24), $\varphi_X(t)$ becomes:

$$\varphi_T(x) = \sum_{j,k,l,m,n=0}^{\infty}\delta_{jk}\frac{(-1)^{m+n}\left(\frac{1}{\alpha}\right)!m^n\gamma^{\frac{n}{\theta}}\Gamma(\frac{n}{\theta}+1)}{\left(\frac{1}{\alpha}-m\right)!l!m!n!}\left(\frac{it}{\beta}\right)_l \quad (30)$$

where $\left(t/\beta\right)_l = t/\beta\left(t/\beta+1\right)\left(t/\beta+2\right)\cdots\left(t/\beta+k-1\right)$.

Moment Generating Function

The moment generating function *mgf* of the *WEENMCR* model denoted by $M_X(t)$ is defined by $M_X(t) = E(e^{tx})$. This is evaluated as:

$$M_X(t) = \int_{-\infty}^{\infty}e^{tx}f(x)dx$$
$$= \int_{-\infty}^{\infty}e^{tx}\sum_{i,j=0}^{\infty}\delta_{jk}f(x|\alpha,\beta,\theta,\gamma)dx \quad (31)$$
$$= \sum_{j,k=0}^{\infty}\delta_{jk}\int_0^{\infty}e^{tx}f(x|\alpha,\beta,\theta,\gamma)dx$$

following the same procedure as the characteristic function, the integral part is evaluated as:

$$\int_0^{\infty}e^{tx}f(x|\alpha,\beta,\theta,\gamma)dx = \sum_{k,l,m=0}^{\infty}\frac{(-1)^{l+m}\left(\frac{k}{\alpha}\right)!l^m\gamma^{\frac{m}{\theta}}\Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)!k!l!m!}\left(\frac{t}{\beta}\right)_k \quad (32)$$

where $\left(t/\beta\right)_k = t/\beta\left(t/\beta+1\right)\left(t/\beta+2\right)\cdots\left(t/\beta+k-1\right)$. Substituting (32) in (31), the moment-generating function of the *WEENMCR* model is derived as:

$$M_X(t) = \sum_{i,j,k,l,m=0}^{\infty} \frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)! l^m \gamma^{\frac{m}{\theta}} \Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)! k! l! m!}\left(\frac{t}{\beta}\right)_k \tag{33}$$

Moments

The r^{th} moment about the origin is easily obtained from the moment-generating function using the relation $E(T^r) = \frac{d^r}{dt^r}\left(M_X(t)\right)\big|_{t=0}$. Hence,

$$E(T^r) = \sum_{i,j,k,l,m=0}^{\infty} \frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)! l^m \gamma^{\frac{m}{\theta}} \Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)! k! l! m!} \frac{d^r}{dt^r}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0} \tag{33}$$

For instance, the first and second moments about the origin are obtained as follows:

$$E(T) = \sum_{i,j,k,l,m=0}^{\infty} \frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)! l^m \gamma^{\frac{m}{\theta}} \Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)! k! l! m!} \frac{d}{dt}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0}$$

but $\frac{d}{dt}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0} = (k-1)!/b$ Hence, substituting in (33), the first moment becomes:

$$E(T) = \sum_{i,j,k,l,m=0}^{\infty} \frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)! l^m \gamma^{\frac{m}{\theta}} \Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)! l! m! kb} \tag{34}$$

on the other hand, the second moment is obtained as:

$$E(T^2) = \sum_{i,j,k,l,m=0}^{\infty} \frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)! l^m \gamma^{\frac{m}{\theta}} \Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)! k! l! m!} \frac{d^2}{dt^2}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0}$$

but the term $\frac{d^2}{dt^2}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0}$ can be evaluated as: $\frac{d^2}{dt^2}\left(\left(\frac{t}{\beta}\right)_k\right)\Big|_{t=0} = \frac{2(k-1)!}{b^2}(\psi(k)-\psi(1))$

Hence, substituting in (33), the second moment becomes:

$$E(T^2)=\frac{2}{b^2}\sum_{i,j,k,l,m=0}^{\infty}\frac{(-1)^{l+m}\delta_{ij}\left(\frac{k}{\alpha}\right)!l^m\gamma^{\frac{m}{\theta}}\Gamma(\frac{m}{\theta}+1)}{\left(\frac{k}{\alpha}-l\right)!l!m!k}(\psi(k)-\psi(1)) \quad (35)$$

Equations (34) and (35) can be used to find the variance of the *WEENMCR* model using the relation $var(X)= E(X^2) - [E(X)]^2$

Simulation Study

In this section, simulation studies was conducted so as to ascertain the performance of the maximum likelihood estimator of $\Theta = (a,\beta,\theta,\phi,p)'$ discussed in section 3. The algorithm discussed under quantile function was used in generating right censored survival times.

Table 1: Maximum Likelihood, bias and standard error (SE)

n	Para- meters	a=1.5; b=2.0; θ=3.0; ϕ =2.0; p=0.05			a=1.5; b=2.0; θ=3.0; ϕ =2.0; p=0.10			a=1.5; b=2.0; θ=3.0 ; ϕ =2.0; p=0.15		
		estim-ates	bias	SE	estim-ates	bias	SE	estim-ates	bias	SE
30	a	2.9505	1.0905	2.1459	2.6343	1.1343	2.896	2.9113	1.4113	3.1985
	b	2.3951	-1.6049	0.9266	2.2513	-1.2487	1.306	2.3826	-1.6174	1.4735
	θ	2.8962	1.3862	1.4916	3.153	1.4153	1.8446	3.7062	1.2762	1.9604
	ϕ	2.9372	1.4372	3.9728	2.9822	1.4822	5.1827	3.0518	1.6518	5.2474
	p	0.0529	0.0029	0.0408	0.1048	0.0048	0.0545	0.1511	0.0011	0.0633
50	a	2.2343	-1.0167	1.8075	2.2982	-1.1218	2.0291	2.5763	1.2763	2.2437
	b	2.1406	-1.2994	0.932	2.1844	-1.2156	1.0383	2.3092	-1.3908	1.3546
	θ	2.9838	1.2993	1.3216	3.1486	1.386	1.6923	3.4071	1.2717	1.5334
	ϕ	2.6591	1.1591	3.1641	2.9014	1.4014	3.4856	2.9007	1.6007	4.2769
	p	0.0529	0.0029	0.0316	0.1044	0.0044	0.0423	0.1515	0.0015	0.0497
75	a	1.8289	-0.9871	1.2122	1.9383	-0.9617	1.3791	2.1755	-1.0245	1.6114
	b	1.8375	-1.2625	0.6114	1.9096	1.2024	0.7957	2.0974	-1.1026	0.8775
	θ	3.1929	1.2929	1.2006	3.1171	1.3701	1.2759	3.3697	1.2697	1.2796
	ϕ	2.5902	1.0902	2.8256	2.8621	1.3621	2.8527	1.7876	1.5876	3.7113
	p	0.0515	0.0025	0.0255	0.1026	0.0026	0.0345	0.1516	0.0016	0.0409
100	a	1.6895	-0.9705	1.0079	1.7765	-0.9235	1.1558	1.9822	-1.0078	1.3903
	b	1.8968	-1.2132	0.6108	1.9174	1.1986	0.7347	2.0233	-1.0267	0.8132
	θ	3.1595	1.2595	1.1692	3.1377	1.3337	1.2704	3.1716	1.2476	1.2705
	ϕ	2.4117	0.9117	2.644	2.6052	1.1052	2.7345	2.5063	1.5263	3.6644
	p	0.0512	0.0021	0.0222	0.1026	0.0026	0.0301	0.1515	0.0015	0.0355
150	a	1.5699	-0.9301	0.7909	1.6296	-0.8704	0.9649	1.806	-0.9694	1.0846
	b	1.753	-1.2047	0.476	1.9382	-1.1618	0.6365	1.8813	-1.1874	0.7064
	θ	3.1298	1.2681	0.9147	3.1505	1.3205	1.0583	3.1265	1.2265	1.1225
	ϕ	2.2931	0.7931	2.0383	2.2626	0.7626	2.0341	2.3861	1.5261	3.1747

	p	0.0519	0.0019	0.0182	0.102	0.0024	0.0246	0.1508	0.0008	0.0291
200	a	1.3985	-0.9015	0.7008	1.5157	-0.7843	0.7218	1.6883	-0.8117	0.834
	b	1.8494	-1.1506	0.4566	1.9585	-1.1415	0.5415	1.8942	-1.1258	0.5395
	θ	3.1295	1.1395	0.8424	3.1352	1.3052	0.9344	3.1147	1.1814	0.8834
	ϕ	2.2045	0.5045	1.665	2.1628	0.6628	1.8247	1.8466	1.3466	2.2782
	p	0.0508	0.002	0.0159	0.1021	0.0021	0.0213	0.1508	0.0008	0.0251
250	a	1.4124	-0.8876	0.6706	1.4933	-0.7067	0.6705	1.6126	-0.8074	0.7504
	b	1.8768	-1.1232	0.4192	1.9601	-1.1399	0.4585	1.9333	-1.1067	0.4662
	θ	3.0846	1.0846	0.8056	3.1295	1.295	0.864	3.1095	1.1795	0.8456
	ϕ	1.9864	0.4864	1.4778	2.0845	0.5845	1.461	2.2909	1.1609	1.7938
	p	0.0509	0.0019	0.0141	0.1017	0.0017	0.019	0.1507	0.0007	0.0225
300	a	1.5636	-0.364	0.6441	1.4963	-0.6037	0.6234	1.5538	-0.7462	0.6454
	b	1.9398	-1.0602	0.4166	1.9647	-1.1053	0.4297	1.9494	-1.0506	0.4484
	θ	3.0514	1.0514	0.7157	3.1074	1.2774	0.7874	3.1076	1.1676	0.7623
	ϕ	1.9857	0.357	1.2282	2.0275	0.4275	1.0718	2.2717	1.0717	1.0573
	p	0.0506	0.0012	0.013	0.1015	0.0015	0.0174	0.1501	0.0001	0.0206

Samples of size n = 30,50,75,100,150,200,250 and 300 with different proportions of cure fraction values were generated for the parameter values a = 2.5,β = 3.0,θ = 2.0, and ϕ = 1.5 as the first setting and a = 1.75, β = 2.0, θ = 1.5 and ϕ = 2.5 as the second setting. In each of these settings, the cure fraction parameter takes the values p = 0.05, 0.10, and 0.15.

Table 2: Maximum Likelihood, bias, and standard error (SE)

N	Para-meters	a=1.25; b=1.5; θ=2.5; ϕ =2.5; p=0.05			a=1.25; b=1.5; θ=2.5; ϕ =2.5; p=0.1			a=1.25; b=1.5; θ=2.5 ; ϕ =2.5; p=0.15		
		estim-ates	bias	SE	estim-ates	bias	SE	estim-ates	bias	SE
30	a	1.9989	0.9489	1.7998	2.0342	0.8242	1.9156	2.3819	0.6319	1.9308
	b	1.8751	0.9849	1.1941	2.1065	0.9565	1.3272	2.4368	0.4368	1.438
	θ	2.3465	-1.2865	1.5477	2.4601	1.3601	1.6506	2.3806	1.6806	1.6032
	ϕ	2.7076	0.2576	4.061	2.6126	0.1626	3.8456	2.3579	-0.1681	3.5531
	p	0.0514	0.0014	0.046	0.1032	0.0032	0.06	0.1495	-0.0008	0.0694
50	a	1.6639	-0.8861	1.2628	1.6491	-0.8109	1.432	1.9073	0.5173	1.4856
	b	1.6234	-0.9766	0.8516	1.9498	-0.9503	1.1499	2.1436	0.3956	1.222
	θ	2.3969	1.1969	1.2088	2.5913	-1.0513	1.3782	2.4407	1.4097	1.3261
	ϕ	2.6907	0.1907	3.3803	2.448	-0.152	2.9477	2.3855	-0.1645	2.918
	p	0.051	0.001	0.0349	0.1033	0.0031	0.0459	0.1506	0.0006	0.0533
75	a	1.3818	-0.8682	1.0119	1.4314	-0.6816	1.1016	1.586	-0.4614	1.2573
	b	1.4461	-0.9039	0.6709	1.7073	-0.8927	0.9369	1.976	-0.3624	1.1228
	θ	2.5113	1.0113	1.0708	2.5559	1.0559	1.3548	2.475	1.0975	1.1507
	φ	2.6843	0.1843	3.3502	2.6037	0.1237	2.5144	2.49	-0.1407	2.6273
	p	0.0508	0.0008	0.0285	0.1018	0.0018	0.0375	0.1501	0.0004	0.0436
100	a	1.3401	-0.7099	0.8425	1.2876	-0.6424	0.8453	1.403	-0.4347	0.9767

	b	1.4734	-0.8666	0.6539	1.6564	-0.7436	0.7428	1.8749	-0.3251	0.9068
	θ	2.5043	1.0043	1.0395	2.5391	1.0191	1.063	2.5931	1.0231	1.1294
	ϕ	2.4244	-0.0756	2.51	2.4832	-0.1168	2.444	2.4637	-0.1363	2.3595
	p	0.0498	-0.0008	0.0247	0.1012	0.0012	0.0324	0.1504	0.0004	0.0377
150	a	1.2741	-0.6759	0.6884	1.1896	-0.5604	0.7082	1.2462	-0.3938	0.8546
	b	1.5347	-0.7753	0.5352	1.5859	-0.6141	0.7142	1.7741	-0.2959	0.9036
	θ	2.492	0.992	0.8984	2.5203	-0.9703	0.9616	2.5566	1.0066	1.1113
	ϕ	2.544	0.044	2.5083	2.4194	-0.1086	1.968	2.5182	-0.1348	2.2903
	p	0.0504	0.0004	0.0202	0.1014	0.001	0.0265	0.1501	0.0003	0.0309
200	a	1.2491	-0.5009	0.5038	1.1375	-0.5125	0.6152	1.1746	-0.3754	0.6459
	b	1.4945	-0.7055	0.4199	1.6431	-0.5569	0.6275	1.7168	-0.2832	0.7885
	θ	2.4587	0.9517	0.7242	2.5187	0.9187	0.8502	2.5706	0.9706	0.8527
	ϕ	2.5019	0.0402	1.7973	2.5965	-0.0635	1.3769	2.5094	-0.1106	1.7401
	p	0.0502	0.0003	0.0174	0.1013	0.001	0.0229	0.1502	0.0002	0.0268
250	a	1.2674	-0.4826	0.5138	1.2725	-0.4775	0.5548	1.1213	-0.3287	0.6269
	b	1.5047	-0.6953	0.4185	1.6034	-0.3966	0.6323	1.733	-0.267	0.6848
	θ	2.5121	0.9121	0.7175	2.5482	0.9182	0.8201	2.5817	0.9581	0.8094
	ϕ	2.4578	-0.022	1.849	2.4371	-0.0529	1.2676	2.4976	-0.1052	1.539
	p	0.05	0.0003	0.0155	0.1009	0.0009	0.0206	0.1499	-0.0001	0.0239
300	a	1.2477	-0.4023	0.4787	1.2595	-0.3905	0.4383	1.0801	-0.3099	0.4969
	b	1.3935	-0.6665	0.3775	1.5204	-0.4796	0.4733	1.6988	-0.2012	0.6065
	θ	2.4981	0.8981	0.6618	2.5365	0.9135	0.7538	2.5159	0.9405	0.7462
	ϕ	2.5344	-0.0656	1.6005	2.5573	-0.0427	1.1764	2.4946	-0.0854	1.4
	p	0.0508	0.0001	0.0143	0.1007	0.0007	0.0187	0.1507	0.0001	0.0219

As mentioned in the algorithm, the censoring variables were assumed to follow the *WEE* susceptible distribution. The performance of the estimates was assessed using bias and standard error (*SE*) of the estimates. Additionally, in this simulation setting, all results were replicated 1000 times for each parameter setting considered.

Table 1 gives the maximum likelihood estimates together with bias and SE for the parameter settings: $a = 1.5$, $b = 2.0$, $\theta = 3.0$, and $\phi = 2.0$. The maximum likelihood estimates for the parameter setting: $a = 1.25$, $b = 1.5$, $\theta = 2.5$, and $\phi = 2.5$ were given in Table 2. Bias and standard error of the estimates were also given in this table. The bias and standard error of the estimates were found to be small and decreased as the sample size increased for all the different parameter settings. Hence, the estimates get closer to the true parameter value as the sample size increases. This indicates that the proposed method of estimation has a good performance overall.

Real Data Applications

In this section, two data sets were used to illustrate the methodology of the *WEENMCR* model. 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 data consists of 287 patients with high-risk melanomas who were accrued to E1684 between 1984 and 1990. The patients were randomized to the observation group or 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) and were treated either with wide local excision or with complete regional lymph node dissection. Two observations were deleted because they obtained missing information. Hence, the analysis of treatment effects against the observation group was based on 285 patients who were randomized to the IFN or observation group. Furthermore, one hundred and forty (140) patients were in the observation group and one hundred and forty-five (145) patients were in the IFN treatment group.

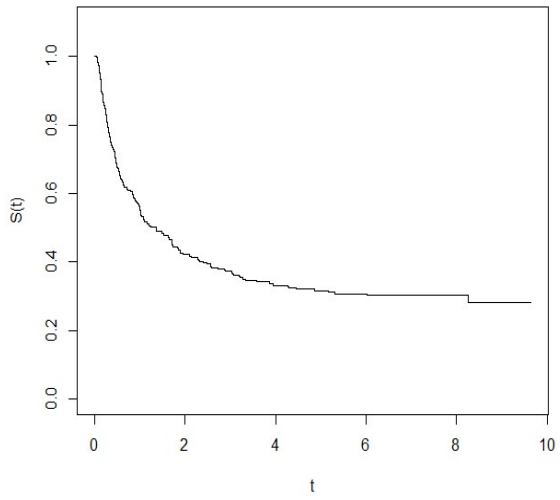


Figure 1: Kaplan-Meier relapse-free survival curve

The Kaplan-Meier survival curve of this data is given in Figure 1. According to (Corbi`ere, Commenges, Taylor, & Joly, 2009), the presence of long-term survivors is usually suggested by the Kaplan-Meier survival curve when the curve level-off. Hence, we observe from the Kaplan-Meier survival curve in Figure 1 that after about 8 years of follow-up, some patients have not experienced any recurrence after treatments. That is, the curve leveled off at a value between 0.15 and 0.2. Hence, we conclude that there is a presence of long-term survivors in the data. The data was fitted to the *WEENMCR*

model and compared its performance with the fits of Weibull non-mixture cure rate *WNMCR*, exponentiated exponential non-mixture cure rate *EENMCR,* and exponential non-mixture cure rate *ENMCR* models. The maximum likelihood estimates of the parameters, their SE, 95% confidence interval (*CI*), and the statistics: *AIC*, *BIC,* and *CAIC* of the fitted models were given in Table 3. From the statistics in this table, we conclude that the *WEENMCR* model is more efficient compared to the *WNMCR*, *EENMCR,* and *ENMRC* models since it has the lowest *AIC*, *BIC,* and *CAIC* values. This can also be observed from the fits of the Kaplan-Meier survival curve overlaid with that of WEENMCR, WNMCR, EENMCR, and ENMCR survival curves shown in Figure 1.

Table 3: Maximum Likelihood Estimates, SE, 95% *CI*, *AIC*, *BIC* and *CAIC* for e1684 data Non-Mixture Models

Model	parameters	estimate	SE	95% CI	AIC	BIC	CAIC
WEENMCR	a	3.8913	0.0021	(3.8873,3.8954)	756.8194	775.0818	747.0345
	β	3.8521	0.0010	(3.8501,3.854)			
	θ	0.5403	0.0396	(0.4626,0.6179)			
	φ	0.3689	0.0534	(0.2642,0.4735)			
	p	0.2672	0.0385	(0.1917,0.3426)			
WNMCR	θ	0.9904	0.0577	(0.8773,1.1036)	772.4918	783.4493	766.5772
	φ	0.6182	0.0613	(0.498,0.7383)			
	p	0.2929	0.0282	(0.2376,0.3483)			
EENMCR	a	1.0741	0.0940	(0.8898,1.2583)	771.7660	782.7235	765.8514
	β	0.6700	0.0868	(0.4998,0.8402)			
	p	0.2911	0.0277	(0.2368,0.3454)			
EENMCR	β	0.6210	0.0612	(0.5011,0.7408)	770.4236	777.7286	766.4662
	p	0.2912	0.0280	(0.2363,0.3461)			

The asymptotic variance-covariance matrix for the proposed *WEENMCR* model fitted to the e1684 data is

$$\begin{pmatrix} & \alpha & \beta & \theta & \phi & p \\ \alpha & 4.29\times10^{-6} & -8.55\times10^{-7} & -4.21\times10^{-7} & 1.86\times10^{-7} & -9.49\times10^{-8} \\ \beta & -8.55\times10^{-7} & 9.47\times10^{-7} & 1.17\times10^{-7} & -1.14\times10^{-7} & 1.59\times10^{-8} \\ \theta & -4.21\times10^{-7} & 1.17\times10^{-7} & 1.57\times10^{-3} & 4.82\times10^{-4} & 6.29\times10^{-4} \\ \phi & 1.86\times10^{-7} & -1.14\times10^{-7} & 4.82\times10^{-4} & 2.8\times10^{-3} & 1.43\times10^{-3} \\ p & -9.49\times10^{-8} & 1.59\times10^{-8} & 6.29\times10^{-4} & 1.43\times10^{-3} & 1.48\times10^{-3} \end{pmatrix}$$

Likelihood Ratio Test

A likelihood ratio test was conducted to test for the superiority of the *WEENMCR* model over its submodels at a 5% significance level. The computed test statistic for the comparison between the *WEENMCR* model with *ENMCR* model, *WEENMCR* model with *WNMCR* model, and *WEENMCR* model with *EENMCR* model are respectively evaluated as:

$\tau_E = 2(- - 373.4097 - (-383.2118)) = 19.6042$ $\tau_W = 2(-373.4097 - (-383.2459)) = 19.6724$

and

$$\tau_{EE} = 2(-373.4097 - (-382.883)) = 18.9466$$

The summary statistics for the likelihood ratio test between the *WEENMCR* model and its sub-models are shown in Table 4. The p-values of *ENMCR*, *WNMCR,* and *EENMCR* models are all significant. This shows that the *WEENMCR* model is more efficient than its sub-models.

Table 4: Likelihood ratio test statistics for the test between *WEENMCR* and *ENMCR*, *WNMCR* and *EENMCR*

Model compared with	Hypothesis	τ	p-value
EMCR Model	$H_0: a = \theta = \varphi = 1$ Vs $H_1: H_0$ is not true	19.6042	0.0002
WMCR Model	$H_0: a = \beta = 1$ Vs $H_1: H_0$ is not true	19.6724	0.0001
EEMCR Model	$H_0: \theta = \varphi = 1$ Vs $H_1: H_0$ is not true	18.9466	0.0001

Malaysian Colorectal Cancer Data

Colorectal cancer has been ranked the third most commonly diagnosed malignancy (Naishadham, Lansdorp-Vogelaar, Siegel, Cokkinides, & Jemal, 2011), the second most frequent cancer in women and the third most frequent cancer in men (Naishadham *et al.*, 2011; Magaji, Moy, Roslani, & Law, 2014). It is also ranked the fourth leading cause of cancer-related death in the world (Magaji *et al.*, 2014; Magaji, Moy, Roslani, & Law, 2017). Medical records of 80 patients diagnosed of colorectal cancer and treated by surgery and chemotherapy/radiotherapy between January 2001 and December 2010 in the University of Malaya Medical Center (UMMC) were obtained. The survival time was defined to be

the time from the date of commencement of treatment to death, loss to follow-up, or end of the study.

Different statistical methods have been used to analyze medical information of patients suffering from colorectal cancer. These include; Magaji *et al.* (2014) analysis of colorectal cancer patients who underwent treatment at the University of Malaya Medical Centre from 2001 to 2010, the rates of survival and its predictors among colorectal cancer patients in Malaysia (Magaji *et al.*, 2017), also in Malaysia, survival analysis and prognostic factors for colorectal cancer patients (Hassan *et al.*, 2016) while Ghazali (2018) modeled the survival time and incidence for colorectal cancer patients. In Thai, Kittrongsiri *et al.,* (2020) assess the overall and stage-specific colorectal cancer survival and identify the prognostic factors among the patients.

However, none of these works model colorectal cancer data using the cure fraction model. Hence, we model this data using the *WEENMCR* model and compared its performance with the fits of generalized Gompertz non-mixture cure rate (GGNMCR), modified Weibull non-mixture cure rate (MWNMCR) and generalized modified Weibull non-mixture cure rate (GMWNMCR) models.

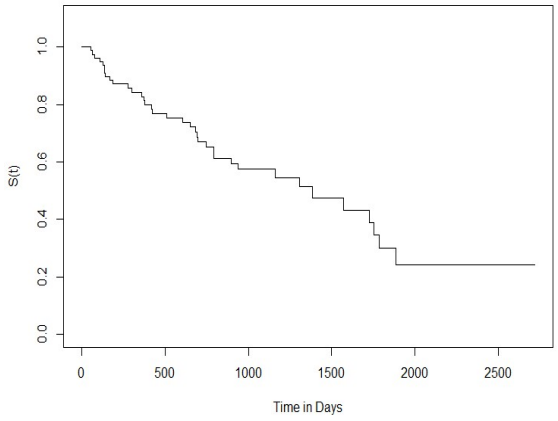


Figure 2: Kaplan-Meier survival curve for Colorectal Cancer patients

The Kaplan-Meier curve for the patients is given in Figure 2. From these graphs, we observed that after about 1800 days, the curve leveled off at a value close to 0.25. This suggests the presence of long-term survivors in the data set as mentioned earlier according to (Corbi`ere *et al.*, 2009). The statistical summaries of the fits of this data to the *WEENMCR*, *GGNMCR*, *MWNMCR,* and *GMWNMCR* models are given in Table 5. The information criteria: *AIC*, *BIC,* and *CAIC* values were also given in this table. The

information criteria values from this table showed that the *WEENMCR* model fits the data better than the *GGNMCR*, *MWNMCR,* and *GMWNMCR* models.

Table 5: Maximum likelihood estimates for the fits on colorectal cancer data

Model	parameters	estimate	SE	95% CI	AIC	BIC	CAIC
WEENMCR	a	1.8010	1.0136	(-0.1856, 3.7875)	622.7430	634.6531	613.5538
	β	0.0039	0.0005	(0.0029, 0.0050)			
	θ	0.8452	0.2440	(0.3669, 1.3236)			
	φ	0.0674	0.0155	(0.0370, 0.0978)			
	p	0.0206	0.0241	(-0.0266, 0.0678)			
GGNMCR	β	0.0007	0.0005	(-0.0011, 0.0017)	632.2858	641.8139	624.5525
	λ	1.9293	0.9631	(0.0417, 3.8170)			
	a	0.0003	0.0007	(-0.0002, 0.0018)			
	p	0.3011	0.1507	(0.0058, 0.5965)			
MWNM	β	0.5249	0.0601	(0.4071, 0.6428)	623.9012	633.4293	616.1679
	λ	0.1285	0.0969	(-0.0615, 0.3184)			
	a	0.0005	0.0002	(0.0002, 0.0009)			
	p	0.3368	0.1205	(0.1006, 0.5730)			
GMWNMCR	β	0.0014	0.0005	(0.0004, 0.0023)	624.4096	636.3197	614.8150
	λ	0.4685	0.1160	(0.2413, 0.6958)			
	a	0.0095	0.0036	(0.0023, 0.0166)			
	γ	0.2976	0.1353	(0.0325, 0.5627)			
	p	0.5351	0.0793	(0.3797, 0.6905)			

Conclusion

In medical applications, the presence of a cure fraction usually occurs in the data. To model such type of data, the cure rate models are used. In this article, a non-mixture cure rate model was introduced using the *WEE* distribution. The model contains the Weibull non-mixture cure rate, exponentiated exponential non-mixture cure rate, and exponential non-mixture cure rate models as special cases. The maximum likelihood estimation method was used to estimate the parameters of the model assuming right censoring. Simulation was conducted to evaluate the performance of the *MLE*s and it was found that on average, the method performs well. Furthermore, the applicability of the model was demonstrated using two real data sets. Results from the fits showed that the proposed *WEENMCR* model is better than the *WNMCR*, *EENMCR*, *ENMCR*, *GGNMCR*, *MWNMCR,* and *GMWNMCR* models.

Acknowledgement

Non-funded

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Achcar, J. A., Coelho-Barros, E. A., & Mazucheli, J. (2012). Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications*, *51*(1), 1 – 9.

Andrei, Y. Y., Asselain, B., et al. (1996). *Stochastic models of tumor latency and their biostatistical applications* (Vol. 1). World Scientific.

Berkson, J., & Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, *47*(259), 501–515.

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.

Cancho, V. G., & Bolfarine, H. (2001). Modeling the presence of immunes by using the exponentiated-Weibull model. *Journal of Applied Statistics*, *28*(6), 659–671.

Cancho, V. G., Rodrigues, J., & Castro, M. de. (2011). A flexible model for survival data with a cure rate: a Bayesian approach. *Journal of Applied Statistics*, *38*(1), 57–70.

Carrasco, J. M., Ortega, E. M., & Cordeiro, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis*, *53*(2), 450–462.

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.

Coelho-Barros, E. A., Achcar, J. A., & Mazucheli, J. (2017). Cure rate models considering the burr xii distribution in the presence of covariate. *Journal of Statistical Theory and Applications*, *16*(2), 150–164.

Corbi`ere, F., Commenges, D., Taylor, J. M., & Joly, P. (2009). A penalized likelihood approach for mixture cure models. *Statistics in medicine*, *28*(3), 510–524.

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, 1041–1046.

Farewell, V. T. (1986). Mixture models in survival analysis: Are they worth the risk? *Canadian Journal of Statistics*, *14*(3), 257–262.

Gamel, J. W., McLean, I. W., & Rosenberg, S. H. (1990). Proportion cured and mean log survival time as functions of tumor size. *Statistics in Medicine*, *9*(8), 999–1006.

Ghazali, A. K. (2018). *Modeling of survival and incidence for colorectal cancer in Malaysia*. Unpublished doctoral dissertation, Lancaster University.

Ghitany, M., & Maller, R. A. (1992). Asymptotic results for exponential mixture models with long-term survivors. *Statistics: A Journal of Theoretical and Applied Statistics*, *23*(4), 321–336.

Goldman, A. I. (1984). Survivorship analysis when cure is a possibility: a Monte Carlo study. *Statistics in Medicine*, *3*(2), 153–163.

Hassan, M. R. A., Suan, M. A. M., Soelar, S. A., Mohammed, N. S., Ismail, I., & Ahmad, F. (2016). Survival analysis and prognostic factors for colorectal cancer patients in Malaysia. *Asian Pacific Journal of Cancer Prevention*, *17*(7), 3575 – 3581.

Herring, A. H., & Ibrahim, J. G. (2002). Maximum likelihood estimation in random effects cure rate models with nonignorable missing covariates. *Biostatistics*, *3*(3), 387–405.

Ishaq, A., Usman, A., Tasiaˆu, M., Aliyu, Y., & Idris, F. (2017). A new Weibull Kumaraswamy distribution: Theory and applications. *Nigerian Journal of Scientific Research*, *16*(2), 158–166.

Jones, D., Powles, R., Machin, D., & Sylvester, R. (1981). On estimating the proportion of cured patients in clinical studies. *Biometrie-Praximetrie*, *21*, 1–11.

Kannan, N., Kundu, D., Nair, P., & Tripathi, R. C. (2010). The generalized exponential cure rate model with covariates. *Journal of Applied Statistics*, *37*(10), 1625 aˆ 1636.

Kittrongsiri, K., Wanitsuwan, W., Prechawittayakul, P., Sangroongruangsri, S., Cairns, J., & Chaikledkaew, U. (2020). Survival analysis of colorectal cancer patients in a Thai hospital-based cancer registry. *Expert review of gastroenterology & hepatology*, *14*(4), 291–300.

Kutal, D., & Qian, L. (2018). A non-mixture cure model for right-censored data with Fr´echet distribution. *Stats*, *1*(1), 176–188.

Liu, H., & Shen, Y. (2009). A semiparametric regression cure model for interval-censored data. *Journal of the American Statistical Association*, *104*(487), 1168 – 1178.

Lopes, C. M. C., & Bolfarine, H. (2012). Random effects in promotion time cure rate models. *Computational statistics & data analysis*, *56*(1), 75–87.

Magaji, B. A., Moy, F. M., Roslani, A. C., & Law, C. W. (2014). Descriptive epidemiology of colorectal cancer in University Malaya Medical Centre, 2001 to 2010. *Asian Pacific Journal of Cancer Prevention*, *15*(15), 6059–6064.

Magaji, B. A., Moy, F. M., Roslani, A. C., & Law, C. W. (2017). Survival rates and predictors of survival among colorectal cancer patients in a Malaysian tertiary hospital. *BMC cancer*, *17*(1), 1 – 8.

Maller, R. A., & Zhou, X. (1996). *Survival analysis with long-term survivors*. Wiley New York.

Martinez, E. Z., Achcar, J. A., Ja´come, A. A., & 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.

Mazucheli, J., Coelho-Barros, E. A., & 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.

Meeker, W. Q. (1987). Limited failure population life tests: application to integrated circuit reliability. *Technometrics*, *29*(1), 51–65.

Mudholkar, G. S., & Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, *42*(2), 299–302.

Naishadham, D., Lansdorp-Vogelaar, I., Siegel, R., Cokkinides, V., & Jemal, A. (2011). State disparities in colorectal cancer mortality patterns in the United States. *Cancer Epidemiology and Prevention Biomarkers*, *20*(7), 1296–1302.

Ng, S., & 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.

Ortega, E. M., Cordeiro, G. M., & Kattan, M. W. (2013). The log-beta Weibull regression model with application to predict recurrence of prostate cancer. *Statistical Papers*, *54*(1), 113–132.

Pal, M., Ali, M. M., & Woo, J. (2006). Exponentiated Weibull distribution. *Statistica*, *66*(2), 139–147.

Peng, Y., Dear, K. B., & Denham, J. (1998). A generalized f mixture model for cure rate estimation. *Statistics in medicine*, *17*(8), 813–830.

Peto, R., Lee, P., & Paige, W. (1972). Statistical analysis of the bioassay of continuous carcinogens. *British journal of cancer*, *26*(4), 258–261.

Ramos, P. L., Nascimento, D., & Louzada, F. (2017). The long term fr\'echet distribution: Estimation, properties and its application. *arXiv preprint arXiv:1709.07593*.

Salem, H. M., & Selim, M. (2014). The generalized Weibull-exponential distribution: properties and applications. *International Journal of Statistics and Applications*, *4*(2), 102–112.

Shao, Q., & Zhou, X. (2004). A new parametric model for survival data with long-term survivors. *Statistics in medicine*, *23*(22), 3525–3543.

Sy, J. P., & Taylor, J. M. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics*, *56*(1), 227–236.S

Tsodikov, A., Ibrahim, J., & Yakovlev, A. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, *98*(464), 1063–1078.

Uddin, M., Islam, M., & Ibrahim, Q. (2006). An analytical approach to cure rate estimation based on uncensored data. *Journal of Applied Sciences*, *6*(3), 548 –552.

Uddin, M., Sen, A., Noor, M., Islam, M., & Chowdhury, Z. (2006). An analytical approach on non-parametric estimation of cure rate based on uncensored data. *Journal of Applied Sciences*, *6*(6), 1258–1264.

Usman, U., Shamsuddeen, S., Arkilla, B. M., & Aliyu, Y. (2020). Inferences on the Weibull exponentiated exponential distribution and applications. *International Journal of Statistical Distributions and Applications*, *6*(1), 1-10.

Usman, U., Shamsuddeen, S., Arkilla, B. M., & Yakubu, A. (2022). Mixture cure model for right censored survival data with Weibull exponentiated exponential distribution. *Pakistan Journal of Statistics*, *38*(4).

Usman, U., Suleiman, S., Arkilla, B. M., & 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.

Yakovlev, A. Y., Asselain, B., Bardou, V., Fourquet, A., Hoang, T., Rochefediere, A., et al. (1993). A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer. *Biometrie et analyse de donnees spatiotemporelles*, *12*, 66–82.

Yakubu, A., & 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.