STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Statistics Opt. Inform. Comput., Vol. **10**, March 2022, pp 548–571.
Published online in International Academic Press (www.IAPress.org)

Exponentiated Weibull Models Applied to Medical Data in Presence of Right-censoring, Cure Fraction and Covariates

Edson Zangiacomi Martinez [1,*], Bruno Caparroz Lopes de Freitas [2], Jorge Alberto Achcar [1], Davi Casale Aragon [1], Marcos Vinicius de Oliveira Peres [1]

[1] *Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil*
[2] *State University of Maringá, Master Program in Biostatistics, Maringá, Brazil*

Abstract Cure fraction models have been widely used to analyze survival data in which a proportion of the individuals is not susceptible to the event of interest. This article considers frequentist and Bayesian methods to estimate the unknown model parameters of the exponentiated Weibull (EW) distribution considering right-censored survival data with a cure fraction and covariates. The EW distribution is as an extension to the Weibull distribution by considering an additional shape parameter to the model. We consider four types of cure fraction models: the mixture cure fraction (MCF), the non-mixture cure fraction (NMCF), the complementary promotion time cure (CPTC), and the cure rate proportional odds (CRPO) models. Bayesian inferences are obtained by using MCMC (Markov Chain Monte Carlo) methods. A simulation study was conducted to examine the performance of the maximum likelihood estimators for different sample sizes. Two real datasets were considered to illustrate the applicability of the proposed model. The EW distribution and its sub-models have the flexibility to accommodate different shapes for the hazard function and should be an attractive choice for survival data analysis when a cure fraction is present.

Keywords Weibull distribution, Survival analysis, Cure fraction, Censored data

AMS 2010 subject classifications 60E05, 62N02, 62P10

DOI:10.19139/soic-2310-5070-1266

1. Introduction

Survival analysis is a collection of statistical procedures used to investigate the time it takes for an event of interest to occur. We can find applications of survival analysis in many areas, including economics and finance, engineering, and medicine. Depending on the area of application, the variable of interest can be the time to failure of an electronic component, the survival time of patients with cancer, the length of patients' hospitalization in a healthcare unit, among many other possible examples. Some of the most popular tools of time-to-event analysis are the Kaplan-Meier method for estimating the survival function, the log-rank test for comparing two survival distributions, and the Cox proportional hazards regression for assessing the effects of covariates on the survival time [5, 19, 22]. If T is a random variable denoting the time-to-failure or failure time and t is an observation of T, the survival function, denoted by $S(t) = P(T \geq t)$, is the probability that the time to the event of interest is greater than some specified time t. While the Kaplan-Meier method gives us a non-parametric estimator of the survival function, alternatively we can obtain parametric estimates of $S(t)$ based on known probability distributions. Common choices are the Weibull distribution [57, 64], the Gompertz distribution [23], the log-normal distribution [2], and the Pareto distribution [31, 35].

*Correspondence to: Edson Zangiacomi Martinez (Email: edson@fmrp.usp.br). Ribeirão Preto Medical School, University of São Paulo, Av. Bandeirantes 3900, Vila Monte Alegre, Ribeirão Preto, SP, 14049-900, Brazil.

A random variable T follows a Weibull distribution if its cumulative distribution function $F(t)$ can be stated as

$$F(t) = P(T < t) = 1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right],$$

where $t > 0$, $\gamma > 0$ is the shape parameter and $\beta > 0$ is the scale parameter of the distribution. The corresponding survival function is given by

$$S(t) = 1 - F(t) = \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right], \tag{1}$$

the probability density function (*pdf*) is given by

$$f(t) = \frac{\gamma}{\beta}\left(\frac{t}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]$$

and the hazard function $h(t)$ is given by

$$h(t) = \frac{f(t)}{S(t)} = \frac{\gamma}{\beta}\left(\frac{t}{\beta}\right)^{\gamma-1}. \tag{2}$$

The hazard function is known as the instantaneous rate of occurrence of a time-related event. In the case of the Weibull distribution, a value of $\gamma = 1$ indicates that $h(t)$ is constant over time. In this situation, the Weibull distribution reduces to an exponential distribution. In addition, the Weibull distribution has a monotone increasing hazard for $\gamma > 1$ and a monotone decreasing hazard for $0 < \gamma < 1$ [73]. Classical methods used to estimate the parameters of the Weibull are discussed by [72]. In a Bayesian framework, Ramos et al. [62] proposed the necessary and sufficient conditions to verify when improper priors lead to proper posteriors for the parameters of the Weibull distribution in the presence of complete or right-censored data. Although the Weibull distribution is widely used in a variety of applications, in some situations we may have data with non-monotone hazard rates, which may require the use of models based on more complex distributions. Useful distributions with three or more parameters are given by the generalized modified Weibull distribution [9], the beta modified Weibull distribution [69], and the Kumaraswamy modified Weibull distribution [16], among many other generalizations and extensions of the Weibull distribution [4, 6, 52, 61].

The three-parameter exponentiated Weibull (EW) distribution was introduced by Mudholkar and Srivastava [48] as an extension of the Weibull distribution. The EW distribution can accommodate both decreasing and increasing failure rates, as well as unimodal and bathtub shaped failure rates, and this can be very useful in some applications [47, 54].

In the present article, we consider frequentist and Bayesian methods to estimate the unknown model parameters of the EW distribution considering right-censored survival data with a cure fraction and covariates. This article is organized as follows. In Section 2, we present the EW distribution and the estimation of its parameters. Under a Bayesian approach, the parameter estimation is based on Markov Chain Monte Carlo (MCMC) methods. Section 3 presents a simulation study illustrating the performance of the frequentist approach when applied to samples from an EW distribution with a cure fraction. Section 3 also describes two applications of the model based on the EW distribution to real data. Some concluding remarks are provided in Section 4.

2. Methods

2.1. The exponentiated Weibull (EW) distribution

Let us assume that T is a continuous nonnegative random variable with probability density function (*pdf*) $f(t)$ and cumulative distribution function $F(t)$. The random variable T follows an exponentiated Weibull (EW) distribution if its survival function is given by

$$S(t) = 1 - F(t) = 1 - \left\{1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}, \tag{3}$$

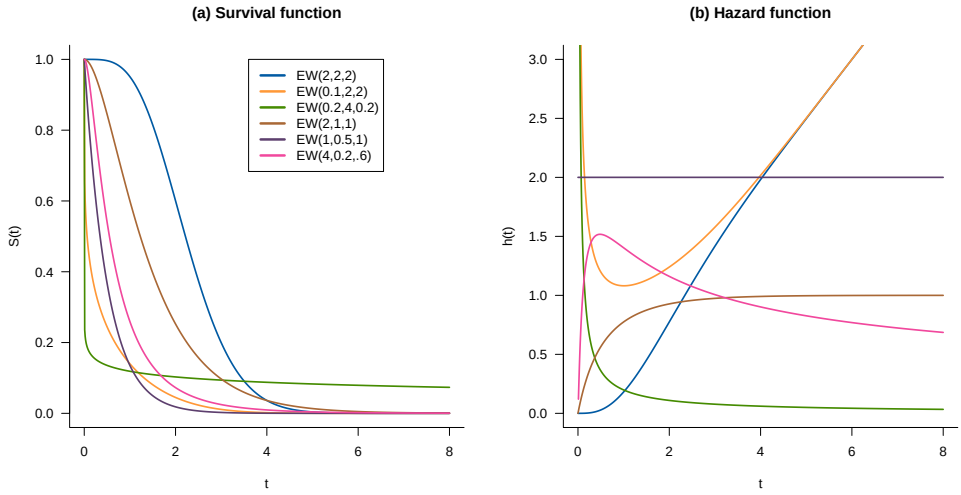

Figure 1. Survival and hazard functions of the EW distribution for different values of α, β and γ.

where $t > 0$, $\alpha > 0$, $\beta > 0$ and $\gamma > 0$. In (3), α and γ are two shape parameters and β is a scale parameter. Hereafter, we write $EW(\alpha, \beta, \gamma)$ to denote an EW distribution with survival function given in (3). The corresponding *pdf* is given by

$$f(t) = \frac{\alpha\gamma}{\beta}\left(\frac{t}{\beta}\right)^{\gamma-1}\left\{1-\exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1}\exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right], \tag{4}$$

and the hazard function $h(t)$ is given by

$$h(t) = \frac{\alpha\gamma}{\beta}\frac{\left(\frac{t}{\beta}\right)^{\gamma-1}\left\{1-\exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1}\exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]}{1-\left\{1-\exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}}. \tag{5}$$

We can note that the EW distribution reduces to a standard Weibull (SW) distribution when $\alpha = 1$. In addition, the EW distribution covers the exponentiated exponential (EE) distribution as a special case when $\gamma = 1$ [29], and the exponentiated Rayleigh (ER) distribution when $\gamma = 2$ [71]. Mudholkar et al. [49] showed that for the EW family, the hazard function (5) is

(a) monotone increasing for $\gamma \geq 1$ and $\alpha\gamma \geq 1$,

(b) monotone decreasing for $\gamma \leq 1$ and $\alpha\gamma \leq 1$,

(c) bathtub-shaped for $\gamma > 1$ and $\alpha\gamma < 1$,

(d) unimodal for $\gamma < 1$ and $\alpha\gamma > 1$, and

(e) constant for $\alpha = \gamma = 1$.

The graphs in Figure 1 illustrate the survival and hazard functions of the EW family for different values of α, β and γ. Panel (b) of Figure 1 shows that the EW distribution and its sub-models have the flexibility to accommodate different shapes for the hazard function and should be an attractive choice for survival data analysis.

Mudholkar and Hutson [47] showed that the r-th moment of a random variable that follows an EW distribution is given by

$$\mu_r = E(T^r) = \alpha\beta^r \int_0^{\infty} t^{r/\gamma} e^{-t}(1-e^{-t})^{\alpha-1} dt,$$

for $r = 1, 2, 3, ...$ For positive integer values of α, the expression for μ_r takes the closed form given by

$$\mu_r = \alpha\beta^r\Gamma\left(\frac{r}{\gamma}+1\right)\left[1+\sum_{j=1}^{\alpha-1}(-1)^j\binom{\alpha-1}{j}\frac{1}{(j+1)^{r/\gamma+1}}\right],$$

for $r = 1, 2, 3, ...$ Other statistical properties of the EW distribution are presented by [50].

2.2. *Maximum likelihood method*

In this subsection, we obtain the maximum likelihood (ML) estimators of the EW distribution parameters. Let $t_1, t_2, ..., t_n$ be observed values of a random sample from EW distribution with parameters (α, β, γ). Based on (4), the likelihood function can be written as

$$L(\theta) = \prod_{i=1}^{n}\frac{\alpha\gamma}{\beta}\left(\frac{t_i}{\beta}\right)^{\gamma-1}\left\{1-\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1}\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right],$$

where $\theta = (\alpha, \beta, \gamma)$ and n is the sample size. The log-likelihood function $\ell(\theta) = \log L(\theta)$ is given by

$$\begin{aligned}\ell(\theta) &= n\log(\alpha)+n\log(\gamma)-n\log(\beta)+(\gamma-1)\sum_{i=1}^{n}\log\left(\frac{t_i}{\beta}\right)\\ &\quad+(\alpha-1)\sum_{i=1}^{n}\log\left\{1-\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\}-\sum_{i=1}^{n}\left(\frac{t_i}{\beta}\right)^{\gamma}.\end{aligned}$$

The first derivatives of $\ell(\theta)$ with respect to α, β and γ, respectively, are given by

$$\frac{\partial\ell(\theta)}{\partial\alpha} = \frac{n}{\alpha}+\sum_{i=1}^{n}\log\left\{1-\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\},$$

$$\frac{\partial\ell(\theta)}{\partial\beta} = \frac{\gamma}{\beta}\left\{-n-\frac{\alpha-1}{\beta^{\gamma}}\sum_{i=1}^{n}\frac{t_i^{\gamma}\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]}{1-\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]}+\sum_{i=1}^{n}\left(\frac{t_i}{\beta}\right)^{\gamma}\right\}$$

and

$$\frac{\partial\ell(\theta)}{\partial\gamma} = \frac{n}{\gamma}+\sum_{i=1}^{n}\log\left(\frac{t_i}{\beta}\right)+\frac{1}{\beta^{\gamma}}\left\{(\alpha-1)\sum_{i=1}^{n}\frac{t_i^{\gamma}\log\left(\frac{t_i}{\beta}\right)\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]}{1-\exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]}-\sum_{i=1}^{n}t_i^{\gamma}\log\left(\frac{t_i}{\beta}\right)\right\}.$$

The ML estimators of the unknown parameters α, β and γ are obtained by solving the following equations simultaneously: $\partial\ell(\theta)/\partial\alpha = 0$, $\partial\ell(\theta)/\partial\beta = 0$ and $\partial\ell(\theta)/\partial\gamma = 0$. It can be seen that these equations do not have closed form solutions, therefore an iterative numerical procedure such as the Newton-Raphson method can be used to obtain the estimates. In this article, we use the maxLik package in the statistical software R, which numerically maximizes the likelihood function [30]. Considering large sample sizes, the ML estimators are asymptotically normal with mean equal to the true parameter values and variance-covariance matrix equal to the inverse of the observed Fisher information matrix, defined to be the matrix of the second partial derivatives of the negative log-likelihood with respect to the model parameters calculated at their respective ML estimators. The asymptotic variances of the ML estimators are thus given by the diagonal elements of the matrix I_0^{-1}, where I_0 is the observed Fisher information matrix. Therefore,

$$I_0^{-1} = -\begin{bmatrix}\frac{\partial^2\ell(\theta)}{\partial\alpha^2} & \frac{\partial^2\ell(\theta)}{\partial\alpha\partial\beta} & \frac{\partial^2\ell(\theta)}{\partial\alpha\partial\gamma}\\ \frac{\partial^2\ell(\theta)}{\partial\beta\partial\alpha} & \frac{\partial^2\ell(\theta)}{\partial\beta^2} & \frac{\partial^2\ell(\theta)}{\partial\beta\partial\gamma}\\ \frac{\partial^2\ell(\theta)}{\partial\gamma\partial\alpha} & \frac{\partial^2\ell(\theta)}{\partial\gamma\partial\beta} & \frac{\partial^2\ell(\theta)}{\partial\gamma^2}\end{bmatrix}^{-1} = \begin{bmatrix}Var(\widehat{\alpha}_{ML}) & cov(\widehat{\alpha}_{ML},\widehat{\beta}_{ML}) & cov(\widehat{\alpha}_{ML},\widehat{\gamma}_{ML})\\ cov(\widehat{\alpha}_{ML},\widehat{\beta}_{ML}) & Var(\widehat{\beta}_{ML}) & cov(\widehat{\beta}_{ML},\widehat{\gamma}_{ML})\\ cov(\widehat{\alpha}_{ML},\widehat{\gamma}_{ML}) & cov(\widehat{\beta}_{ML},\widehat{\gamma}_{ML}) & Var(\widehat{\gamma}_{ML})\end{bmatrix},$$

where $\widehat{\alpha}_{ML}$, $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ are the ML estimators of the parameters α, β and γ, respectively. Based on the asymptotic normality of the ML estimators, approximate $(1-\delta)100\%$ Wald-type confidence intervals for the parameters α, β and γ are given by

$$\widehat{\alpha}_{ML} \mp z_{\delta/2}\sqrt{Var(\widehat{\alpha}_{ML})}, \quad \widehat{\beta}_{ML} \mp z_{\delta/2}\sqrt{Var(\widehat{\beta}_{ML})} \quad \text{and} \quad \widehat{\gamma}_{ML} \mp z_{\delta/2}\sqrt{Var(\widehat{\gamma}_{ML})}, \tag{6}$$

respectively, where $z_{\delta/2}$ is the $(\delta/2)\,th$ percentile of the standard normal distribution.

However, the confidence intervals in (6) may lead to negative lower bounds sometimes. To overcome this problem, we can use the logarithmic transformation and the delta method [51] to get the asymptotic normal distributions for $\log(\widehat{\alpha}_{ML})$, $\log(\widehat{\beta}_{ML})$ and $\log(\widehat{\gamma}_{ML})$ given by

$$\sqrt{n}\left[\log(\widehat{\alpha}_{ML}) - \log(\alpha)\right] \xrightarrow{D} N\left(0, \frac{\sigma_\alpha^2}{\alpha^2}\right),$$

$$\sqrt{n}\left[\log(\widehat{\beta}_{ML}) - \log(\beta)\right] \xrightarrow{D} N\left(0, \frac{\sigma_\beta^2}{\beta^2}\right)$$

and

$$\sqrt{n}\left[\log(\widehat{\gamma}_{ML}) - \log(\gamma)\right] \xrightarrow{D} N\left(0, \frac{\sigma_\gamma^2}{\gamma^2}\right),$$

respectively, where $\xrightarrow{D}$ denotes convergence in distribution, $\sigma_\alpha^2 = nVar(\widehat{\alpha}_{ML})$, $\sigma_\beta^2 = nVar(\widehat{\beta}_{ML})$, and $\sigma_\gamma^2 = nVar(\widehat{\gamma}_{ML})$. Asymptotic $(1-\delta)100\%$ confidence intervals for $\log(\alpha)$, $\log(\beta)$ and $\log(\gamma)$ are respectively given by

$$\log(\widehat{\alpha}_{ML}) \mp z_{\delta/2}\frac{\sqrt{Var(\widehat{\alpha}_{ML})}}{\widehat{\alpha}_{ML}} \equiv (L_\alpha, U_\alpha),$$

$$\log(\widehat{\beta}_{ML}) \mp z_{\delta/2}\frac{\sqrt{Var(\widehat{\beta}_{ML})}}{\widehat{\beta}_{ML}} \equiv (L_\beta, U_\beta)$$

and

$$\log(\widehat{\gamma}_{ML}) \mp z_{\delta/2}\frac{\sqrt{Var(\widehat{\gamma}_{ML})}}{\widehat{\gamma}_{ML}} \equiv (L_\gamma, U_\gamma). \tag{7}$$

Using the inverse logarithmic transformation, the approximate confidence intervals for α, β and γ are respectively obtained as $\left(e^{L_\alpha}, e^{U_\alpha}\right)$, $\left(e^{L_\beta}, e^{U_\beta}\right)$ and $\left(e^{L_\gamma}, e^{U_\gamma}\right)$.

In this article, we use the Akaike information criterion (AIC) to compare the fit of different models to a particular data set [10]. The AIC was proposed by Akaike [3] as an asymptotic approximate estimator of the Kullback-Leibler information between the model generating the data and a candidate model fitted by the ML method. The AIC value is given by $AIC = 2\log(L) + 2k$, where L is the maximum value of the likelihood function for each candidate model, and k is the number of estimated parameters of a candidate distribution function. The AIC allows for a direct comparison between models, with lower values indicating a better model fit.

The TTT-plot can be used to examine trends of the hazard function $h(t)$ before we fit a model to the data and then assist us to choose the best model [36]. For uncensored data, the TTT statistic is given by $TTT_i^0 = \sum_{k=1}^{i}(n-k+1)(T_k^* - T_{k-1}^*)$, where T^* is the time-to-event variable T in ascending sorted order and T_0 is defined to be zero. The scaled TTT is then defined as $TTT_i = TTT_i^0/TTT_n^0$ and it is ranged from 0 to 1. The TTT plot is then a plot of TTT_i against i/n. For an exponential distribution (with $h(t)$ constant), the TTT plot is close to the diagonal from (0,0) to (1,1). The TTT plot suggests an increasing $h(t)$ when it concaves downward, and a decreasing $h(t)$ when it concaves upward. An upside down "S" shaped curve suggests an unimodal $h(t)$ function and a "S" shaped curve suggests a bathtub-shaped $h(t)$ function. In these last two cases, a model based on the standard Weibull distribution may not be appropriate for the data, since its hazard function (2) does not have these shapes. These situations can justify the use of models based on more complex distributions.

2.3. *Maximum-likelihood estimation in presence of censored data*

Censoring occurs when we have some information on the individual survival time, but we do not know the exact survival time [70]. Among different types of censoring mechanisms, right censoring is considered the most common. Right-censored data refers to observations for which the corresponding individuals have not yet experienced the event of interest at the end of the follow-up period. Given a sample of n individuals, the contribution of the ith subject for the likelihood function under a right-censoring mechanism is given by

$$L_i = [f(t_i)]^{d_i} [S(t_i)]^{1-d_i},$$

where $d_i = 1$ denotes a complete observation and $d_i = 0$ denotes a censored observation $(i = 1, ..., n)$ [37]. Assuming a model based on the EW distribution and from (4) and (3), the likelihood function is given by

$$L(\theta) = \prod_{i=1}^{n} \left[\frac{\alpha\gamma}{\beta} \left(\frac{t_i}{\beta}\right)^{\gamma-1} \left\{1 - \exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1} \exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right] \right]^{d_i} \left[1 - \left\{1 - \exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}\right]^{1-d_i},$$

where $\theta = (\alpha, \beta, \gamma)$, and the log-likelihood function $\ell(\theta)$ is given by

$$\begin{aligned} \ell(\theta) \;=\; & \log(\alpha) \sum_{i=1}^{n} d_i + \log(\gamma) \sum_{i=1}^{n} d_i - \log(\beta) \sum_{i=1}^{n} d_i + (\gamma - 1) \sum_{i=1}^{n} d_i \log\left(\frac{t_i}{\beta}\right) \\ & + (\alpha - 1) \sum_{i=1}^{n} d_i \log\left\{1 - \exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\} - \sum_{i=1}^{n} d_i \left(\frac{t_i}{\beta}\right)^{\gamma} \\ & + \sum_{i=1}^{n} (1 - d_i) \log\left[1 - \left\{1 - \exp\left[-\left(\frac{t_i}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}\right]. \end{aligned}$$

To find the ML estimators, we derive the log-likelihood function $\ell(\theta)$ with respect to the three parameters (α, β and γ) and we equate these derivatives to zero. As expected, the ML estimators cannot be obtained in closed form, and numerical procedures are required. Asymptotic $(1 - \delta)100\%$ confidence intervals for α, β and γ can be obtained in a way analogous to that presented in the previous subsection.

2.4. *Cure fraction models*

Cure models are defined by Lambert [38] as "a special type of survival analysis model where it is assumed that there are a proportion of subjects who will never experience the event of interest and thus the survival curve will eventually reach a plateau". The standard survival analysis techniques assume that all subjects have the same susceptibility to the event of interest, and at some point, everyone will experience it within a sufficiently long period of time. Mathematically speaking, this implies that $S(t)$ approaches zero as t goes to infinity. However, this assumption may not be true in many practical situations. For example, in randomized clinical trials of cancer treatments, we can have a fraction of patients who will be cured and consequently will not die due to cancer. Several articles were published in medical journals that use cure fraction models in their data analysis (see, for example, Smoll et al. [68]; Jácome et al. [33]; Meshkat et al. [46]; Looha et al. [41]; Omer et al. [53]; Ricci et al. [63]). A simple way to identify the presence of cured or immune patients in a dataset is to visualize the shape of the Kaplan-Meier curve corresponding to the time-to-event variable. If the survival curve shows a plateau at the end, then a cure model may be appropriate to these data. However, expert opinion is also important to state that a model with a cure fraction is suitable for a given data set. For example, when analysing data from a clinical trial, it is essential to use medical arguments to justify whether a cure for the disease under study can actually occur. Mixture and non-mixture cure fraction models are the two most common forms of cure fraction models [21, 42, 56], but we can find some alternatives in the literature, as we will discuss below.

2.4.1. The mixture cure fraction (MCF) model The mixture cure fraction (MCF) model introduced by Boag [7] assumes that the population of interest is divided into two groups [42]. The first group is composed of "cured"

or "immune" individuals with probability η, and the second group is composed of susceptible individuals with probability $1-\eta$. In this model, the survival function is given by

$$S(t) = \eta + (1-\eta)S_0(t),$$

where η is a parameter which represents the proportion of "long-term survivors" or "cured patients", regarding the event of interest ($0 < \eta < 1$), and $S_0(t)$ is the baseline survival function for the susceptible individuals [45]. We can note that if $\lim_{t\longrightarrow\infty} S_0(t) = 0$, thus $S(t)$ approach η as t goes to infinity. The corresponding *pdf* for the lifetime T is

$$f(t) = \frac{dF(t)}{dt} = (1-\eta)f_0(t),$$

where $f_0(t)$ is the baseline *pdf* for the susceptible individuals and $F(t) = 1 - S(t)$. The MCF model is also called Bernoulli cure rate model by some authors such as Davies et al. [18]. An MCF model where $S_0(t)$ corresponds to the survival function of a standard Weibull distribution (1) was described by Achcar et al. [1]. Considering a random sample (t_i, d_i) of size n, $i = 1, \ldots, n$, the contribution of the ith subject for the likelihood function is given by

$$L_i = [f(t_i)]^{d_i} [S(t_i)]^{1-d_i} = [(1-\eta)f_0(t_i)]^{d_i} [\eta + (1-\eta)S_0(t_i)]^{1-d_i},$$

where d_i is the censoring indicator variable. The likelihood function for the parameters of the MCF model based on the EW distribution is given by

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} \left[(1-\eta)\frac{\alpha\gamma}{\beta}\left(\frac{t}{\beta}\right)^{\gamma-1} \left\{1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1} \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right]^{d_i} \\ &\quad \times \left[\eta + (1-\eta)\left[1 - \left\{1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}\right]\right]^{1-d_i}, \end{aligned}$$

where $\theta = (\alpha, \beta, \gamma, \eta)$. On differentiating the log-likelihood function $\ell(\theta) = \log L(\theta)$ with respect to α, β, γ and η and equating to zero, we obtain the estimating equations. These non-linear equations do not have an explicit solution and they are solved by means of numerical methods.

2.4.2. The non-mixture cure fraction (NMCF) model The non-mixture cure fraction (NMCF) model defines an asymptote for the cumulative hazard and hence for the cure fraction [75]. In this case, the survival function is given by

$$S(t) = \eta^{F_0(t)} = \exp\left[\log(\eta) F_0(t)\right],$$

where $F_0(t) = 1 - S_0(t)$ is the proper baseline cumulative probability function for the susceptible individuals and $0 < \eta < 1$ represents the cure fraction. The corresponding *pdf* for the lifetime T is

$$f(t) = \frac{dF(t)}{dt} = -\log(\eta) f_0(t_i) \exp\left[\log(\eta) F_0(t)\right],$$

where $F(t) = 1 - S(t)$, and the hazard function $h(t)$ is given by

$$h(t) = \frac{f(t)}{S(t)} = -\log(\eta) f_0(t_i).$$

Given a sample of n individuals, the contribution of the ith subject for the likelihood function is thus given by

$$L_i = [h(t_i)]^{d_i} S(t_i) = [-\log(\eta) f_0(t_i)]^{d_i} \exp\left[\log(\eta) F_0(t_i)\right],$$

$i = 1, ..., n$, where d_i is the censoring indicator variable. The likelihood function for the parameters of the NMCF model based on the EW distribution is given by

$$L(\theta) = \prod_{i=1}^{n} \left[-\log(\eta) \frac{\alpha\gamma}{\beta} \left(\frac{t}{\beta}\right)^{\gamma-1} \left\{1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha-1} \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right] \right]^{d_i} \times \exp\left[\log(\eta) \left\{1 - \exp\left[-\left(\frac{t}{\beta}\right)^{\gamma}\right]\right\}^{\alpha}\right].$$

As in the previous cases, ML parameter estimates are obtained by first differentiating the log-likelihood function $\ell(\theta) = \log L(\theta)$ with respect to the parameters and then by equating those derivatives to zero. Numerical methods should be used to solve the estimating equations. Lambert [38] claims that one of the advantages of the NMCF model is that it has a proportional hazards model as a particular case. The identifiability of the MCF and NMCF models is discussed by Li et al. [40].

2.4.3. Alternatives to the mixture and non-mixture cure fraction models Some alternatives to the mixture and non-mixture cure fraction models are the complementary promotion time cure (CPTC) model [79] and the cure rate proportional odds (CRPO) model studied by Gu et al. [27]. The promotion time cure (PTC) model proposed by and Tsodikov [78] and Chen et al. [11] is analogous to the NMCF model described in the previous subsection. The survival function $S(t)$ of the CPTC model is given by

$$S(t) = 1 + \eta - \eta^{S_0(t)},$$

and the correspondent *pdf* is given by

$$f(t) = -\log(\eta)\, f_0(t) \eta^{S_0(t)},$$

where $f_0(t)$ and $S_0(t)$ are the baseline *pdf* and the baseline survival function for the susceptible individuals, respectively, and $0 < \eta < 1$ represents the cure fraction.

In the CRPO model, the survival function is given by

$$S(t) = \frac{1}{1 + \left(\frac{1}{\eta} - 1\right) F_0(t)}, \tag{8}$$

where $F_0(t) = 1 - S_0(t)$. The *pdf* and the hazard function corresponding to (8) are given, respectively, by

$$f(t) = \frac{\left(\frac{1}{\eta} - 1\right) f_0(t)}{\left[1 + \left(\frac{1}{\eta} - 1\right) F_0(t)\right]^2} \quad \text{and} \quad h(t) = \frac{\left(\frac{1}{\eta} - 1\right) f_0(t)}{1 + \left(\frac{1}{\eta} - 1\right) F_0(t)}.$$

Under the CRPO model, the contribution of the ith subject for the likelihood function is thus given by

$$L_i = [h(t_i)]^{d_i} S(t_i) = \frac{\left[\left(\frac{1}{\eta} - 1\right) f_0(t)\right]^{d_i}}{\left[1 + \left(\frac{1}{\eta} - 1\right) F_0(t)\right]^{1+d_i}},$$

where d_i is the censoring indicator variable ($i = 1, ..., n$).

We can note that the cure fraction of each of these models is given by $\eta = \lim_{t \longrightarrow \infty} S(t)$. Models based on defective distributions are also alternatives to the mixture and non-mixture models, but they will not be considered in the present paper. Interest readers can find details in Rocha et al. [65], Martinez and Achcar [44] and Scudilio et al. [66].

2.5. *Simulating samples from an EW distribution with a cure fraction*

Based on the Algorithms 6 and 7 presented in the article from Ramos et al. [60], we can simulate a sample of size n from the EW distribution with right-censoring and a cure fraction following the steps:

Step 1. Fix values of α, β, γ and η.

Step 2. Generate n random samples from $u_i \sim Uniform(0,1)$.

Step 3. Consider $t'_{0i} = F^{-1}(u_i, \alpha, \beta, \gamma)$, where

$$F^{-1}(u_i, \alpha, \beta, \gamma) = \beta \left[-\log\left(1 - u_i^{1/\alpha}\right)\right]^{1/\gamma}.$$

Step 4. Generate n random samples from a Bernoulli distribution with parameter η, denoted by $b_i \sim Bernoulli(\eta)$.

Step 5. If $b_i = 1$, then $t'_i = \infty$. Otherwise, $t'_i = t'_{0i}$.

Step 6. Generate n random samples from $c_i \sim Uniform\left(0, \max\left(t'_{0i}\right)\right)$. This variable controls the censorship mechanism.

Step 7. Pairs of values (t_i, d_i), $i = 1, ..., n$, are thus obtained, where d_i is the censoring indicator variable. If $t'_i \leq c_i$, then $t_i = t'_i$ and $d_i = 1$. Otherwise, $t_i = c_i$ and $d_i = 0$ (the censored observations).

2.6. *Model with cure fraction and covariates*

Let t_i be an observation of a nonnegative random variable T denoting the time-to-event for the ith individual, and let $x_i = (x_{i1}, ..., x_{iJ})$ be the observed values of a covariate vector, for $i = 1, ..., n$. For the ith individual, the baseline survival function of a model based on the EW distribution with covariates is given by

$$S_0(t_i|x_i, \theta_0) = 1 - \left\{1 - \exp\left[-\left(\frac{t_i}{\beta(x_i)}\right)^{\gamma(x_i)}\right]\right\}^{\alpha(x_i)}, \tag{9}$$

where

$$\log\alpha(x_i) = \alpha_0 + \sum_{j=1}^{J}\alpha_j x_{ij}, \quad \log\beta(x_i) = \beta_0 + \sum_{j=1}^{J}\beta_j x_{ij}, \quad \log\gamma(x_i) = \gamma_0 + \sum_{j=1}^{J}\gamma_j x_{ij},$$

and $\theta_0 = (\alpha_0, \alpha_1, ..., \alpha_J, \beta_0, \beta_1, ..., \beta_J, \gamma_0, \gamma_1, ..., \gamma_J)$. According to the complexity of the study, we can also assume different covariate vectors for each parameter of the EW distribution. The MCF model with covariates is thus given by

$$S(t_i|x_i, \theta) = \eta(x_i) + [1 - \eta(x_i)]\, S_0(t_i|x_i, \theta_0),$$

where

$$\eta(x_i) = \frac{\exp\left(\eta_0 + \sum_{j=1}^{J}\eta_j x_{ij}\right)}{1 + \exp\left(\eta_0 + \sum_{j=1}^{J}\eta_j x_{ij}\right)}, \tag{10}$$

and θ is the complete vector of parameters given by $\theta = (\alpha_0, \alpha_1, ..., \alpha_J, \beta_0, \beta_1, ..., \beta_J, \gamma_0, \gamma_1, ..., \gamma_J, \eta_0, \eta_1, ..., \eta_J)$ [77]. In this case, the contribution of the ith subject for the likelihood function is given by

$$L_i = \{[1 - \eta(x_i)]\, f_0(t_i|x_i, \theta_0)\}^{d_i} \{\eta(x_i) + [1 - \eta(x_i)]\, S_0(t_i|x_i, \theta_0)\}^{1-d_i},$$

where $f_0(t_i|x_i, \theta_0)$ is the correspondent baseline *pdf* for the susceptible individuals. Similarly, we can specify models with covariates assuming the other forms of cure fraction described in subsection 2.4.

Cox-Snell residuals can be used to graphically test the fit of the various models to the data. The Cox-Snell residuals r_i corresponding to time t_i, are given by $r_i = -\log S(t|x, \theta)$, where θ is the complete vector of parameters

and $\widehat{\theta}$ is the corresponding vector estimate parameter [17]. The estimated Cox-Snell residuals under the MCF model are thus given by

$$\widehat{r}_i = -\log\left\{\widehat{\eta}(x_i) + [1-\widehat{\eta}(x_i)]\, S_0(t_i|x_i,\widehat{\theta}_0)\right\}.$$

Diagnostics based on Cox-Snell residuals are usually based on fitting a Kaplan-Meier curve to r_i and comparing it to that of the standard exponential distribution. However, Peng and Taylor [55] warn that under the MCF model, the Cox-Snell residuals are not a sample from an unit exponential distribution when the cure model is correctly specified, given that $P\left[-\log S(t|x,\theta) > t\right] = P\left[S(t|x,\theta) \leq e^{-t}\right]$ is equal to e^{-t} if $t < -\log\eta(x)$ and 0 otherwise. Despite this, Peng and Taylor [55] argue that "the usual approach of checking the estimated cumulative hazard rate of the (r_i, d_i)'s against the cumulative hazard function of the unit exponential distribution is still appropriate since the residuals that are equal to $-\log\eta(x)$ are always censored".

2.7. Bayesian estimation

Under a Bayesian framework, the joint posterior distribution for the model parameters is obtained by combining the joint prior distribution of the parameters and the likelihood function [14]. The model based on the EW distribution without the presence of covariates but including a cure fraction has four parameters, α, β, γ and η. Since $\alpha > 0$, $\beta > 0$, $\gamma > 0$ and $0 < \eta < 1$, we assume independent prior distributions $\alpha \sim Gamma(a_\alpha, b_\alpha)$, $\beta \sim Gamma(a_\beta, b_\beta)$, $\gamma \sim Gamma(a_\gamma, b_\gamma)$, and $\eta \sim Beta(a_\eta, b_\eta)$, where a_α, b_α, a_β, b_β, a_γ, b_γ, a_η and b_η are known hyperparameters; $Gamma(a,b)$ denotes a gamma distribution with mean a/b and variance a/b^2, and $Beta(a,b)$ denotes a beta distribution with mean $a/(a+b)$ and variance $ab/[(a+b)^2(a+b+1)]$ [28].

Considering the model in the presence of cure fraction and covariates introduced in the Subsection 2.6, we assume independent prior distributions $\alpha_j \sim Normal(a_{\alpha_j}, b^2_{\alpha_j})$, $\beta j \sim Normal(a_{\beta_j}, b^2_{\beta_j})$, $\gamma_j \sim Normal(a_{\gamma_j}, b^2_{\gamma_j})$, and $\eta_j \sim Normal(a_{\eta_j}, b^2_{\eta_j})$, $j = 1, ..., J$, where a_{α_j}, $b^2_{\alpha_j}$, a_{β_j}, $b^2_{\beta_j}$, a_{γ_j}, $b^2_{\gamma_j}$, a_{η_j}, and $b^2_{\eta_j}$ are known hyperparameters, J is the number of covariates, and $Normal(a, b^2)$ denotes a normal distribution with mean a and variance b^2.

To simulate samples from the joint posterior distribution, we consider the use of MCMC (Markov Chain Monte Carlo) algorithms implemented in the package MCMCpack of the R software [43]. Interested readers can refer to Chib and Greenberg [13] for a review of standard MCMC methods. We generated 1,005,000 samples for each parameter of interest. The 5,000 first simulated samples were discarded as a burn-in period, which is usually used to minimize the effect of the initial values. The Bayes estimates of the parameters can be obtained as the mean of samples drawn from the joint posterior distribution (use of squared error loss function). The 95% highest posterior density (HPD) intervals were obtained from the simulated posterior distributions for the parameters. HPD intervals are the shortest possible intervals for any given coverage probability [12]. Convergence was assessed visually from traceplots of each MCMC chain and quantitatively via Geweke criterion, which is based on a test for equality of the means of the first and last part of a Markov chain [26]. The Geweke criterion involves calculating a Z-score from the difference of means between the first 10% and last 50% of the sampled chain divided by their pooled standard deviation. If the MCMC samples are drawn from a stationary distribution, this Z-score has an asymptotically standard normal distribution. We calculated Geweke's convergence diagnostics using the geweke.diag function in the R package coda [58].

The logarithm of the pseudo-marginal likelihood (LPML) is a well-known Bayesian criterion for model comparison [25]. This comparison criterion is based on the conditional predictive ordinate (CPO) statistics [24]. In a general manner, let $\mathbf{y} = \{y_1, \cdots, y_n\}$ be an observed sample from $f(\cdot|\theta)$. CPO estimates the probability of observing y_i in the future after having already observed y_{-1}, the vector of all observations except the ith observation. The CPO for the ith observation is given by

$$CPO_i = f(y_i|y_{-1}) = \int f(y_i|\theta) f(\theta|y_{-1}) d\theta,$$

where y_i is each observation, θ is the parameter vector of interest, $f(y_i|\theta)$ is the probability density function of the random variable Y and $f(\theta|y_{-1})$ is the joint posterior distribution of θ given y_{-1}. The CPO is easily calculated using MCMC methods. By considering the inverse likelihood across M iterations, the CPO for each individual i is

given by

$$\widehat{CPO}_i = \left\{ \frac{1}{M} \sum_{i=1}^{M} \frac{1}{f(y_i|\theta_m)} \right\}^{-1}$$

[20]. The logarithm of the pseudo marginal likelihood (LPML) [32] is defined as

$$LPML = \sum_{i=1}^{n} \log \widehat{CPO}_i.$$

Under a Bayesian framework, models with larger LPML values are preferred over models with lower LPML values.

3. Results

3.1. Simulation study

In this subsection, we present a simulation study to illustrate the performance of the ML approach when applied to samples from an EW distribution with a cure fraction. The coverage probabilities of the confidence intervals for the parameters, bias and mean square error (MSE) were applied to evaluate the simulation results. We generated $B = 1,000$ samples of size n from the EW distribution following the steps described in the Subsection 2.5 and computed the ML estimates for the B samples, say $(\widehat{\alpha}_{ML}^{(b)}, \widehat{\beta}_{ML}^{(b)}, \widehat{\gamma}_{ML}^{(b)}, \widehat{\eta}_{ML}^{(b)})$ for $b = 1, ..., B$. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method was used to maximize the likelihood function with respect to the model parameters [8]. The average bias in the estimation of a parameter θ is estimated by

$$\widehat{Bias}(\widehat{\theta}_{ML}) = \frac{1}{B} \sum_{b=1}^{B} \left(\widehat{\theta}_{ML}^{(b)} - \theta_N \right)$$

and the corresponding MSE is estimated by

$$\widehat{MSE}(\widehat{\theta}_{ML}) = \frac{1}{B} \sum_{b=1}^{B} \left(\widehat{\theta}_{ML}^{(b)} - \theta_N \right)^2,$$

where $\widehat{\theta}_{ML} \in (\widehat{\alpha}_{ML}, \widehat{\beta}_{ML}, \widehat{\gamma}_{ML}, \widehat{\eta}_{ML})$ is the ML estimate for a given parameter, $\widehat{\theta}_{ML}^{(b)}$ is the ML estimate obtained for θ considering the b-th simulated sample, θ_N is the corresponding nominal value for θ, $\theta \in (\alpha, \beta, \gamma, \eta)$, and B is the number of simulated samples. We repeated these steps for $n = 25, 30, 35, 40, ..., 400$ with different values for α, β, γ and η, including censored data. Figure 2 shows how the coverage probabilities, average biases and MSE vary with respect to the sample size n, considering a nominal confidence coefficient of 95% and four sets of arbitrary values for the parameters of the EW distribution and η, given by (a) $(\alpha, \beta, \gamma) = (1, 0.5, 1)$ and $\eta = 0.2$, (b) $(\alpha, \beta, \gamma) = (2, 1, 1)$ and $\eta = 0.25$, (c) $(\alpha, \beta, \gamma) = (2, 2, 2)$ and $\eta = 0.2$, and (d) $(\alpha, \beta, \gamma) = (4, 0.2, 0.6)$ and $\eta = 0.25$.

We can note in Figure 2 that in all simulations the MSE values decrease and the bias values tend to zero as the sample size increases, as it is expected. In addition, the coverage probability of all parameters approaches to the nominal value 95% when n increases. In this simulation study, confidence intervals for the parameters α, β and γ were based on the delta method (7), while confidence intervals for η were obtained by the Wald method (similar to (6)). For comparison, we also performed a simulation considering all confidence intervals based on the Wald method (6), and in general the coverage probabilities for all parameters were slightly lower than those shown in Figure 2.

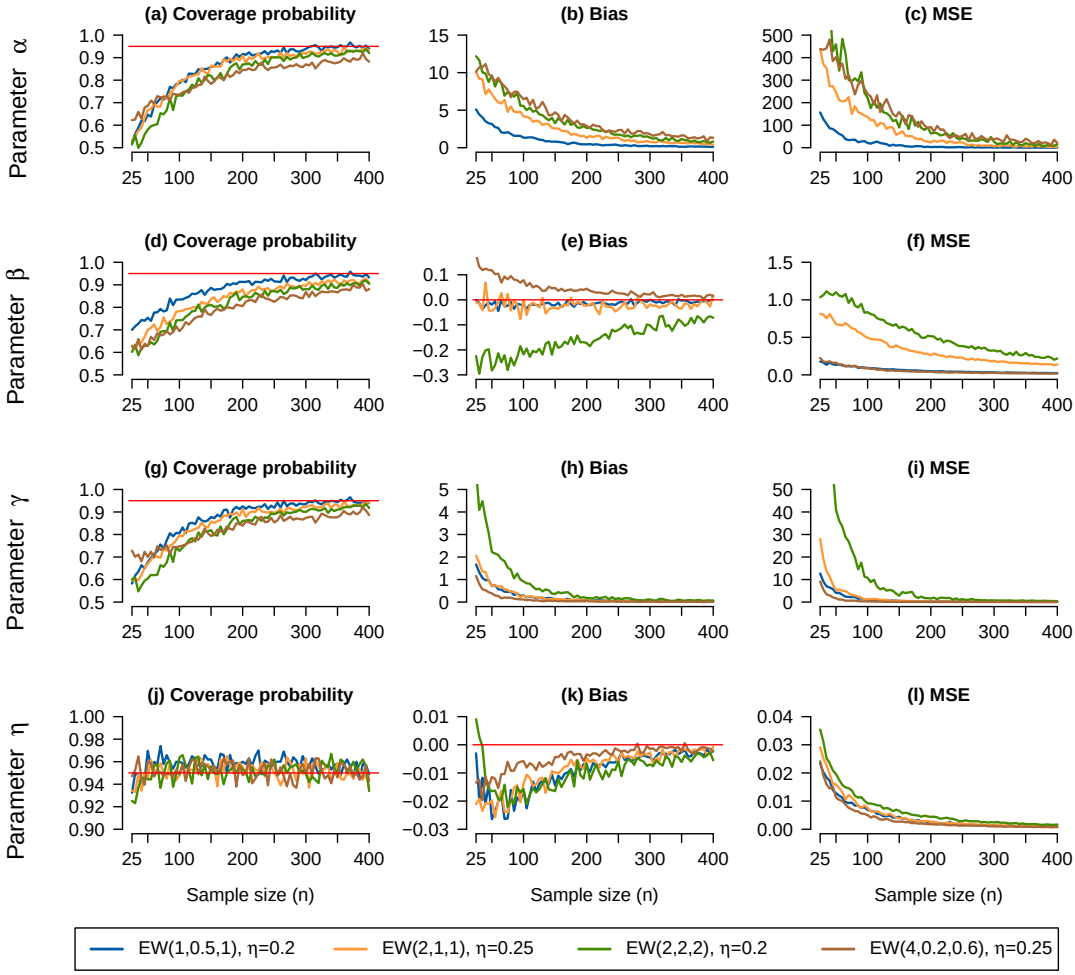

Figure 2. Results from the simulation study. The graphs show the estimates of the coverage probabilities of the 95% confidence intervals for the parameters α, β, γ and η, the average bias, and the mean squared error (MSE).

3.2. Applications to real data sets

In this subsection, we provide two applications to real data to illustrate the potential usefulness of the EW distribution to analyse lifetime data. In all applications, under the frequentist framework, the BFGS method was used to maximize the likelihood function with respect to the model parameters [8].

3.2.1. Survival of ESCC patients with radical resection A retrospective study included 124 patients with esophageal squamous cell carcinoma (ESCC) and radical resection [39]. The patients were treated with esophagectomy in a hospital in Wuhan, China, between March 2010 and December 2012. The variable of interest is the length of time (in months) from the date of surgery to the death from any cause. If a patient did not die during the study period, he/she was treated as censored and the time to the last follow-up was recorded. Approximately one-third of the observations are censored (41/124). Figure 3 shows the survival function estimated by the Kaplan-Meier method and the corresponding TTT plot. The reversed S-shape of the TTT plot suggests that a standard Weibull distribution (without cure fraction) may not be a good choice to be fitted by the data.

Table 1 shows the ML estimates and associated standard errors for all parameters of the mixture cure fraction (MCF), non-mixture cure fraction (NMCF), complementary promotion time cure (CPTC), and the cure rate

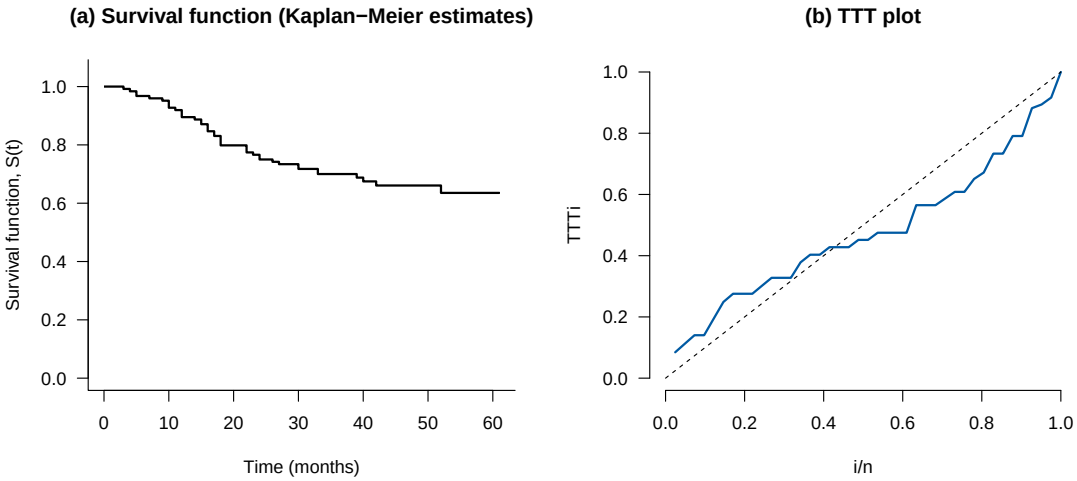

Figure 3. (a) Survival function estimated by the Kaplan-Meier method for 124 ESCC patients with radical resection, and (b) the corresponding TTT plot. Data available from [39].

Table 1. ML estimates and associated standard errors (SE) of the model based on the EW distribution and its special cases, including a cure fraction, fitted to the 124 ESCC patients with radical resection. Data available from Li et al. [39].

Baseline distribution and parameters	ML estimates			
	MCF model Estimate (SE)	**NMCF model** Estimate (SE)	**CPTC model** Estimate (SE)	**CRPO model** Estimate (SE)
EW distribution				
α	7.4238 (9.1605)	4.9756 (3.4251)	7.4884 (5.4300)	4.2735 (4.4570)
β	4.6984 (7.5051)	8.3273 (6.0021)	4.1429 (3.8000)	11.0194 (11.0329)
γ	0.6024 (0.3565)	0.7153 (0.2702)	0.6051 (0.2077)	0.7684 (0.4705)
η	0.6083 (0.0709)	0.6118 (0.0667)	0.6136 (0.0623)	0.6100 (0.0753)
AIC	464.38	464.36	464.35	464.36
SW distribution				
β	24.3646 (3.1641)	26.4089 (2.9267)	22.4209 (3.0238)	28.3995 (4.2929)
γ	1.7285 (0.2481)	1.8016 (0.2542)	1.6460 (0.2395)	1.8859 (0.2678)
η	0.6402 (0.0493)	0.6381 (0.0483)	0.6422 (0.0486)	0.6371 (0.0504)
AIC	463.87	463.56	464.17	463.29
EE distribution				
α	2.6848 (0.7663)	2.7040 (0.8097)	2.6151 (0.7826)	2.7261 (0.6955)
β	13.1873 (3.3313)	14.5737 (4.5148)	12.1242 (2.8455)	16.2194 (4.4703)
η	0.6302 (0.0521)	0.6264 (0.0557)	0.6341 (0.0502)	0.6220 (0.0545)
AIC	462.55	462.46	462.67	462.40
ER distribution				
α	0.8473 (0.1613)	0.8974 (0.2127)	0.7851 (0.1735)	0.9507 (0.1925)
β	26.1254 (3.0412)	27.4334 (6.0296)	25.1545 (3.5975)	28.8785 (5.6066)
η	0.6438 (0.0467)	0.6406 (0.0505)	0.6474 (0.0466)	0.6376 (0.0501)
AIC	464.40	463.86	465.03	463.41

proportional odds (CRPO) models, with baseline survival functions for the susceptible individuals given by the EW, the standard Weibull (SW), the exponentiated exponential (EE), and the exponentiated Rayleigh (ER) distributions. The R code used to obtain the ML estimates and associated standard errors for all parameters of the MCF model based on the EW distribution is provided in an Appendix at the end of the article. The AIC values for the different models are also presented in Table 1. The AIC values are close to each other, not suggesting a better model among those fitted to the data. Panel (a) of Figure 4 compares the estimated values of the survival function obtained by the Kaplan-Meier method and by the parametric MCF model using the ML approach. This graph visually suggests that the EW distribution and its special cases are well fitted by the data. In the panels (b) to (e) of Figure 3, we have

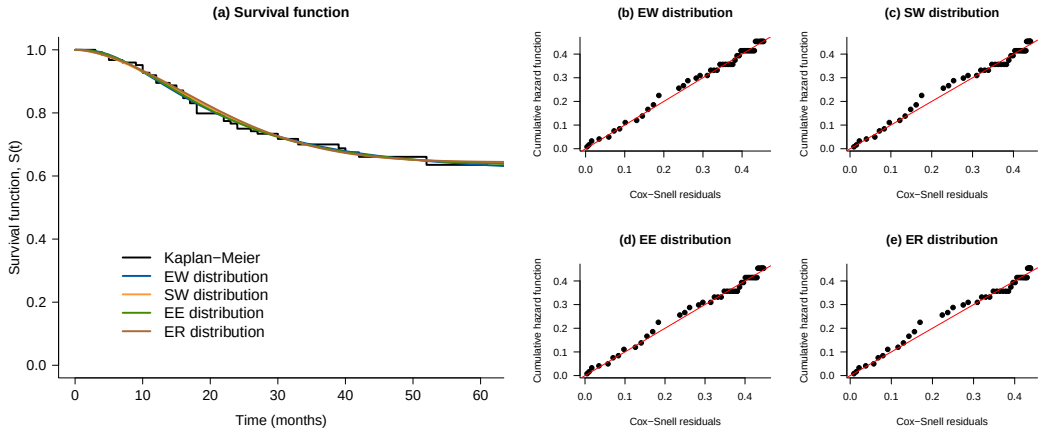

Figure 4. (a) Survival function estimated by the Kaplan-Meier method and assuming the parametric MCF model based on the EW, SW, EE and ER distributions for the 124 ESCC patients with radical resection; and the corresponding Cox-Snell residual plots for the MCF models based on the (b) EW, (c) SW, (d) EE and (e) ER distributions. Data available from [39].

the corresponding Cox-Snell residuals plots for the MCF models based on the EW, SW, EE, and ER distributions, respectively. The estimated cumulative hazard rates are close to the diagonal line, indicating a proper fit of the MCF models to the data. Cox-Snell residuals plots for the NMCF, CPTC, and CRPO models also suggest a reasonable fit of these models to the data (for parsimony, we do not show these graphs here).

Table 2. Bayesian estimates and associated 95% highest probability density (HPD) intervals, considering the model based on the EW distribution and its special cases and including a cure fraction, fitted to the 124 ESCC patients with radical resection. Data available from Li et al. [39].

Baseline distribution and parameters	Bayesian estimates			
	MCF model Estimate (95%HPD)	**NMCF model** Estimate (95%HPD)	**CPTC model** Estimate (95%HPD)	**CRPO model** Estimate (95%HPD)
EW distribution				
α	6.2243 (3.0193, 10.215)	6.6234 (3.3860, 10.887)	5.7193 (2.7484, 9.4317)	7.0517 (3.6068, 11.355)
β	4.7747 (1.3925, 9.0085)	5.0466 (1.4855, 9.4299)	4.4373 (1.2909, 8.6229)	5.3251 (1.7098, 9.7897)
γ	0.5423 (0.2785, 0.8201)	0.5346 (0.2750, 0.8109)	0.5464 (0.3003, 0.8290)	0.5254 (0.2880, 0.8135)
η	0.5786 (0.2828, 0.7358)	0.5860 (0.3681, 0.7284)	0.5714 (0.2520, 0.7231)	0.5918 (0.4098, 0.7240)
LPML	-231.58	-231.65	-231.85	-231.37
SW distribution				
β	19.251 (15.522, 23.533)	20.598 (16.740, 24.891)	17.783 (13.851, 21.862)	21.840 (18.083, 26.550)
γ	1.5681 (1.1402, 2.0022)	1.6779 (1.2513, 2.1186)	1.4448 (1.0212, 1.8565)	1.7904 (1.3755, 2.2532)
η	0.6539 (0.5635, 0.7364)	0.6653 (0.5760, 0.7447)	0.6391 (0.5501, 0.7287)	0.6741 (0.5852, 0.7478)
LPML	-234.04	-232.93	-234.12	-234.12
EE distribution				
α	3.3098 (2.0518, 4.8910)	3.4290 (2.1882, 5.0291)	3.1372 (1.9049, 4.7288)	3.5940 (2.3125, 5.1547)
β	10.204 (7.6164, 13.521)	10.653 (7.9899, 14.239)	9.7886 (7.4011, 12.925)	11.113 (8.1302, 14.739)
η	0.6474 (0.5570, 0.7368)	0.6522 (0.5657, 0.7422)	0.6418 (0.5460, 0.7260)	0.6572 (0.5664, 0.7399)
LPML	-231.74	-231.88	-231.69	-232.06
ER distribution				
α	1.0892 (0.7106, 1.4845)	1.1626 (0.7572, 1.5718)	1.0079 (0.6307, 1.4078)	1.2435 (0.8529, 1.6907)
β	20.348 (17.244, 24.234)	20.763 (17.498, 24.764)	20.007 (16.906, 23.700)	21.162 (17.660, 25.168)
η	0.6589 (0.5727, 0.7434)	0.6611 (0.5756, 0.7423)	0.6571 (0.5764, 0.7486)	0.6644 (0.5797, 0.7455)
LPML	-234.41	-234.30	-234.54	-234.29

Table 2 shows the Bayesian estimates and associated 95%HPD intervals for all parameters of the same models used in the frequentist analysis. The point estimates were determined by the median of the correspondent posterior distribution simulated by the MCMC method. Prior distributions of α, β and η were given by $Gamma(0.1, 0.1)$, $Gamma(0.1, 0.1)$ and $Beta(0.5, 0.5)$, respectively, that is, approximately non-informative priors. The Bayesian

estimates are approximately equal to the ML estimates. Analogously to the previous analysis, the LPML values showed in Table 2 are close to each other, not suggesting a better model among those fitted to the data. Since all absolute Z-scores for each parameter are less than 1.96 (not shown), the Geweke's diagnostic evidenced the convergence of the correspondent chains. Trace plots of the simulated samples for each parameter show convergence of the Gibbs sampling algorithm, and the plots of the auto-correlation function suggest that the posterior samples are approximately uncorrelated (graphics not shown). The R code used to fit the Bayesian model to the data is provided in the Appendix.

3.2.2. Treatments for COVID-19 symptoms The year 2020 was historically marked by the global pandemic crisis caused by the COVID-19 coronavirus [80]. In that year, many studies were conducted in an attempt to investigate new treatments and therapies for individuals affected with the disease. In a randomized clinical trial of 214 patients with confirmed SARS-CoV-2 infection receiving outpatient care [74], volunteers were divided into four groups:

Group 1: usual care without any study medications (a control group).

Group 2: 8000 mg of ascorbic acid (to be divided over 2-3 times per day with meals),

Group 3: 50 mg of zinc gluconate at bedtime,

Group 4: both therapies used in Groups 2 and 3.

In all groups, patients were treated for ten days after a positive diagnosis for COVID-19. The variable of interest is the number of days required to achieve a 50% reduction in disease symptoms. The sample observations are:

Group 1: 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 7, 7, 7, 7, 8, 8, 9, 9, 9, 9, 9, 9, 9, 10, 10, 11, 11, 12, 15, 16, 23, 28+, 28+, 28+, 28+, 28+ and 28+.

Group 2: 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 8, 9, 9, 10, 25, 25, 28+ and 28+.

Group 3: 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 8, 8, 9, 10, 10, 10, 11, 14, 25, 25, 28+, 28+, 28+, 28+, 28+, 28+ and 28+.

Group 4: 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 8, 8, 8, 9, 9, 9, 11, 12, 16, 18, 28+, 28+, 28+, 28+, 28+, 28+, 28+ and 28+.

Plus sign indicates censored observations, representing the patients who did not present the event of interest after 28 days of follow-up. Panel (a) of Figure 5 shows the Kaplan-Meier survival curves comparing the time to symptom reduction between patients submitted to these treatments. As the Figure suggests, the curves reach a plateau at the right tail. In addition, some medical researchers have shown that patients with COVID-19 can experience persistent symptoms over a relatively long period [59, 67], which suggests that the use of models with a cure fraction to study the duration of COVID-19 symptoms is clinically feasible. The corresponding TTT plots are shown in Panel (b) of Figure 5. The reversed S-shape of these curves suggests that the use of models based on the standard Weibull distribution may not be suitable for these data, and models based on more complex distributions are required.

Suppose that x_1, x_2 and x_3 are dummy variables, where $x_{1i} = 0$, $x_{2i} = 0$ and $x_{3i} = 0$ if the ith patient was assigned to receive usual care without any study medications (Group 1), $x_{1i} = 1$, $x_{2i} = 0$ and $x_{3i} = 0$ if the ith patient was assigned to Group 2, $x_{1i} = 0$, $x_{2i} = 1$ and $x_{3i} = 0$ if the ith patient was assigned to Group 3, and $x_{1i} = 0$, $x_{2i} = 0$ and $x_{3i} = 1$ if the ith patient was assigned to Group 4, for $i = 1, ..., n$. In this way, we have a regression model based on the expression (9) where $\alpha(x_i) = \exp(\alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \alpha_3 x_{i3})$, $\beta(x_i) = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3})$ and $\gamma(x_i) = \exp(\gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \gamma_3 x_{i3})$. In addition, considering the MCF model, from (10) we have

$$\eta(x_i) = \frac{\exp\left(\eta_0 + \eta_1 x_{i1} + \eta_2 x_{i2} + \eta_3 x_{i3}\right)}{1 + \exp\left(\eta_0 + \eta_1 x_{i1} + \eta_2 x_{i2} + \eta_3 x_{i3}\right)}. \tag{11}$$

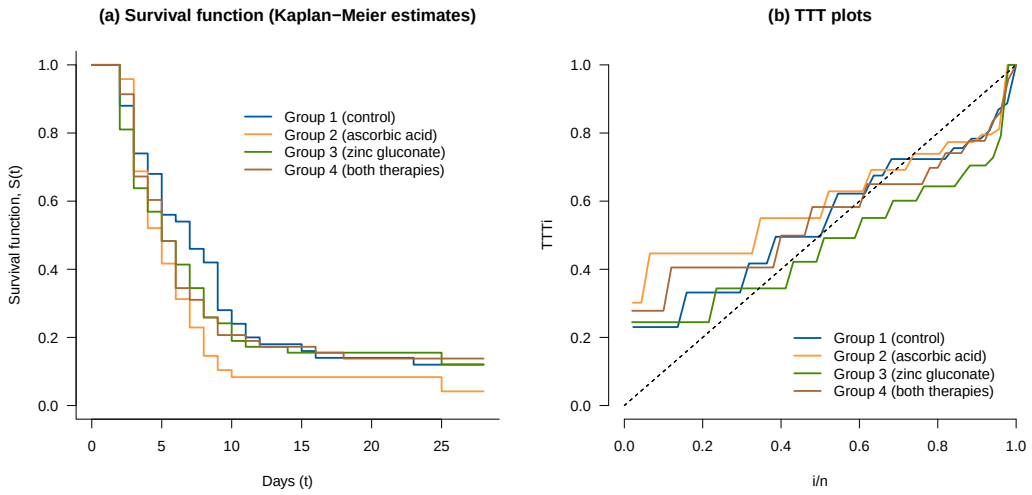

Figure 5. (a) Survival functions estimated by the Kaplan-Meier method, comparing the four treatment groups for COVID-19, and (b) TTT plots. Data available from [74].

Panels (a) and (c) of Figure 6 compare the Kaplan-Meier and parametric ML estimates of the survival function for COVID-19 data, considering the frequentist MCF and NMCF models, respectively. The AIC values for these model are 1122.28 and 1116.91, respectively. In panels (b) and (d) of Figure 6, the Cox-Snell residuals are plotted against the cumulative hazard function to assess the overall fit of the MCF and NMCF models, respectively [17]. All points on the graph are close to the diagonal line, indicating no evidence of lack of fit in a general sense. Figures showing the survival curves estimated by the CPTC and CRPO models are similar to those shown in Figure 6, as are the corresponding Cox-Snell residual plots, and for parsimony, they are not shown.

Table 3 shows the ML and Bayesian estimates of the MCF, NMCF, and CPTC models, and Table 4 shows the estimates of the CRPO models, all based on the EW distribution. We can observe in these tables that the ML and the Bayesian estimates are relatively close. In the frequentist analysis, we have no evidence of differences between the survival curves, given that all p-values are greater than the significance level of 0.05, except the one related to the η_0 intercept. Choices of 0.05 for the significance level are common in applied research [15]. In addition, under the Bayesian framework, we can see that all 95% HPD intervals include the 0 value except for the η_0 intercept in the MCF, NMCF, CPTC, and CRPO models, and also for the γ_0 intercept in the MCF, NMCF, and CPTC models. All Geweke's Z-score values were in the range $(-1.96, 1.96)$, suggesting the convergence of the Markov chains [26].

4. Concluding remarks

In applications of survival analysis methods to medical data, we may find situations where the hazard function is not a monotonically increasing or decreasing curve. This motivates us to explore the use of distributions with more flexible shapes of the hazard functions. In this article, we presented a model based on the EW distribution considering right-censored survival data with a cure fraction and covariates. The hazard function of the EW distribution can have a bathtub, unimodal, increasing, and decreasing shapes. Maximum likelihood (ML) and Bayesian approaches were used to estimate the parameters of the model. Bayesian estimation was based on MCMC methods. The simulation study showed that for small sample sizes, the ML estimates seem to be biased, but the bias reduction of the estimators occurs as the sample size increases. In addition, the ML method returns good coverage probabilities as the sample size increases.

Table 3. ML and Bayesian estimates of the MCF, NMCF and CPTC models based on the EW distribution including covariates. Data available from Thomas et al. [74].

	ML estimates				Bayesian estimates		
Parameter	Estimate	Std. Error	95%CI	p-value	Estimate	95%HPD	Z-score
MCF model							
α_0	2.2606	1.4243	(-0.5310, 5.0522)	0.112	2.5720	(1.5260, 3.4660)	-0.08
α_1	0.5705	1.6877	(-2.7373, 3.8783)	0.735	0.7472	(-0.5039, 1.9430)	0.77
α_2	0.1741	1.5670	(-2.8972, 3.2454)	0.912	0.3067	(-0.8014, 1.5400)	0.95
α_3	0.3154	1.8937	(-3.3963, 4.0271)	0.868	0.4397	(-0.8918, 1.6070)	-0.77
β_0	0.2057	1.5693	(-2.8700, 3.2814)	0.896	-0.2659	(-1.4130, 0.8832)	0.24
β_1	-0.3709	1.7633	(-3.8268, 3.0850)	0.833	-0.4535	(-1.7990, 0.7302)	-1.31
β_2	-0.3957	1.7262	(-3.7789, 2.9876)	0.819	-0.5428	(-1.8960, 0.8142)	-0.86
β_3	-0.1911	1.9461	(-4.0054, 3.6232)	0.922	-0.2349	(-1.5010, 1.0510)	0.89
γ_0	-0.4401	0.4892	(-1.3990, 0.5189)	0.368	-0.5789	(-0.9295, -0.2366)	0.46
γ_1	0.0292	0.5369	(-1.0231, 1.0814)	0.957	0.0389	(-0.3284, 0.4185)	-1.26
γ_2	-0.0745	0.5224	(-1.0985, 0.9495)	0.887	-0.0773	(-0.4579, 0.3023)	-0.75
γ_3	0.0966	0.6064	(-1.0920, 1.2852)	0.873	0.1030	(-0.2877, 0.5109)	0.44
η_0	-1.8267	0.4236	(-2.6569, -0.9965)	$< .001$	-1.9570	(-2.6700, -1.2760)	-1.12
η_1	-0.7803	0.7388	(-2.2282, 0.6677)	0.291	-0.9363	(-2.1080, 0.2539)	0.21
η_2	-0.2378	0.5983	(-1.4105, 0.9349)	0.691	-0.1335	(-1.1650, 0.8381)	0.22
η_3	-0.1087	0.5740	(-1.2336, 1.0162)	0.850	0.0658	(-0.8528, 1.0070)	1.88
			AIC: 1122.28			LPML: -555.05	
NMCF model							
α_0	2.4710	2.2848	(-2.0072, 6.9492)	0.279	2.3090	(1.3800, 3.2780)	-0.75
α_1	0.5973	2.0478	(-3.4163, 4.6110)	0.770	0.5174	(-0.5990, 1.6730)	0.35
α_2	0.1724	2.7581	(-5.2333, 5.5781)	0.950	0.1901	(-0.9092, 1.2640)	1.33
α_3	0.2929	3.1272	(-5.8362, 6.4220)	0.925	0.2770	(-0.8958, 1.4340)	-0.33
β_0	0.2027	3.0714	(-5.8172, 6.2225)	0.947	0.2980	(-0.9714, 1.5160)	0.99
β_1	-0.4247	2.7529	(-5.8204, 4.9709)	0.877	-0.3353	(-1.6500, 1.0530)	-0.39
β_2	-0.4176	3.7226	(-7.7138, 6.8785)	0.910	-0.4215	(-1.8860, 0.9831)	-1.31
β_3	-0.2040	3.9707	(-7.9865, 7.5785)	0.959	-0.2057	(-1.5030, 1.0800)	0.12
γ_0	-0.6103	0.8327	(-2.2424, 1.0218)	0.463	-0.5877	(-1.0200, -0.1358)	1.20
γ_1	0.0270	0.7420	(-1.4273, 1.4813)	0.971	0.0582	(-0.3933, 0.5549)	-0.93
γ_2	-0.0727	0.9664	(-1.9667, 1.8214)	0.940	-0.0515	(-0.5264, 0.4241)	-1.88
γ_3	0.1206	1.0916	(-2.0188, 2.2600)	0.912	0.1591	(-0.2967, 0.6861)	-0.67
η_0	-1.9045	0.5143	(-2.9125, -0.8965)	$< .001$	-1.9690	(-2.7240, -1.3130)	-0.30
η_1	-0.8591	0.8428	(-2.5109, 0.7927)	0.308	-0.9802	(-2.0940, 0.2043)	0.02
η_2	-0.2289	0.7676	(-1.7334, 1.2755)	0.765	-0.1674	(-1.1220, 0.7993)	-1.47
η_3	-0.0588	0.6727	(-1.3772, 1.2597)	0.930	0.0516	(-0.9139, 0.9925)	0.08
			AIC: 1116.91			LPML: -553.81	
CPTC model							
α_0	2.2307	1.4565	(-0.6239, 5.0853)	0.126	2.4240	(1.3500, 3.4070)	1.65
α_1	0.5112	1.6836	(-2.7885, 3.8110)	0.761	0.6293	(-0.6615, 1.9010)	-0.03
α_2	0.1102	1.9040	(-3.6215, 3.8420)	0.954	0.2809	(-0.9135, 1.5370)	0.13
α_3	0.3046	1.7578	(-3.1406, 3.7499)	0.862	0.5091	(-0.7995, 1.8620)	-1.00
β_0	-0.2148	1.5847	(-3.3208, 2.8912)	0.892	-0.6357	(-1.8490, 0.4543)	-1.32
β_1	-0.5506	1.7656	(-4.0111, 2.9098)	0.755	-0.5972	(-1.8000, 0.7313)	0.88
β_2	-0.5037	2.0624	(-4.5460, 3.5387)	0.807	-0.5904	(-1.9880, 0.7151)	-0.16
β_3	-0.2096	1.7955	(-3.7287, 3.3096)	0.907	-0.2504	(-1.5470, 1.0370)	0.85
γ_0	-0.4989	0.4260	(-1.3338, 0.3361)	0.242	-0.6123	(-0.9196, -0.2978)	-1.47
γ_1	-0.0292	0.4627	(-0.9361, 0.8778)	0.950	-0.0027	(-0.3117, 0.3407)	1.14
γ_2	-0.1059	0.5234	(-1.1318, 0.9199)	0.840	-0.0881	(-0.4402, 0.2607)	0.47
γ_3	0.0632	0.4818	(-0.8811, 1.0076)	0.896	0.0814	(-0.2875, 0.4314)	1.04
η_0	-1.8216	0.4462	(-2.6961, -0.9470)	$< .001$	-1.9090	(-2.5930, -1.2760)	0.48
η_1	-0.8133	0.8202	(-2.4208, 0.7943)	0.321	-0.9854	(-2.1530, 0.2090)	1.58
η_2	-0.2065	0.6255	(-1.4323, 1.0194)	0.741	-0.0898	(-1.0760, 0.8253)	0.67
η_3	-0.0921	0.6259	(-1.3189, 1.1347)	0.883	0.0529	(-0.8325, 0.9951)	-0.10
			AIC: 1131.16			LPML: -559.39	

Two applications to real data were used to illustrate the potential usefulness of the EW distribution to analyze lifetime data. The first application is a clinical trial including 124 patients with ESCC and radical resection treated with esophagectomy. We also presented an application of the proposed model to a dataset from a trial comparing treatments for COVID-19 symptoms [74]. We emphasise that the results found in this data analysis should not be

Table 4. ML and Bayesian estimates of the CRPO model based on the EW distribution including covariates. Data available from Thomas et al. [74].

	ML estimates				Bayesian estimates		
Parameter	Estimate	Std. Error	95%CI	p-value	Estimate	95%HPD	Z-score
CRPO model							
α_0	2.5895	1.9272	(-1.1877, 6.3666)	0.179	2.1410	(1.0810, 3.1080)	-0.74
α_1	0.6101	2.1193	(-3.5437, 4.7639)	0.773	0.3842	(-0.7093, 1.3280)	1.92
α_2	0.1988	2.4002	(-4.5054, 4.9030)	0.934	0.1163	(-0.9436, 1.1840)	-0.68
α_3	0.2776	1.6696	(-2.9947, 3.5498)	0.868	0.1240	(-1.1090, 1.2370)	0.52
β_0	0.3615	3.1350	(-5.7830, 6.5061)	0.908	0.9223	(-0.4052, 2.2310)	0.45
β_1	-0.4386	3.3758	(-7.0551, 6.1779)	0.897	-0.1633	(-1.4710, 1.1650)	-0.70
β_2	-0.4343	3.9620	(-8.1998, 7.3311)	0.913	-0.2789	(-1.7960, 1.1460)	0.74
β_3	-0.2395	2.2624	(-4.6737, 4.1947)	0.916	-0.2005	(-1.5810, 1.1250)	0.24
γ_0	-0.7651	0.9135	(-2.5554, 1.0253)	0.402	-0.5463	(-1.0990, 0.0438)	0.54
γ_1	0.0251	1.0242	(-1.9822, 2.0324)	0.980	0.0641	(-0.5241, 0.6418)	-0.76
γ_2	-0.0573	1.1038	(-2.2208, 2.1061)	0.959	-0.0619	(-0.6073, 0.6082)	1.06
γ_3	0.1736	0.6190	(-1.0396, 1.3868)	0.779	0.2051	(-0.4242, 0.8233)	-0.28
η_0	-1.9869	0.6307	(-3.2231, -0.7507)	$< .001$	-1.9560	(-2.7000,-1.2350)	0.58
η_1	-0.9789	1.1052	(-3.1451, 1.1873)	0.376	-1.0510	(-2.2170, 0.1386)	1.60
η_2	-0.2138	0.8693	(-1.9176, 1.4900)	0.806	-0.3080	(-1.3770, 0.8873)	-0.13
η_3	0.0364	0.7485	(-1.4307, 1.5035)	0.961	0.0195	(-0.9139, 1.0200)	-0.55
		AIC: 1115.26				LPML: -552.68	

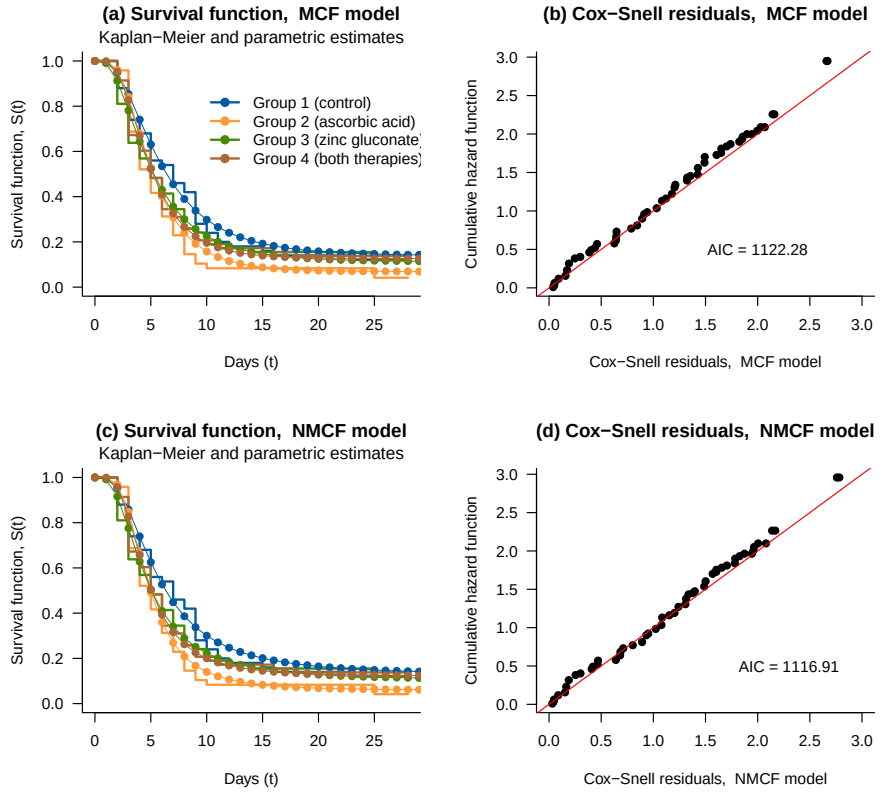

Figure 6. Survival functions estimated by the Kaplan-Meier method and assuming the parametric (a) MCF and (c) NMCF models based on the EW distribution, comparing the four treatment groups for COVID-19. Panels (b) and (d) show the corresponding Cox-Snell residual plot. Data available from [74].

used as clinical evidence in favour or against the treatments under investigation, as it is necessary to consider the presence of some factors that could lead to bias in these results from a medical perspective. In addition, the usual concept of statistical significance (suggested by p-values less than 0.05) is not synonymous with clinical relevance, as has been described by many authors [34, 76]. However, this application was helpful to illustrate the performance of the model based on the EW distribution when applied to real data.

Our findings indicate that both the ML and Bayesian approaches are computationally feasible to estimate the parameters of EW distribution. The EW distribution and its sub-models have the flexibility to accommodate different shapes for the hazard function and should be an attractive choice for survival data analysis when a cure fraction is present.

Acknowledgement

The researchers of EZM and JAA are supported by the Brazilian organization CNPq.

Appendix: R Code

The following R code can be used to obtain the ML estimates and associated standard errors for all parameters of the mixture cure fraction based on the EW distribution. Let us consider the example introduced in subsection 3.2.1.

```
# Clear existing data
rm(list=ls())
# Load maxLik library
library(maxLik)
# Read Data (Li et al.)
# t is the survival time and d is the censoring indicator
# d = 1: complete observation, d = 0: censored observation
t<-c(49, 49, 30, 10, 7, 37, 57, 17, 15, 33, 60, 42, 36, 17, 39,
38, 35, 45, 3, 50, 52, 24, 40, 34, 18, 60, 39, 5, 24, 4, 5, 54,
10, 14, 52, 18, 18, 22, 42, 49, 38, 12, 10, 44, 36, 26, 61, 54,
34, 31, 33, 18, 54, 38, 29, 46, 39, 11, 59, 61, 33, 59, 49, 32,
22, 49, 22, 41, 59, 37, 36, 34, 35, 51, 57, 42, 47, 9, 36, 16,
43, 43, 12, 57, 54, 57, 23, 30, 16, 38, 59, 16, 41, 58, 51, 37,
32, 34, 53, 33, 41, 53, 30, 41, 51, 50, 52, 34, 32, 39, 59, 35,
12, 36, 61, 30, 46, 15, 57, 33, 45, 27, 48, 56)
d<-c(0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0,
1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,
1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1,
0, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0)

# The log-likelihood function with a cure fraction
log.f <- function(parms) {
alpha   <- parms[1]
beta    <- parms[2]
gamma   <- parms[3]
eta     <- parms[4]  # The cure fraction
if (parms[1]<0) return(-Inf)
if (parms[2]<0) return(-Inf)
```

```
if (parms[3]<0) return(-Inf)
if (parms[4]<0) return(-Inf)
S0t <- 1-(1-exp(-(t/beta)^gamma))^alpha
f0t <- alpha*(gamma/beta)*(t/beta)^(gamma-1)*
(1-exp(-(t/beta)^gamma))^(alpha-1)*exp(-(t/beta)^gamma)
St  <- eta + (1-eta)*S0t
ft  <- (1-eta)*f0t
like <- ft^d * St^(1-d)
L <- sum(log(like))
if (is.na(L)==TRUE) {return(-Inf)} else {return(L)} }

mle  <- c()
# Setting the initial values for the parameters
init <- c(35,0.36,0.33,0.571)
# Obtaining the ML estimates
mle  <- maxLik(logLik=log.f,start=init,method="BFGS")
summary(mle)
AIC(mle) # AIC value
```

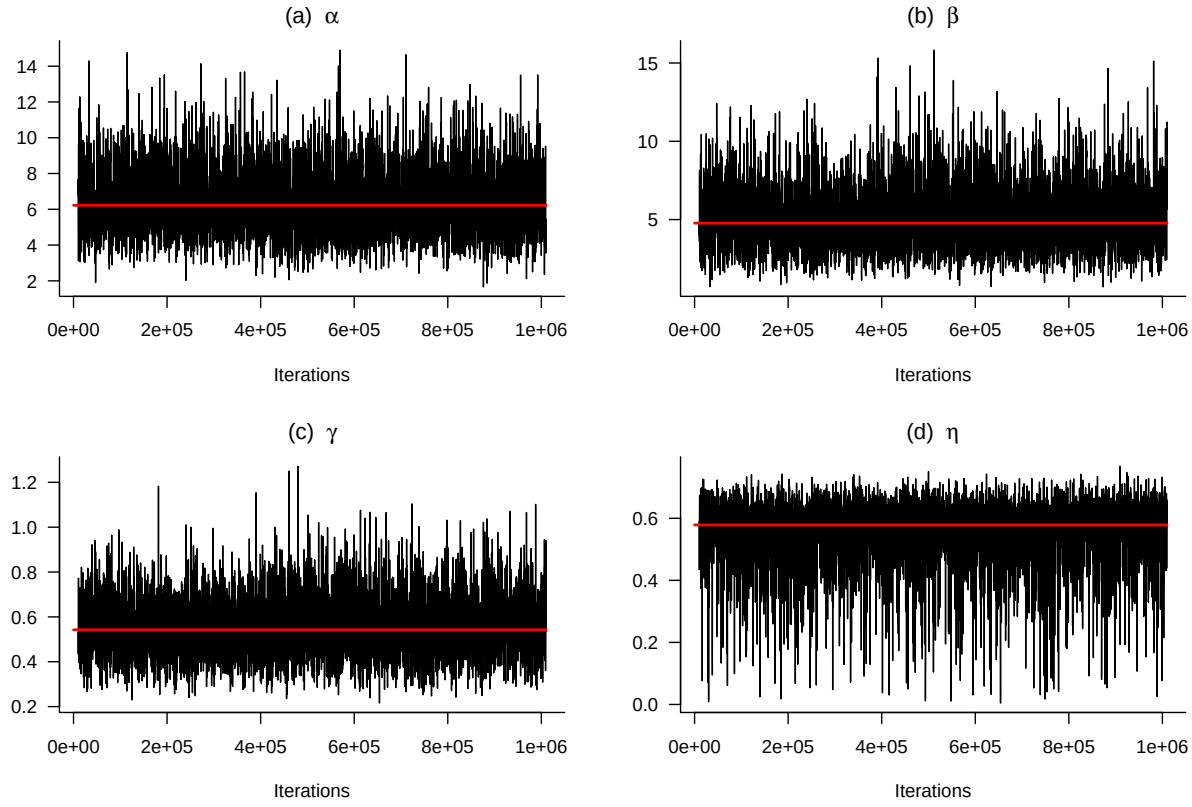

Figure 7. Traceplots of the MCMC samples. The horizontal red lines indicate the posterior means.

Similarly, Bayesian estimates can be obtained by using the following R code.

```
# Load MCMCpack library
```

```
library(MCMCpack)
# The log posterior function with a cure fraction
log.post <- function(t,d,parms) {
alpha <- parms[1]
beta  <- parms[2]
gamma <- parms[3]
eta   <- parms[4]
if (parms[1]<0) return(-Inf)
if (parms[2]<0) return(-Inf)
if (parms[3]<0) return(-Inf)
if (parms[4]<0) return(-Inf)
S0t  <- 1-(1-exp(-(t/beta)^gamma))^alpha
f0t  <- alpha*(gamma/beta)*(t/beta)^(gamma-1)*
(1-exp(-(t/beta)^gamma))^(alpha-1)*exp(-(t/beta)^gamma)
St   <- eta + (1-eta)*S0t
ft   <- (1-eta)*f0t
like <- ft^d * St^(1-d)
log.like  <- sum(log(like))
prior     <- dgamma(alpha,.1,.1)*dgamma(beta,.1,.1)*
dgamma(gamma,.1,.1)*dbeta(eta,0.5,0.5)
log.prior <- log(prior)
L <- log.like + log.prior
if (is.na(L)==TRUE) {return(-Inf)} else {return(L)} }

# Obtaining the MCMC estimates
posterior <- MCMCmetrop1R(log.post,
theta.init=c(alpha=7.5,beta=4.5,gamma=0.6,eta=0.6),
burnin=5000, mcmc=1000000, thin=200, logfun=T, t=t,
d=d, verbose=100000, tune = 1)
varnames(posterior) <- c("alpha","beta","gamma","eta")
summary(posterior)
# Obtaining the HPD intervals
HPDinterval(posterior, prob = 0.95)
# Geweke Z scores
geweke.diag(posterior)
```

Figure 7 shows the traceplots of the corresponding MCMC chains, and the horizontal red lines indicate the posterior means.

These R codes can be easily adapted to the models based on the standard Weibull (SW), exponentiated exponential (EE), and exponentiated Rayleigh (ER) distributions.

REFERENCES

1. J. A. Achcar, E. A. Coelho-Barros, and J. Mazucheli, *Cure fraction models using mixture and non-mixture models*, Tatra Mountains Mathematical Publications, vol. 51, no. 1, pp. 1–9, 2012.
2. J. Aitchison, and J. A. Brown, *The lognormal distribution*, Cambridge, Cambridge University Press, 1957.
3. H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, vol. 19, no. 6, pp. 716–723, 1974.
4. M. Alizadeh, M. N. Khan, M. Rasekhi, and G. G. Hamedani, *A new generalized modified Weibull distribution*, Statistics, Optimization & Information Computing, vol. 9, no. 1, pp. 17–34, 2021.
5. I. Arano, T. Sugimoto, T. Hamasaki, and Y. Ohno, *Practical application of cure mixture model for long-term censored survivor data from a withdrawal clinical trial of patients with major depressive disorder*, BMC Medical Research Methodology, vol. 10, no. 1, pp. 1–13, 2010.

6. L. Benkhelifa, *The Weibull Birnbaum-Saunders distribution and its applications*, Statistics, Optimization & Information Computing, vol. 9, no. 1, pp. 61–81, 2021.
7. J. W. Boag, *Maximum likelihood estimates of the proportion of patients cured by cancer therapy*, Journal of the Royal Statistical Society Series B, vol. 11, no. 1, pp. 15–53, 1949.
8. C. G. Broyden, *The convergence of a class of double-rank minimization algorithms 1. General Considerations*, IMA Journal of Applied Mathematics, vol. 6, no. 1, pp. 76–90, 1970.
9. J. M. Carrasco, E. M. Ortega, and G. M. Cordeiro, *A generalized modified Weibull distribution for lifetime modeling*, Computational Statistics & Data Analysis, vol. 53, no. 2, pp. 450–462, 2008.
10. J. E. Cavanaugh, and A. A. Neath, *The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements*, Wiley Interdisciplinary Reviews: Computational Statistics, vol. 11, no. 3, pp. e1460, 2019.
11. M. H. Chen, J. G. Ibrahim, and D. Sinha, *A new Bayesian model for survival data with a surviving fraction*, The Journal of the American Statistical Association, vol. 94, no. 447, pp. 909–919, 1999.
12. M. H. Chen, and Q. M. Shao, *Monte Carlo estimation of Bayesian credible and HPD intervals*, Journal of Computational and Graphical Statistics, vol. 8, no. 1, pp. 69–92, 1999.
13. S. Chib, and E. Greenberg, *Understanding the Metropolis-Hastings algorithm*, The American Statistician, vol. 49, no. 4, pp. 327–335, 1995.
14. R. Christensen, W. Johnson, A. Branscum, T. E. Hanson, *Bayesian ideas and data analysis: an introduction for scientists and statisticians*, Boca Raton, CRC press, 2011.
15. J. Cohen, *The Earth is round ($p < .05$)*, American Psychologist, vol. 49, no. 12, pp. 997–003, 1994.
16. G. M. Cordeiro, E. M., Ortega, G. O. and Silva, *The Kumaraswamy modified Weibull distribution: theory and applications*, Journal of Statistical Computation and Simulation, vol. 84, no. 7, pp. 1387–1411, 2014.
17. D. R. Cox, and E. J. Snell, *A general definition of residuals*, Journal of the Royal Statistical Society Series B, vol. 30, no. 2, pp. 248–275, 1968.
18. K. Davies, S. Pal, and J. A. Siddiqua, *Stochastic EM algorithm for generalized exponential cure rate model and an empirical study*, Journal of Applied Statistics, vol. 48, no. 12, p. 2112–2135, 2021.
19. S. V. Deo, V. Deo, and V. Sundaram, *Survival analysis - part 2: Cox proportional hazards model*, Indian Journal of Thoracic and Cardiovascular Surgery, vol. 37, pp. 229–233, 2021.
20. D. K. Dey, M. H. Chen, and H. Chang, *Bayesian approach for nonlinear random effects models*, Biometrics, vol. 53, no. 4, pp. 1239–1252, 1957.
21. V. T. Farewell, *The use of mixture models for the analysis of survival data with long-term survivors*, Biometrics, vol. 38, no. 4, pp. 1041–1046, 1982.
22. T. R. Fleming, and D. Y. Lin, *Survival analysis in clinical trials: past developments and future directions*, Biometrics, vol. 56, no. 4, pp. 971–983, 2000.
23. M. L. Garg, B. R. Rao, and C. K. Redmont, *Maximum-likelihood estimation of the parameters of the Gompertz survival function*, Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 19, no. 2, pp. 152–159, 1970.
24. S. Geisser, and W. F. Eddy, *A predictive approach to model selection*, Journal of the American Statistical Association, vol. 74, no. 365, pp. 153–160, 1979.
25. A. E. Gelfand, and D. K. Dey, *Bayesian model choice: asymptotics and exact calculations*, Journal of the Royal Statistical Society. Series B, vol. 56, no. 3, pp. 501–514, 1994.
26. J. Geweke, *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, Bayesian Statistics, vol. 4, pp. 641-649, 1992.
27. Y. Gu, D. Sinha, and S. Banerjee, *Analysis of cure rate survival data under proportional odds model*, Lifetime Data Analysis, vol. 17, no. 1, pp. 123–134, 2011.
28. A. K. Gupta, and S. Nadarajah, *Handbook of beta distribution and its applications*, Boca Raton, CRC press, 2004.
29. R. C. Gupta, R. D. Gupta, and P. L. Gupta, *Modeling failure timedata by Lehman alternatives*, Communication in Statistics - Theory and Methods, vol. 27, no. 4, pp. 887–904, 1998.
30. A. Henningsen, and O. Toomet, *maxLik: A package for maximum likelihood estimation in R*, Computational Statistics, vol. 26, no. 3, pp. 443–458, 2011.
31. H. A. Howlader, and A. M. Hossain, *Bayesian survival estimation of Pareto distribution of the second kind based on failure-censored data*, Computational Statistics & Data Analysis, vol. 38, no. 3, pp. 301–314, 2002.
32. J. G. Ibrahim, M. H. Chen, and D. Sinha, *Bayesian Survival Analysis*, Springer-Verlag, New York, 2001.
33. A. A. Jácome, D. R. Wohnrath, C. Scapulatempo-Neto, E. C. Carneseca, S. V. Serrano, L. S. Viana, J. S. Nunes, E. Z. Martinez, J. S. Santos, *Prognostic value of epidermal growth factor receptors in gastric cancer: a survival analysis by Weibull model incorporating long-term survivors*, Gastric Cancer, vol. 17, no. 1, pp. 76–86, 2014.
34. M. Kieser, T. Friede, and M. Gondan, *Assessment of statistical significance and clinical relevance*, Statistics in Medicine, vol. 32, no. 10, pp. 1707–1719, 2013.
35. D. H. Kim, W. D. Lee, and S. G. Kang, *Bayesian survival estimation of Pareto distribution of the second kind based on type II censored data*, Communications for Statistical Applications and Methods, vol. 12, no. 3, pp. 729–742, 2005.
36. B. Klefsjö, *TTT-plotting - a tool for both theoretical and practical problems*, Journal of Statistical Planning and Inference, vol. 29, no. 1–2, pp. 99–110, 1991.
37. J. P. Klein, and M. L. Moeschberger, *Survival analysis: techniques for censored and truncated data*, Springer, New York, 2003.
38. P. C. Lambert, *Modeling of the cure fraction in survival studies*, The Stata Journal, vol. 7, no. 3, pp. 351–375, 2007.
39. J. Li, Y. Tang, L. Huang, Q. Yu, G. Hu, and X. Yuan, *Genetic variants in the p14ARF/ MDM2/ TP53 pathway are associated with the prognosis of esophageal squamous cell carcinoma patients treated with radical resection*, PloS one, vol. 11, pp. e0158613, 2016.
40. C. S. Li, J. M. Taylor, and J. P. Sy, *Identifiability of cure models*, Statistics & Probability Letters, vol. 54, no. 4, pp. 389–395, 2001.
41. M. A. Looha, E. Zarean, F. Masaebi, M. A. Pourhoseingholi, and M. R. Zali, *Assessment of prognostic factors in long-term survival of male and female patients with colorectal cancer using non-mixture cure model based on the Weibull distribution*, Surgical Oncology,

pp. 101562, 2021.
42. R. A. Maller, and X. Zhou, *Survival analysis with long-term survivors*, Wiley, New York, 1996.
43. A. D. Martin, K. M. Quinn, and J. H. Park, *MCMCpack: Markov chain Monte Carlo in R*, Journal of Statistical Software, vol. 42, no. 9, pp. 1–21, 2011.
44. E. Z. Martinez, and J. A. Achcar, *A new straightforward defective distribution for survival analysis in the presence of a cure fraction*, Journal of Statistical Theory and Practice, vol. 12, no. 4, pp. 688–703, 2018.
45. E. Z. Martinez, J. A. Achcar, A. A. Jácome, and J. S. Santos, *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, vol. 112, no. 3, pp. 343–355, 2013.
46. M. Meshkat, A. R. Baghestani, F. Zayeri, M. Khayamzadeh, and M. E. Akbari, *Survival probability and prognostic factors of Iranian breast cancer patients using cure rate model*, The Breast Journal, vol. 24, no. 6, pp. 1015–1018, 2018.
47. G. S. Mudholkar, and A. D. Hutson, *The exponentiated Weibull family: some properties and a flood data application*, Communication in Statistics - Theory and Methods, vol. 25, no. 12, pp. 3059–3083, 1996.
48. G. S. Mudholkar, and D. K. Srivastava, *Exponentiated Weibull family for analyzing bathtub failure-rate data*, IEEE Transactions on Reliability, vol. 42, no. 2, pp. 299–302, 1993.
49. G. S. Mudholkar, D. K. Srivastava, and M. Freimer, *The exponentiated Weibull family: a reanalysis of the bus-motor-failure data*, Technometrics, vol. 37, no. 4, pp. 436–445, 1995.
50. M. M. Nassar, and F. H. Eissa, *On the exponentiated Weibull distribution*, Communications in Statistics - Theory and Methods, vol. 32, no. 7, pp. 1317–1336, 2003.
51. G. W. Oehlert, *A note on the delta method*, The American Statistician, vol. 46, no. 1, pp. 27–29, 1992.
52. R. P. Oliveira, M. V. Oliveira-Peres, M. R. Santos, E. Z. Martinez, and J. A. Achcar, *A Bayesian inference approach for bivariate Weibull distributions derived from Roy and Morgenstern methods*, Statistics, Optimization & Information Computing, vol. 9, no. 3, pp. 529–554, 2021.
53. M. E. Omer, M. A. Bakar, M. Adam, and M. Mustafa, *Utilization of a mixture cure rate model based on the generalized modified Weibull distribution for the analysis of leukemia patients*, Asian Pacific Journal of Cancer Prevention, vol. 22, no. 4, pp. 1045–1053, 2021.
54. S. Pasari, and O. Dikshit, O. *Stochastic earthquake interevent time modeling from exponentiated Weibull distributions*, Natural Hazards, vol. 90, no. 2, pp. 823–842, 2018.
55. Y. Peng, and J. M. Taylor, *Residual-based model diagnosis methods for mixture cure models*, Biometrics, vol. 73, no. 2, pp. 495–505, 2017.
56. Y. Peng, and B. Yu, *Cure Models: Methods, Applications, and Implementation*, CRC Press, New York, 2021.
57. J. E. Pinder III, J. G. Wiener, and M. H. Smith, *The Weibull distribution: a new method of summarizing survivorship data*, Ecology, vol. 59, no. 1, pp. 175–179, 1978.
58. M. Plummer, N. Best, K. Cowles, and K. Vines, *CODA: convergence diagnosis and output analysis for MCMC*, R News, vol. 6, no. 1, pp. 7–11, 2006
59. A. Ramakrishnan, J. Zreloff, M. A. Moore, S. H. Bergquist, M. Cellai, J. Higdon, J. B. O'Keefe, D. Roberts, and H. M. Wu, *Prolonged symptoms after COVID-19 infection in outpatients*, Open Forum Infectious Diseases, vol. 8, no. 3, pp. ofab060, 2021.
60. P. Ramos, D. Guzman, A. Mota, F. Rodrigues, and F. Louzada, *Sampling with censored data: a practical guide*, arXiv preprint, vol. arXiv:2011.08417, 2020.
61. P. L. Ramos, D. C. Nascimento, C. Cocolo, M. J. Nicola, C. Alonso, L. G. Ribeiro, A, Ennes, and F. Louzada, *Reliability-centered maintenance: analyzing failure in harvest sugarcane machine using some generalizations of the Weibull distribution*, Modelling and Simulation in Engineering, vol. 2018, pp. 1241856, 2018.
62. E. Ramos, P. L. Ramos, and F. Louzada, *Posterior properties of the Weibull distribution for censored data*, Statistics & Probability Letters, vol. 166, pp. 108873, 2020.
63. C. Ricci, S. Partelli, L. Landoni, M. Rinzivillo, C. Ingaldi, V. Andreasi, C. Nessi, F. Muffatti, M. Fontana, D. Tamburrino, G. Deiro, L. Alberici, D. Campana, F. Panzuto, C. Bassi, M. Falconi, and R. Casadei, *Sporadic non-functioning pancreatic neuroendocrine tumours: multicentre analysis*, British Journal of Surgery, vol. 108, no. 7, pp. 811–816, 2021.
64. H. Rinne, *The Weibull distribution: a handbook*, CRC Press, New York, 2008.
65. R. Rocha, S. Nadarajah, V. Tomazella, and F. Louzada, *A new class of defective models based on the Marshall-Olkin family of distributions for cure rate modeling*, Computational Statistics & Data Analysis, vol. 107, pp. 48–63, 2017.
66. J. Scudilio, V. F. Calsavara, R. Rocha, F. Louzada, V. Tomazella, and A. S. Rodrigues, *Defective models induced by gamma frailty term for survival data with cured fraction*, Journal of Applied Statistics, vol. 46, no. 3, pp. 484–507, 2019.
67. W. Shah, T. Hillman, E. D. Playford, and L. Hishmeh, *Managing the long term effects of covid-19: summary of NICE, SIGN, and RCGP rapid guideline*, BMJ, vol. 372, pp. n136, 2021.
68. N. R. Smoll, K. Schaller, and O. P. Gautschi, *The cure fraction of glioblastoma multiforme*, Neuroepidemiology, vol. 39, no. 1, pp. 63–69, 2012.
69. G. O. Silva, E. M. Ortega, and G. M. Cordeiro, *The beta modified Weibull distribution*, Lifetime Data Analysis, vol. 16, no. 3, pp. 409–430, 2010.
70. R. Singh, and K. Mukhopadhyay, *Survival analysis in clinical trials: basics and must know areas*, Perspectives in Clinical Research, vol. 2, no. 4, pp. 145–148, 2011.
71. J. G. Surles, and W. J. Padgett, *Inference for reliability and stress-strength for a scaled Burr type X distribution*, Lifetime Data Analysis, vol. 7, no. 2, pp. 187–200, 2001.
72. M. Teimouri, S. M. Hoseini, S. Nadarajah, *Comparison of estimation methods for the Weibull distribution*, Statistics, vol. 47, no. 1, pp. 93–109, 2013.
73. D. R. Thoman, L. J. Bain, and C. E. Antle, *Inferences on the parameters of the Weibull distribution*, Technometrics, vol. 11, no. 3, pp. 445–460, 1969.

74. S. Thomas, D. Patel, B. Bittel, K. Wolski, Q. Wang, A. Kumar, Z. J. ll'Giovine, R. Mehra, C. McWilliams, S. E. Nissen, M. Y. Desai, *Effect of high-dose zinc and ascorbic acid supplementation vs usual care on symptom length and reduction among ambulatory patients with SARS-CoV-2 infection: the COVID A to Z Randomized Clinical Trial*, JAMA Network Open, vol. 4, no. 2, pp. e210369, 2021.
75. A. D. Tsodikov, J. G. Ibrahim, and A. Y. Yakovlev, *Estimating cure rates from survival data: an alternative to two-component mixture models*, Journal of the American Statistical Association, vol. 98, no. 464, pp. 1063–1078, 2003.
76. M. H. van Rijn, A. Bech, J. Bouyer, and J. A. van den Brand, *Statistical significance versus clinical relevance*, Nephrology Dialysis Transplantation, vol. 32, pp. ii6-ii12, 2017.
77. M. V. P. Vigas, M. B. Fatoretto, G. S. Slanzon, E. M. M. Ortega, C. G. B. Demétrio, and C. M. M. Bittar, *Red propolis effect analysis of dairy calves health based on Weibull regression model with long-term survivors*, Research in Veterinary Science, vol. 136, pp. 464–471, 2021.
78. A. Y. Yakovlev, and A. D. Tsodikov, *Stochastic models of tumor latency and their biostatistical applications*, World Scientific, New Jersey, 1996.
79. B. Yiqi, V. G. Cancho, D. K. Dey, N. Balakrishnan, and A. K. Suzuki, *Power series cure rate model for spatially correlated interval-censored data based on generalized extreme value distribution*, Journal of Computational and Applied Mathematics, vol. 364, pp. 112362, 2020.
80. P. Zhai, Y. Ding, X. Wu, J. Long, Y. Zhong, and Y. Li, *The epidemiology, diagnosis and treatment of COVID-19*, International Journal of Antimicrobial Agents, vol. 55, no. 5, pp. 105955, 2020.