Malaysian Journal of Mathematical Sciences 19(3): 789–810 (2025)
https://doi.org/10.47836/mjms.19.3.02

Malaysian Journal of Mathematical Sciences

Journal homepage: https://mjms.upm.edu.my

Flexible Proportional Odds Regression Model with Applications

Ishag, M. A. S.[1], Wanjoya, A.[2], Adem, A.[3], Alghamdi, A. S.[4], and Afify, A. Z.*[5]

[1]*Department of Mathematics, , Institute for Basic Sciences, Technology and Innovation (PAUSTI), Pan African University Nairobi 62000-00200, Kenya*
[2]*Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi 62000-00200, Kenya*
[3]*Department of Mathematics and Physics, Technical University of Mombasa, Kenya*
[4]*Department of Mathematics, College of Science & Arts, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi Arabia*
[5]*Department of Statistics, Mathematics, and Insurance, Benha University, Benha 13511, Egypt*

E-mail: ahmed.afify@fcom.bu.edu.eg
**Corresponding author*

Received: 15 October 2024
Accepted: 2 March 2025

Abstract

This study proposes a flexible parametric proportional odds regression model that incorporates the exponentiated-Weibull distribution as a baseline for analyzing censored lifetime data. The proposed model is referred to as the exponentiated-Weibull proportional odds regression model. This model provides greater flexibility in capturing a wider range of hazard shapes and survival patterns. The paper discusses the theoretical framework as well as estimation methods for the model parameters. Additionally, extensive simulation studies are conducted to evaluate the proposed model's performance under different scenarios. The results demonstrate that the model effectively accommodates the unique characteristics of the exponentiated-Weibull distribution. Furthermore, two real-world datasets are presented to illustrate and compare the model's practical application and performance with existing proportional odds regression models. The findings highlight the advantages of using the proposed model and its potential to enhance the analysis of survival data and capture complex survival patterns.

Keywords: proportional odds regression model; survival analysis; censored data; exponentiated Weibull distribution; maximum likelihood estimation; simulation study.

1 Introduction

The proportional odds (PO) model, also known as the cumulative odds model or the ordinal logistic regression model, is a versatile and intuitive framework for analyzing ordinal response variables, as demonstrated in [2, 6]. It has demonstrated exemplary performance in many applications and assumes that the cumulative probabilities of each response category follow a PO structure. This indicates that the influence of covariates remains unchanged over time [24, 30]. Regarding probability distributions, they serve as the foundation for survival analysis and may be categorized as nonparametric, semiparametric, or parametric [8]. The Cox proportional hazard (PH) model is commonly used to analyze time-to-event data. However, including too many predictors in the model can lead to complications. To identify key genes and improve classification accuracy, a new method for selecting tuning parameters has been proposed [1, 9].

The parametric survival models are more valuable and efficient for handling various censored data if the distribution assumption is correct [22]. Yang and Prentice [29] developed semiparametric inference in the PO regression model. Royston and Parmar [24] presented parametric PH and PO models for censored data, applying to predictive modeling and the estimation of treatment effects. Hsieh and Chen [10] proposed two methods for assessing the regression parameters for the PH and PO models using dependent truncated data. Vieira et al. [26] created a PO model using log-logistic and discrete Weibull distributions as foundational models. Additionally, Muse et al. [21] proposed a parametric framework of hazard-based and odds-based regression models specifically for analyzing right-censored survival data. Mahanta and Hazarika [16] developed a new multivariate PO frailty model by using a Weibull hazard function (HF) in the context of the Bayesian mechanism. Zhu et al. [31] examined the efficient odds ratio estimation for the PO model with censored time-lagged outcomes. Wang and Wang [27] has highlighted the computational complexity and inefficiency of the PO model with right-censored data. Huang et al. [11] discussed the efficient estimation and inference in the PO model for survival data. Nonparametric inference under right-censored data and under interval-censored survival data are discussed by [3, 4], respectively.

To model survival data using a parametric approach, choose a suitable baseline distribution that captures relevant observations' properties [25]. Traditionally, logistic, Weibull, or log-normal distributions have been employed as baseline distributions. However, these distributions may not adequately capture the heterogeneity and complexity of survival patterns and hazard shapes observed in real-world scenarios [28].

To address these limitations, there is a growing need for flexible extensions of PO models, which are capable of capturing a more comprehensive range of survival patterns. Integrating these distributions can improve the model's capacity to reflect various hazard shapes and the impact of covariates, resulting in more precise and informative outcomes.

Motivated by these considerations, this paper proposes a flexible PO regression model by using the exponentiated-Weibull (EW) distribution as a baseline distribution. The proposed model is called the exponentiated-Weibull proportional odds (EWPO) regression model. The EW distribution is a generalization of the Weibull distribution by incorporating an additional shape parameter [19]. The weighted likelihood estimation method for the EW distribution parameters was developed to provide accurate estimates, especially when the dataset contains contamination [7]. Bashir et al. [5] developed the bounded EW (BEW) distribution, designed to model datasets with support in the unit interval $[0, 1]$. A novel extension of the PHs model has been proposed by Ishag et al. [13], which incorporates the EW distribution to model the baseline HF. This new model offers greater flexibility in capturing various shapes of failure rates and can accommodate both monotonic and non-monotonic hazard patterns.

The EW distribution provides greater flexibility for modeling survival data, allowing for a diverse range of hazard rate (HR) shapes, such as monotonically increasing, decreasing, bathtub-shaped, and unimodal configurations, as demonstrated in [12, 15]. By incorporating this distribution into the PO framework, we contribute to ordinal regression and survival analysis by introducing a novel and versatile approach for analyzing ordinal response variables with complex hazard structures. This extension expands the applicability of the PO model to a broader range of research domains. Additionally, it provides a valuable tool for researchers seeking more profound insights into the relationships between covariates and ordinal outcomes.

The paper is organized as follows: Section 2 offers an overview of the PO model. Section 3 introduces the EW distribution and its theoretical background. In Section 4, the proposed PO regression model is presented. Section 5 focuses on the estimation and inference procedures for the proposed model. Section 6 presents a simulation study aimed at assessing the model's performance. The application of the model to censored real-world datasets is illustrated in Section 7. Finally, Section 8 summarizes the findings and presents some directions for future research.

2 The PO Regression Model

The proportional odds (PO) model was first introduced by Bennett [6] and is a widely used regression framework for analyzing ordinal response variables. According to Bennett [6], the PO model is comparable to Cox's PH model and can be utilized in similar situations. The multiplicative term $\exp(\beta x^{'})$ likewise provides the regression framework, but at this point, it is modeling the odds function R(t) with the corresponding baseline $R_0(t)$ for an initial level. The assumption of converging death rates given two individuals with different covariate values is analogous to assuming a constant odds ratio (OR) for the same individuals. The odds function (OF) can be written as follows,

$$R(t;\beta,x) = \frac{F(t|\beta,x)}{1-F(t|\beta,x)} = R_0(t)\exp(\beta x^{'}), \tag{1}$$

where $R_0(t)$ is the baseline odds function. The associated derivative of the odds function is given as follows,

$$r(t;\beta,x) = r_0(t)\exp(\beta x^{'}), \tag{2}$$

where $r_0(t)$ is the baseline derivative odds function.

The survival function (SF) is expressed by,

$$S(t;\beta,x) = \frac{1}{1+R_0(t)\exp(\beta x')}. \tag{3}$$

The HR function (HRF) can be written as follows,

$$h(t;\beta,x) = \frac{r_0(t)\exp(\beta x^{'})}{1+R_0(t)\exp(\beta x')}. \tag{4}$$

The probability density function (PDF) can be written as follows,

$$f(t;\beta,x) = \frac{r_0(t)\exp(\beta x^{'})}{[1+R_0(t)\exp(\beta x')]^2}. \tag{5}$$

The OR, when compared to any two individuals, such as (x_1 and x_2), can be written as follows,

$$OR(x_1, x_2, \beta) = \frac{h_0(t)\ \exp(\beta\, x_1^{'})}{h_0(t)\ \exp(\beta\, x_2^{'})} = \exp\left[\beta(x_1^{'} - x_2^{'})\right]. \tag{6}$$

3 The EW Distribution

The baseline parametric distribution plays a crucial role in capturing the diverse HR shapes of the PH model. This paper offers an overview of the EW distribution, which is both flexible and widely utilized in survival analysis. The EW distribution serves as a generalization of the Weibull distribution by incorporating an additional shape parameter [19]. The EW regression model for time-to-event data, within the framework of the AFT model, was developed by [15]. The EW distribution has several advantages over other parametric distributions, as it can accommodate a broad range of shapes and various survival patterns [23, 13]. Many conventional distributions that fit within the PH framework, such as the exponential, Gompertz, and Weibull distributions, struggle to model unimodal and bathtub-shaped HRs. Therefore, exploring distributions that can effectively manage both monotonic and non-monotonic HRs is a worthwhile pursuit.

Consider a random variable T that follows the exponentiated-Weibull (EW) distribution. The PDF, HRF, cumulative distribution function (CDF), SF, and cumulative HRF (CHRF) of T are defined as follows:

The PDF of the EW has the form,

$$f(t) = \rho\lambda\upsilon(\lambda t)^{\rho-1}\left(1 - \exp(-(\lambda t)^{\rho})\right)^{\upsilon-1}\exp(-(\lambda t)^{\rho}), \qquad t > 0. \tag{7}$$

The HRF is given by,

$$h(t) = \frac{\rho\lambda\upsilon(\lambda t)^{\rho-1}(1 - \exp(-(\lambda t)^{\rho}))^{\upsilon-1}\exp(-(\lambda t)^{\rho})}{1 - (1 - \exp(-(\lambda t)^{\rho}))^{\upsilon}}. \tag{8}$$

The CDF of the EW model reduces to

$$F(t) = (1 - \exp(-(\lambda t)^{\rho}))^{\upsilon}. \tag{9}$$

The SF takes the form,

$$S(t) = 1 - (1 - \exp(-(\lambda t)^{\rho}))^{\upsilon}. \tag{10}$$

The CHRF of the EW model is

$$H(t) = -\log(1 - (1 - \exp(-(\lambda t)^{\rho}))^{\upsilon}). \tag{11}$$

The OF of the EW distribution and its derivative are defined by,

$$R(t) = \frac{F(t)}{S(t)} = (1 - \exp(-(\lambda t)^{\rho}))^{\upsilon} - 1 \tag{12}$$

and

$$r(t) = \frac{dR(t)}{d(t)} = \frac{h(t)}{S(t)} = \rho\lambda\upsilon(\lambda t)^{\rho-1}(1 - \exp(-(\lambda t)^{\rho}))^{\upsilon-1}\exp(-(\lambda t)^{\rho}). \tag{13}$$

where ρ and υ are positive shape parameters and λ is a positive scale parameter.

It is important to note that setting $\upsilon = 1$ simplifies the EW distribution to the Weibull distribution. Mudholkar and Srivastava [18] demonstrated that the HRF exhibits the following characteristics:

(i) it is monotone increasing when $\rho \geq 1$ and $\rho\upsilon \geq 1$,

(ii) it is monotone decreasing when $\rho \leq 1$ and $\rho\upsilon \leq 1$,

(iii) it is unimodal when $\rho < 1$ and $\rho, \upsilon > 1$,

(iv) it is bathtub-shaped when $\rho > 1$ and $\rho\upsilon < 1$.

Figure 1 illustrates the HRF shapes of the EW model, which accommodate constant, increasing, decreasing, bathtub, and unimodal.

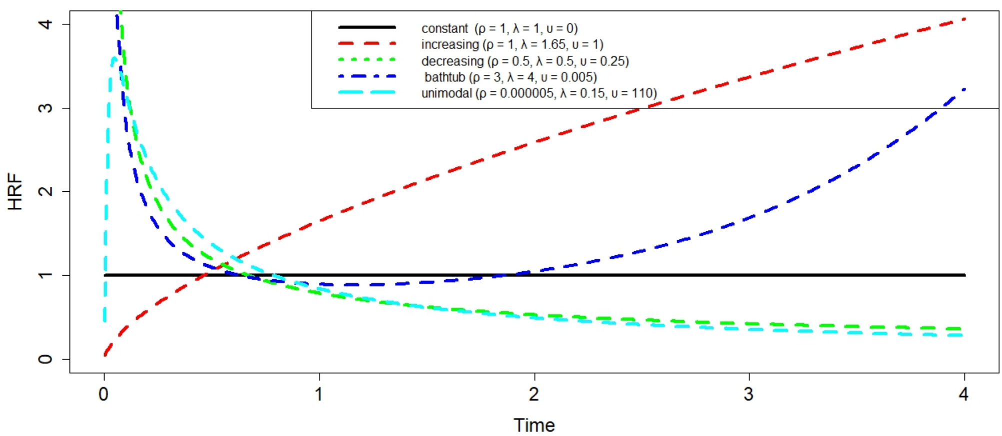


Figure 1: Demonstrate the shapes of the HRF for the EW distribution using various scale and shape parameter values.

4 The Proposed EWPO Regression Model

The proposed PO regression model can be developed by incorporating the covariates into the EW $(\rho, \lambda, \upsilon)$ distribution. The OF of the EWPO regression model is

$$R_{EWPO}(t;x) = R_0(t)\exp(\beta x^{'}) = [(1-\exp(-(\lambda t)^{\rho}))^{\upsilon} - 1]\exp(\beta x^{'}). \tag{14}$$

The first derivative of the OF reduces to

$$r_{EWPO}(t;x) = r_0(t)\exp(\beta x^{'}) = \rho\lambda\upsilon(\lambda t)^{\rho-1}(1-\exp(-(\lambda t)^{\rho}))^{\upsilon-1}\exp(-(\lambda t)^{\rho})\exp(\beta x^{'}). \tag{15}$$

The HRF, SF, and CHRF of the EWPO model are defined as follows,

$$\begin{aligned} h_{EWPO}(t;x) &= \frac{r_0(t)\exp(\beta x^{'})}{1+R_0(t)\exp(\beta x')} \\ &= \frac{\rho\lambda\upsilon(kt)^{\rho-1}(1-\exp(-(\lambda t)^{\rho}))^{\upsilon-1}\exp(-(\lambda t)^{\rho})\exp(\beta x^{'})}{1+[(1-\exp(-(\lambda t)^{\rho}))^{\upsilon}-1]\exp(\beta x')}, \end{aligned} \tag{16}$$

$$\begin{aligned} S_{EWPO}(t;x) = S(t|\beta,x) &= \frac{1}{1+R_0(t)\exp(x'\beta)} \\ &= \frac{1}{1+[(1-\exp(-(\lambda t)^{\rho}))^{\upsilon}-1]\exp(\beta x')}, \end{aligned} \tag{17}$$

and

$$H_{EWPO}(t;x) = -\log\left[\frac{1}{1+[(1-\exp(-(\lambda t)^{\rho}))^{\upsilon}-1]\exp(\beta x')}\right]. \tag{18}$$

5 Estimation of Parameters

In this section, the maximum likelihood (ML) estimation is applied to estimate the parameters of the EWPO model. Let n independent individuals have lifetimes, say, T_i, censoring time C_i for individual i, and the covariate of interest X_i. Assuming the data are subjected to the right censoring, then it is found that $t_i = \min(T_i, C_i)$, refers to the observed time of the occurrence, and $\delta_i = I(T_i, C_i)$ refers the censorship indicator based on the observed data.

Assuming that the censored observations for individual i are represented as (t_i, δ_i, x_i), where $i = 1, 2, 3, \ldots, n$, the likelihood function for the EWPO model is expressed as follows,

$$\begin{aligned} L(t;\phi,\beta) &= \prod_{i=1}^{n}[S(t_i;\phi,\beta,x_i)]^{1-\delta_i}\,[f(t_i;\phi,\beta,x_i)]^{\delta_i} = \prod_{i=1}^{n}\Bigg[S(t_i;\phi,\beta,x_i)\,[h(t_i;\phi,\beta,x_i)\Bigg]^{\delta_i}, \\ &= \prod_{i=1}^{n}\left[[S(t_i;\phi,\beta,x_i)]^{1-\delta_i}\,[\frac{h(t_i;\phi,\beta,x_i}{S(t_i;\phi,\beta,x_i)}\right]^{\delta_i} \\ &= \prod_{i=1}^{n}\left[\frac{r(t_i;\phi,\beta,x_i)}{1+R(t_i;\phi,\beta,x_i)}\right]^{\delta_i}\frac{1}{1+R(t_i;\phi,\beta,x_i)} \\ &= \prod_{i=1}^{n}\left[\frac{r_0(t)\exp(\beta x_i')}{1+R_0(t)\exp(\beta x_i')}\right]^{\delta_i}\frac{1}{1+R_0(t)\exp(\beta x_i')}, \end{aligned} \tag{19}$$

where $\phi = (\rho, \lambda, \upsilon)$ represents a vector of parameters of the baseline distributions, x_i is a covariate, and β refers to regression coefficients. Based on (19), the log-likelihood function for a parametric EWPO regression model can be written as,

$$\ell(t;\phi,\beta) = \sum_{i=1}^{n}\delta_i\left[\log(h_0(t_i;\phi,\beta,x_i))\right] - \sum_{i=1}^{n}H_0(t_i;\phi,\beta,x_i). \tag{20}$$

Using the HRF and CHRF in (16) and (18), respectively, the log-likelihood function for the EWPO regression model simplifies to,

$$\ell(t;\phi,\beta) = \sum_{i=1}^{n}\delta_i\left[\log\left(\frac{\rho\lambda\upsilon(\lambda t_i)^{\rho-1}(1-\exp(-(\lambda t_i)^{\rho}))^{\upsilon-1}\exp(-(\lambda t_i)^{\rho})\exp(\beta x_i')}{1+[(1-\exp(-(\lambda t_i)^{\rho}))^{\upsilon}-1]\exp(\beta x_i')}\right)\right] \tag{21}$$

$$-\sum_{i=1}^{n}\left[\frac{1}{1+[(1-\exp(-(\lambda t_i)^{\rho}))^{\upsilon}-1]\exp(\beta x_i')}\right]. \tag{22}$$

In this case, let us assume that, $w_i = 1-\exp(-(\lambda t_i)^{\rho})$, $c_i = \exp(-(\lambda t_i)^{\rho})$ and $q_i = \exp(\beta x_i')$, then (21) reduces to:

$$\ell(t;\phi,\beta) = \sum_{i=1}^{n}\delta_i\left[\log\left(\frac{\rho\lambda\upsilon(\lambda t_i)^{\rho-1}(w_i)^{\upsilon-1}(c_i)(q_i)}{1+[q_i w_i^{\upsilon}-q_i]}\right)\right] - \sum_{i=1}^{n}\left[\frac{1}{1+[q_i w_i^{\upsilon}-q_i]}\right]. \tag{23}$$

To obtain the maximum likelihood estimators $\hat{\varphi} = (\hat{\rho}, \hat{\lambda}, \hat{\upsilon})$ and $\hat{\beta}$, (23) can be maximized directly with respect to ρ, λ, υ and β using the Newton-Raphson optimization method. The first derivatives of $\ell(\phi, \beta; t)$ are expressed as follows:

$$\begin{aligned}\frac{\partial \ell(\phi)}{\partial \rho} &= \log(\lambda)\left(\sum_{i=1}^{n} \frac{\delta_i}{1 + q_i\left(w_i\right)^{\upsilon} - q_i}\right) + \left(\sum_{i=1}^{n} \frac{\delta_i \log\left(t_i\right)}{1 + q_i\left(w_i\right)^{\upsilon} - q_i}\right) \\ &\quad + \left(\sum_{i=1}^{n} \frac{\delta_i}{\rho(1 + q_i\left(w_i\right)^{\upsilon} - q_i)}\right),\end{aligned} \tag{24}$$

$$\begin{aligned}\frac{\partial \ell(\phi)}{\partial \upsilon} &= \sum_{i=1}^{n}\left(\frac{\delta_i\left(\rho\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i + \rho\upsilon\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} \log\left(w_i\right) c_i q_i\right) \lambda^{-\rho} t_i^{1-\rho} w_i^{-\upsilon+1}}{\rho\upsilon c_i q_i\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)}\right) \\ &\quad - \sum_{i=1}^{n}\left(\frac{\delta_i \log\left(\rho\upsilon\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i\right) q_i\left(w_i\right)^{\upsilon} \log\left(q_i\left(w_i\right)\right)}{\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^2}\right) \\ &\quad + \sum_{i=1}^{n}\left(\frac{q_i\left(w_i\right)^{\upsilon} \log\left(q_i\left(w_i\right)\right)}{\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^2}\right),\end{aligned} \tag{25}$$

$$\frac{\partial \ell(\varphi)}{\partial \lambda} = \sum_{i=1}^{n}\left(\frac{\delta_i \rho}{\lambda\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)}\right), \tag{26}$$

and

$$\begin{aligned}\frac{\partial \ell(\varphi)}{\partial \beta} &= \sum_{i=1}^{n}\left(\frac{\delta_i}{q_i\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)}\right) + \sum_{i=1}^{n}\left(\frac{\delta_i \log\left(\rho\upsilon\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i\right)}{\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^2}\right) \\ &\quad - \sum_{i=1}^{n}\left(\frac{1}{\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^2}\right).\end{aligned} \tag{27}$$

The second derivatives of $\ell(\phi, \beta; t)$ are given as follows,

$$\frac{\partial^2 \ell(\phi)}{\partial \rho^2} = \sum_{i=1}^{n}\left(-\frac{\delta_i}{\rho^2\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)}\right), \tag{28}$$

$$\begin{aligned}\frac{\partial^2 \ell(\phi)}{\partial \upsilon^2} &= \sum_{i=1}^{n}\left(\frac{-q_i\left(w_i\right)^{\upsilon} \log\left(q_i\left(w_i\right)\right)^2 \upsilon^2\left(q_i\left(w_i\right)^{\upsilon} - 1 + q_i\right) \log\left(\rho\upsilon\,\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i\right) \upsilon^2\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^3}{\upsilon^2\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^3}\right) \\ &\quad + \sum_{i=1}^{n}\left(\frac{\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)\left(\left(\frac{1}{2} + \left(\upsilon^2 \log\left(w_i\right) + \upsilon\right) \log\left(q_i\left(w_i\right)\right)\right) q_i\left(w_i\right)^{\upsilon} - \frac{q_i}{2} + \frac{1}{2}\right)}{\left(\upsilon^2\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)^3\right) \delta_i}\right) \\ &\quad - \left(\sum_{i=1}^{n}\left(-\frac{q_i\left(w_i\right)^{\upsilon} \log\left(q_i\left(w_i\right)\right)^2\left(q_i\left(w_i\right)^{\upsilon} - 1 + q_i\right)}{\left(-1 - q_i\left(w_i\right)^{\upsilon} + q_i\right)^3}\right)\right),\end{aligned} \tag{29}$$

$$\frac{\partial^2 \ell(\varphi)}{\partial \lambda^2} = \sum_{i=1}^{n}\left(-\frac{\delta_i \rho}{\lambda^2\left(1 + q_i\left(w_i\right)^{\upsilon} - q_i\right)}\right), \tag{30}$$

$$
\begin{aligned}
\frac{\partial^2 \ell(\varphi)}{\partial \beta^2} =& \sum_{i=1}^{n}\left(-\frac{\delta_i}{q_i^2\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)}\right)+\sum_{i=1}^{n}\left(\frac{2\delta_i}{q_i\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}\right) \\
&+\sum_{i=1}^{n}\left(\frac{2\delta_i \log\left(\rho\upsilon\,\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i\right)}{\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^3}\right)-\left(\sum_{i=1}^{n}\frac{2}{\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^3}\right),
\end{aligned}
\tag{31}
$$

$$
\frac{\partial^2 \ell(\varphi)}{\partial \rho \partial \lambda} = \sum_{i=1}^{n}\frac{\delta_i}{\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)\lambda},
\tag{32}
$$

$$
\frac{\partial^2 \ell(\varphi)}{\partial \upsilon \partial \lambda} = \sum_{i=1}^{n}\left(-\frac{\delta_i \rho q_i\left(w_i\right)^{\upsilon}\log\left(q_i\left(w_i\right)\right)}{\lambda\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}\right),
\tag{33}
$$

$$
\frac{\partial^2 \ell(\varphi)}{\partial \rho \partial \upsilon} = \sum_{i=1}^{n}\left(-\frac{\delta_i q_i\left(w_i\right)^{\upsilon}\log\left(q_i\left(w_i\right)\right)\left(\rho\log\left(t_i\right)+\rho\log\left(\lambda\right)+1\right)}{\rho\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}\right),
\tag{34}
$$

$$
\begin{aligned}
\frac{\partial^2 \ell(\varphi)}{\partial \beta \partial \upsilon} =& \sum_{i=1}^{n}\left(\frac{\delta_i\left(\rho\,\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i+\rho\upsilon\,\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1}\log\left(w_i\right) c_i q_i\right)\lambda^{-\rho} t_i^{1-\rho} w_i^{1-\upsilon}}{\rho\upsilon c_i q_i\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}\right) \\
&-\sum_{i=1}^{n}\left(\frac{2\delta_i\log\left(\rho\upsilon\,\lambda^{\rho} t_i^{\rho-1} w_i^{\upsilon-1} c_i q_i\right) q_i\left(w_i\right)^{\upsilon}\log\left(q_i\left(w_i\right)\right)}{\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^3}\right) \\
&-\sum_{i=1}^{n}\left(\frac{\delta_i q_i\left(w_i\right)^{\upsilon}\log\left(q_i\left(w_i\right)\right)}{q_i\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}\right)+\sum_{i=1}^{n}\left(\frac{2q_i\left(w_i\right)^{\upsilon}\log\left(q_i\left(w_i\right)\right)}{\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^3}\right),
\end{aligned}
\tag{35}
$$

$$
\frac{\partial^2 \ell(\varphi)}{\partial \beta \partial \lambda} = \sum_{i=1}^{n}\frac{\delta_i\rho}{\lambda\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2},
\tag{36}
$$

and

$$
\frac{\partial^2 \ell(\varphi)}{\partial \beta \partial \rho} = \sum_{i=1}^{n}\frac{\delta_i\left(\rho\log\left(t_i\right)+\rho\log\left(\lambda\right)+1\right)}{\rho\left(1+q_i\left(w_i\right)^{\upsilon}-q_i\right)^2}.
\tag{37}
$$

6 Simulation Study

This section provides a simulation study to illustrate the inferential characteristics of the proposed EWPO regression model. The Akaike information criterion (AIC) is obtained to show how to select the best models that accurately capture the basic HR shapes and the impact of censored percentages on the model characteristics.

Assuming the EWPO regression model in (16). The covariates are considered: two binary covariates (x_1 and x_3) are generated from the Bernoulli distribution with a probability of 0.5, and another continuous covariate (x_2) is generated using the standard normal distribution.

The inverse transform function of the EW distribution is utilized to generate survival times, allowing for the accommodation of all fundamental HR shapes.

Let us recall the CHRF of the EW model, which is given by,

$$H(t;\rho,\lambda,\upsilon) = -\log\left[1-(1-e^{-(\lambda t)^{\rho}})^{\upsilon}\right]. \tag{38}$$

The inverse of the CHRF follows as,

$$H^{-1}(t;\rho,\lambda,\upsilon) = -\frac{\log\left[(e^{-t}-1)^{\frac{1}{\upsilon}}-1\right]^{\frac{1}{\rho}}}{\lambda}. \tag{39}$$

The simulation study concentrated on evaluating the performance and accuracy of the proposed model's estimators, specifically assessing the average bias (AB), standard error (SE), and mean squared error (MSE). The simulation's findings were derived with 100, 300, 500, and 2000 samples for each parameter value, with approximately 30% and 20% censoring, respectively.

6.1 Simulation scenarios

Four simulation scenarios are conducted to evaluate the performance of the proposed EWPO model and compare it with other models, such as the WPO and LLPO regression models, based on different HRFs, including increasing, decreasing, bathtub, and unimodal. The goal is to investigate how the HR shape specification affects the inferential aspects of the PO model. The lifetime data in the following four scenarios are generated using the EW model.

Scenario 1: Increasing HRF
The lifetime data in this scenario are obtained via the EW model using the parameter values for $\rho = 1.65$, $\lambda = 1.20$, and $\upsilon = 1.0$. The censored data are generated, assuming that administrative censoring Tc at different time point values,

(i) Tc = 4 and (ii) Tc = 7,

which resulted in approximately 30% and 20% censoring, respectively.

Scenario 2: Decreasing HRF
The lifetime data in this scenario are obtained via the EW model using the parameter values for $\rho = 0.70$, $\lambda = 0.60$, and $\upsilon = 1.0$. The censored data are generated, assuming that administrative censoring Tc at different time point values,

(i) Tc = 14 and (ii) Tc = 8,

which resulted in approximately 20% and 30% censoring, respectively.

Scenario 3: Bathtub HRF
The lifetime data in this scenario are obtained via the EW model using the parameter values for $\rho = 4$, $\lambda = 7$, and $\upsilon = 0.08$. The censored data are generated, assuming that administrative censoring Tc at different time point values,

(i) Tc = 7 and (ii) Tc = 3,

which resulted in approximately 20% and 30% censoring, respectively.

Scenario 4: Unimodel HRF
The lifetime data in this scenario are obtained via the EW model using the parameter values for $\rho = 0.15$, $\lambda = 0.00006$, and $\upsilon = 40$. The censored data are generated, assuming that administrative censoring Tc at different time point values

(i) Tc = 12 and (ii) Tc = 5,
which resulted in approximately 20% and 30% censoring, respectively.

6.2 Simulation results

Table 1 shows the results for the EWPO regression model, which includes the mean estimate (estimate), SE, AB, MSE, and confidence interval (CI) for the ML approach. The averages of the estimates are similar, and both the SE and MSE tend to decrease with larger sample sizes. Furthermore, as the sample sizes increase, the estimates for all assessed parameters show improved performance.

Table 1: Simulation outcomes including the true values, estimates, SE, AB, MSE, and 95% CI for the parameters of the for EWPO regression model.

					$n = 100$						
		20% Censoring					30% Censoring				
Parameter	True	Estimate	SE	AB	MSE	CI 95 %	Estimate	SE	AB	MSE	CI 95 %
β_1	0.25	0.323	0.354	0.073	0.042	(-0.371, 1.017)	0.312	0.354	0.062	0.035	(-0.382, 1.006)
β_2	0.35	0.316	0.138	-0.033	0.022	(0.046, 0.587)	0.311	0.137	-0.039	0.026	(0.042, 0.579)
β_3	0.45	0.997	0.363	0.547	0.792	(0.286, 1.709)	0.970	0.363	0.521	0.740	(0.259, 1.682)
λ	1.20	0.904	0.349	-0.296	0.622	(0.220, 1.588)	0.795	0.363	-0.404	0.807	(0.084 , 1.507)
ρ	1.65	1.250	0.376	-0.400	1.159	(0.513, 1.987)	1.119	0.356	-0.531	1.469	(0.422 , 1.817)
υ	1.00	1.529	0.755	0.530	1.340	(0.050, 3.010)	1.810	0.978	0.811	2.279	(-0.106, 3.728)
					$n = 300$						
		20% Censoring					30% Censoring				
Parameter	True	Estimate	SE	AB	MSE	CI 95 %	Estimate	SE	AB	MSE	CI 95 %
β_1	0.25	0.632	0.202	0.383	0.338	(0.237, 1.028)	0.643	0.203	0.394	0.352	(0.246, 1.041)
β_2	0.35	0.349	0.075	-0.001	0.001	(0.202, 0.496)	0.353	0.075	0.003	0.002	(0.206, 0.500)
β_3	0.45	0.942	0.203	0.493	0.686	(0.545, 1.340)	0.958	0.204	0.509	0.717	(0.559, 1.359)
λ	1.20	1.074	0.187	-0.126	0.286	(0.708, 1.441)	1.177	0.202	-0.022	0.053	(0.782, 1.573)
ρ	1.65	1.307	0.182	-0.342	1.013	(0.951, 1.664)	1.460	0.242	-0.190	0.590	(0.986, 1.935)
υ	1.00	1.499	0.336	0.500	1.250	(0.841, 2.159)	1.290	0.320	0.290	0.664	(0.663, 1.917)
					$n = 500$						
		20% Censoring					30% Censoring				
Parameter	True	Estimate	SE	AB	MSE	CI 95 %	Estimate	SE	AB	MSE	CI 95 %
β_1	0.25	0.276	0.155	0.027	0.014	(-0.027, 0.581)	0.383	0.156	0.133	0.084	(0.077, 0.689)
β_2	0.35	0.415	0.058	0.065	0.050	(0.302, 0.529)	0.470	0.057	0.120	0.099	(0.359, 0.582)
β_3	0.45	0.447	0.156	-0.003	0.002	(0.142, 0.753)	0.719	0.158	0.270	0.316	(0.410, 1.029)
λ	1.20	0.863	0.144	-0.337	0.695	(0.581, 1.145)	1.095	0.155	-0.104	0.239	(0.792, 1.400)
ρ	1.65	1.210	0.142	-0.439	1.257	(0.932, 1.489)	1.449	0.186	-0.200	0.620	(1.085, 1.814)
υ	1.00	1.648	0.337	0.649	1.718	(0.988, 2.309)	1.278	0.251	0.279	0.635	(0.787, 1.771)

The key finding from Scenario 1, as demonstrated in Tables 2 and 3, is the significant advantages of the proposed EWPO model over others. This is not just a marginal improvement but a clear and substantial advancement in hazard modeling. The lower AIC values of the EWPO model, which indicate its superior performance, are a testament to this. The SE, AB, and MSE values further reinforce this, showing that our model consistently outperforms the others. The impact of sample size and censoring percentage on the model's accuracy is also crucial to the research. As the censoring percentage increased, the proposed EWPO model consistently outperforms the WPO and LLPO models. Figure 2 shows that all the models are equally integrated into the increasing HRF, but in the case of heavy censoring, our proposed model stood out as the best performer, demonstrating its robustness and adaptability.

Table 2: Simulation outcomes for Scenario 1 (increasing HRF) with $n = 100$ and approximately 20% and 30% censoring respectively to compare model performance.

Model	Parameter	True value	20% Censoring MLE	SE	AB	MSE	AIC	30% Censoring MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.275	0.357	0.025	0.013	152.892	0.818	0.359	0.569	0.608	146.320
	β_2	0.35	0.323	0.127	-0.026	0.018		0.165	0.150	-0.184	0.095	
	β_3	0.45	-0.207	0.350	-0.657	0.159		0.687	0.354	0.237	0.270	
	λ	1.20	0.453	0.274	-0.746	1.235		0.851	0.379	-0.348	0.715	
	ρ	1.65	0.905	0.265	-0.745	1.904		1.165	0.388	-0.484	1.364	
	υ	1.00	2.543	1.554	1.543	5.471		1.777	0.963	0.777	2.161	
WPO	β_1	0.25	1.088	0.353	0.838	1.122	280.394	1.547	0.363	1.297	2.333	249.54
	β_2	0.35	0.342	0.128	-0.007	0.005		0.113	0.157	-0.237	0.110	
	β_3	0.45	0.694	0.342	0.244	0.280		1.364	0.356	0.914	1.658	
	λ	1.20	1.550	0.173	0.350	0.963		1.808	0.184	0.608	1.831	
	ρ	1.65	1.263	0.112	-0.386	1.126		1.417	0.129	-0.232	0.714	
LLPO	β_1	0.25	1.109	0.353	0.860	1.170	283.356	1.564	0.365	1.314	2.386	251.549
	β_2	0.35	0.320	0.126	-0.029	0.020		0.093	0.157	-0.256	0.114	
	β_3	0.45	0.707	0.343	0.257	0.298		1.371	0.358	0.921	1.680	
	λ	1.20	1.142	0.149	-0.057	0.136		1.411	0.169	0.211	0.553	
	ρ	1.65	1.574	0.134	-0.075	0.243		1.663	0.144	0.013	0.046	

Table 3: Simulation outcomes for Scenario 1 (Increasing HRF) with $n = 2000$ and approximately 20% and 30% censoring, respectively, to compare model performance.

Model	Parameter	True value	20% Censoring MLE	SE	AB	MSE	AIC	30% Censoring MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.385	0.077	0.135	0.086	3119.944	0.396	0.078	0.146	0.095	3035.441
	β_2	0.35	0.564	0.031	0.214	0.196		0.575	0.031	0.225	0.209	
	β_3	0.45	0.681	0.078	0.231	0.261		0.699	0.078	0.249	0.287	
	λ	1.20	0.854	0.067	-0.345	0.710		0.956	0.075	-0.243	0.525	
	ρ	1.65	1.172	0.062	-0.477	1.348		1.298	0.081	-0.351	1.037	
	υ	1.00	1.888	0.177	0.888	2.565		1.617	0.166	0.617	1.615	
WPO	β_1	0.25	1.0725	0.077	0.822	1.088	5469.276	1.055	0.078	0.805	1.051	5337.317
	β_2	0.35	0.570	0.031	0.220	0.203		0.583	0.032	0.233	0.218	
	β_3	0.45	1.401	0.078	0.951	1.761		1.392	0.078	0.942	1.737	
	λ	1.20	1.860	0.039	0.660	2.020		1.813	0.037	0.613	1.849	
	ρ	1.65	1.446	0.028	-0.203	0.631		1.474	0.030	-0.175	0.548	
LLPO	β_1	0.25	0.527	0.078	0.839	1.125	5539.723	1.066	0.078	0.816	1.076	5337.317
	β_2	0.35	1.402	0.030	0.177	0.156		0.536	0.030	0.186	0.166	
	β_3	0.45	0.450	0.079	0.952	1.764		1.392	0.078	0.942	1.737	
	λ	1.20	1.408	0.034	0.208	0.543		1.395	0.034	0.195	0.506	
	ρ	1.65	1.780	0.034	0.130	0.449		1.775	0.034	0.125	0.430	

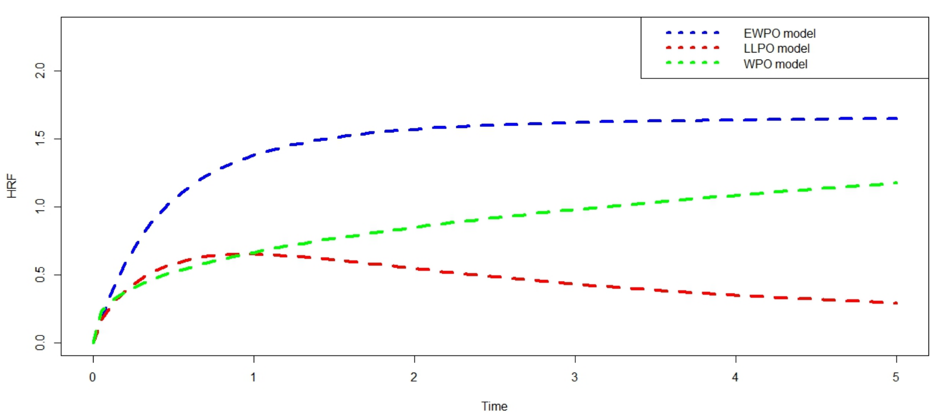


Figure 2: Estimated HRFs of Scenario 1 for the competing baseline HRFs.

According to the findings of Scenario 2, which are presented in Tables 4 and 5, all competing models can integrate the decreasing HRF shape. However, based on the AIC, the proposed EWPO model outperforms the others. It surpasses the WPO and LLPO models and the genuine produced model regarding SE, AB, and MSE.

Table 4: Simulation outcomes for Scenario 2 (Decreasing HRF) with $n = 100$ and approximately 20% and 30% censoring respectively to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.205	0.359	-0.044	0.020		0.407	0.370	0.157	0.103	
	β_2	0.35	0.344	0.137	-0.005	0.004		0.384	0.136	0.034	0.025	
	β_3	0.45	0.304	0.359	-0.145	0.110	50.538	0.237	0.363	-0.212	0.146	43.660
	λ	0.60	0.156	0.178	-0.443	0.336		0.053	0.088	-0.546	0.357	
	ρ	0.70	0.432	0.125	-0.267	0.303		0.329	0.101	-0.370	0.382	
	υ	1.00	1.334	0.718	0.334	0.780		2.209	1.406	1.209	3.880	
WPO	β_1	0.25	1.107	0.350	0.857	1.163		1.147	0.367	0.897	1.254	
	β_2	0.35	0.297	0.132	-0.052	0.034		0.345	0.137	-0.004	0.003	
	β_3	0.45	1.137	0.352	0.687	1.091	133.877	1.234	0.352	0.784	1.320	120.298
	λ	0.60	1.302	0.306	0.702	1.337		1.445	0.361	0.845	1.730	
	ρ	0.70	0.552	0.049	-0.147	0.184		0.540	0.048	-0.159	0.198	
LLPO	β_1	0.25	1.115	0.352	0.865	1.182		1.189	0.370	0.939	1.352	
	β_2	0.35	0.255	0.128	-0.094	0.057		0.310	0.132	-0.039	0.026	
	β_3	0.45	1.116	0.354	0.666	1.043	137.909	1.191	0.354	0.741	1.217	120.339
	λ	0.60	0.643	0.180	0.043	0.054		0.677	0.196	0.077	0.099	
	ρ	0.70	0.674	0.058	-0.025	0.034		0.669	0.058	-0.030	0.041	

Table 5: Simulation outcomes for Scenario 2 (Decreasing HRF) with $n = 2000$ and approximately 20% and 30% censoring respectively to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.183	0.079	-0.066	0.029		0.184	0.079	-0.065	0.029	
	β_2	0.35	0.582	0.031	0.232	0.217		0.582	0.031	0.232	0.217	
	β_3	0.45	0.478	0.080	0.028	0.027	1702.950	0.479	0.080	0.029	0.028	1619.241
	λ	0.60	0.220	0.048	-0.379	0.311		0.223	0.050	-0.376	0.310	
	ρ	0.70	0.468	0.028	-0.231	0.270		0.470	0.029	-0.229	0.268	
	υ	1.00	1.397	0.152	0.397	0.954		1.387	0.156	0.387	0.924	
WPO	β_1	0.25	1.072	0.078	0.822	1.089		1.067	0.078	0.817	1.077	
	β_2	0.35	0.575	0.031	0.225	0.209		0.579	0.031	0.229	0.214	
	β_3	0.45	1.403	0.078	0.954	1.769	3338.996	1.399	0.078	0.949	1.757	3204.628
	λ	0.60	1.663	0.081	1.063	2.407		1.636	0.080	1.036	2.319	
	ρ	0.70	0.618	0.012	-0.081	0.108		0.620	0.012	-0.079	0.105	
LLPO	β_1	0.25	1.087	0.078	0.837	1.121		1.081	0.078	0.832	1.108	
	β_2	0.35	0.528	0.031	0.178	0.156		0.533	0.030	0.183	0.162	
	β_3	0.45	1.400	0.078	0.950	1.758	3413.309	1.394	0.079	0.944	1.742	3271.891
	λ	0.60	0.871	0.050	0.271	0.400		0.867	0.050	0.267	0.392	
	ρ	0.70	0.755	0.015	0.055	0.081		0.754	0.015	0.054	0.079	

The results from Scenario 2, detailed in Tables 4 and 5, indicate that all competing models can accommodate the decreasing HRF shape. However, according to the AIC, the proposed EWPO model outperforms the others, including the WPO and LLPO models, as well as the actual generated model, in terms of SE, AB, and MSE. Additionally, our proposed model proves to be the most suitable option as censoring increases and effectively handles heavy censoring. Figure 3 demonstrates that all models are similarly effective in integrating the decreasing HRF. Moreover, the EWPO model remains the most appropriate choice when censoring increases, and it makes a wise decision regarding heavy censoring. Figure 3 illustrates that all the models are equally integrated into the decreasing HRF. However, in the case of heavy censoring, our proposed model emerged as the best performer.

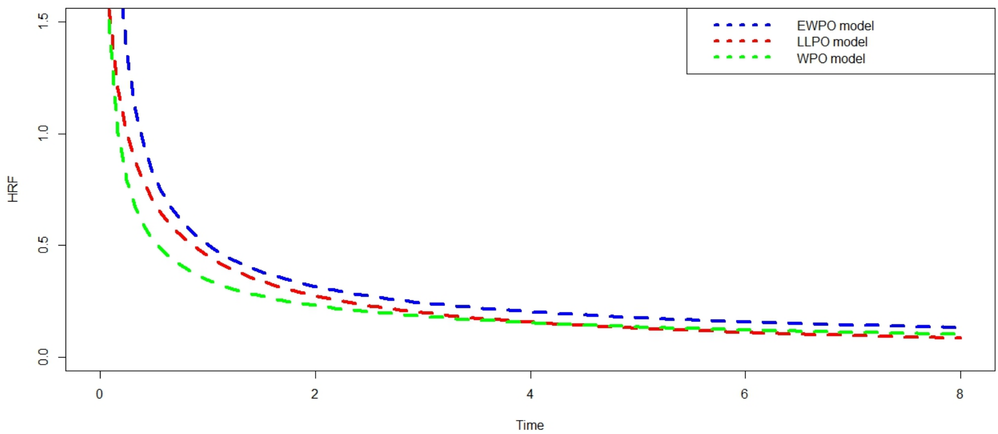


Figure 3: Estimated HRFs of Scenario 2 for the competing baseline HRFs.

Based on the findings presented in Tables 6 and 7, Scenario 3 highlights a significant advantage of the EWPO model. It is the only model consistently delivering the lowest values for SE, AB, and MSE, even in challenging conditions of heavy censoring and bathtub hazards. This adaptability is a crucial and essential aspect of our model, which makes it a reliable tool for hazard modeling in real-world scenarios. As expected, the EWPO model had the least accurate estimates for AB, SE, and MSE in Scenario 3, further underscoring the superiority of our proposed model. Figure 4 illustrates that only the EWPO model is capable of integrating the bathtub HRF, while the other models fail to accommodate the bathtub HRF. In this scenario, the EWPO model emerged as the best performer.

Table 6: Simulation outcomes for Scenario 3 (Bathtub HRF) with $n = 100$ and approximately 20% and 30% censoring, respectively, to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
	β_1	0.25	0.004	0.351	0.246	0.062		0.085	0.352	-0.165	0.055	
EWPO	β_2	0.35	0.411	0.135	0.062	0.047		0.395	0.140	0.045	0.034	
	β_3	0.45	0.958	0.351	0.508	0.715	106.678	0.890	0.356	0.440	0.590	57.577
	λ	7.00	5.126	1.786	1.874	22.721		4.371	2.588	-2.628	29.888	
	ρ	4.00	1.970	1.357	2.029	12.116		1.155	1.047	-2.844	14.665	
	υ	0.08	0.190	0.140	0.110	0.030		0.334	0.327	0.254	0.105	
	β_1	0.25	0.861	0.333	0.611	0.680		0.803	0.338	0.554	0.583	
WPO	β_2	0.35	0.251	0.129	-0.098	0.059		0.339	0.142	-0.011	0.008	
	β_3	0.45	1.671	0.337	1.221	2.593	185.809	1.543	0.339	1.093	2.179	104.078
	λ	7.00	11.857	2.634	4.857	91.608		11.628	3.729	4.628	86.214	
	ρ	4.00	0.505	0.046	-3.494	15.744		0.479	0.048	-3.520	15.770	
	β_1	0.25	0.871	0.334	0.621	0.697		0.815	0.338	0.565	0.602	
LLPO	β_2	0.35	0.227	0.127	-0.122	0.071		0.323	0.141	-0.026	0.018	
	β_3	0.45	1.628	0.339	1.178	2.449	191.001	1.520	0.340	1.071	2.110	106.308
	λ	7.00	6.730	1.833	-0.270	3.701		6.942	2.485	-0.057	0.798	
	ρ	4.00	0.549	0.048	-3.450	15.698		0.515	0.049	-3.484	15.734	

Table 7: Simulation outcomes for Scenario 3 (Bathtub HRF) with $n = 2000$ and approximately 20% and 30% censoring respectively to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
	β_1	0.25	0.399	0.076	0.149	0.097		0.402	0.076	0.152	0.099	
EWPO	β_2	0.35	0.580	0.031	0.230	0.215		0.568	0.031	0.218	0.201	
	β_3	0.45	0.693	0.076	0.243	0.279	1692.519	0.708	0.077	0.258	0.300	1146.876
	λ	7.00	5.714	0.484	-1.285	16.347		6.0387	0.499	-0.961	12.533	
	ρ	4.00	2.000	0.267	-2.241	12.906		2.427	0.772	-1.572	10.108	
	υ	0.08	0.210	0.034	0.130	0.038		0.151	0.049	0.071	0.016	
	β_1	0.25	1.217	0.071	0.967	1.420		1.153	0.073	0.903	1.268	
WPO	β_2	0.35	0.507	0.031	0.157	0.135		0.531	0.032	0.181	0.160	
	β_3	0.45	1.505	0.072	1.055	2.063	3246.712	1.469	0.073	1.019	1.956	2375.638
	λ	7.00	14.684	0.852	7.684	166.631		14.081	0.944	7.081	149.300	
	ρ	4.00	0.493	0.010	-3.506	15.757		0.484	0.010	-3.515	15.765	
	β_1	0.25	0.485	0.073	1.137	1.474		1.158	0.074	0.908	1.280	
LLPO	β_2	0.35	1.490	0.031	0.235	0.173		0.514	0.032	0.165	0.142	
	β_3	0.45	0.450	0.073	1.140	2.098	3361.522	1.463	0.074	1.014	1.940	2462.112
	λ	7.00	8.711	0.619	1.712	26.898		8.629	0.678	1.630	25.468	
	ρ	4.00	0.530	0.010	-3.469	15.718		0.519	0.011	-3.481	15.731	

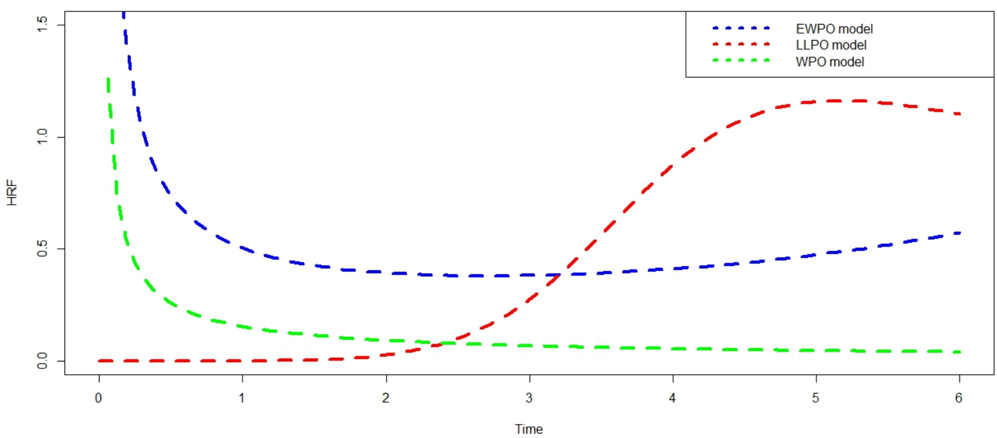


Figure 4: Estimated HRFs of Scenario 3 for the competing baseline HRFs.

The findings in Scenario 4 in Tables 8 and 9 indicate that the EWPO model produces estimates with small AB and MSR values for all regression coefficients. The model produced estimates that are similar to those of the true model, as indicated by the AIC value. Figure 5 shows that only the EWPO model can integrate the Unimode HRF, while the other models cannot accommodate it. Moreover, the proposed EWPO regression model outperforms all competing models, even in cases of heavy censorship.

Table 8: Simulation outcomes for Scenario 4 (Unimodel HRF) with $n = 100$ and approximately 20% and 30% censoring respectively to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.005	0.355	-0.245	0.062		0.108	0.359	-0.142	0.051	
	β_2	0.35	0.333	0.143	-0.017	0.012		0.461	0.137	0.111	0.090	
	β_3	0.45	0.525	0.357	0.075	0.073	170.626	0.626	0.358	0.176	0.189	131.552
	λ	0.00006	0.00001	0.001	0.000	0.000		0.0001	0.001	0.000	0.000	
	ρ	0.15	0.166	0.075	0.016	0.005		0.168	0.083	0.018	0.006	
	υ	40.00	36.302	63.611	-3.698	282.130		34.180	62.916	-5.820	431.728	
WPO	β_1	0.25	0.660	0.350	0.410	0.373		0.750	0.360	0.500	0.500	
	β_2	0.35	0.358	0.150	0.008	0.006		0.497	0.143	0.148	0.125	
	β_3	0.45	1.269	0.353	0.820	1.410	296.982	1.266	0.361	0.817	1.403	244.2114
	λ	0.00006	2.603	0.634	2.604	6.779		2.436	0.588	2.436	5.9364	
	ρ	0.15	0.552	0.049	0.402	0.282		0.585	0.053	0.435	0.320	
LLPO	β_1	0.25	0.646	0.352	0.397	0.356		0.702	0.362	0.453	0.431	
	β_2	0.35	0.357	0.148	0.008	0.005		0.490	0.141	0.141	0.118	
	β_3	0.45	1.259	0.357	0.809	1.383	290.939	1.232	0.362	0.782	1.316	238.714
	λ	0.00006	1.145	0.314	1.145	1.312		1.110	0.305	1.110	1.233	
	ρ	0.15	0.716	0.061	0.566	0.491		0.737	0.064	0.588	0.522	

Table 9: Simulation outcomes for Scenario 4 (Unimodel HRF) with $n = 2000$ and approximately 20% and 30% censoring respectively to compare model performance.

			20% Censoring					30% Censoring				
Model	Parameter	True value	MLE	SE	AB	MSE	AIC	MLE	SE	AB	MSE	AIC
EWPO	β_1	0.25	0.250	0.080	0.0001	0.0001		0.250	0.080	0.0001	0.0001	
	β_2	0.35	0.350	0.031	0.0001	0.0001		0.350	0.031	0.0001	0.0001	
	β_3	0.45	0.450	0.080	0.0001	0.0001	4539.585	0.450	0.080	0.0001	0.0001	3871.114
	λ	0.00006	0.00003	0.000	0.0001	0.0001		0.00004	0.000	0.0001	0.0001	
	ρ	0.15	0.150	0.002	0.0001	0.0001		0.150	0.002	0.0001	0.0001	
	υ	40.00	40.000	1.800	0.0001	0.0001		40.000	1.873	0.0001	0.0001	
WPO	β_1	0.25	1.067	0.078	0.817	1.077		1.036	0.079	0.787	1.012	
	β_2	0.35	0.642	0.033	0.292	0.290		0.635	0.033	0.285	0.281	
	β_3	0.45	1.420	0.079	0.970	1.814	6024.243	1.400	0.080	0.950	1.759	5152.698
	λ	0.00006	4.625	0.250	4.625	21.393		3.541	0.183	3.542	12.545	
	ρ	0.15	0.581	0.012	0.431	0.315		0.634	0.014	0.484	0.380	
LLPO	β_1	0.25	1.003	0.079	0.753	0.944		0.983	0.079	0.734	0.906	
	β_2	0.35	0.640	0.033	0.291	0.288		0.623	0.033	0.274	0.267	
	β_3	0.45	1.350	0.080	0.901	1.622	5824.980	1.338	0.080	0.888	1.588	5041.362
	λ	0.00006	1.900	0.118	1.900	3.610		1.722	0.104	1.722	2.966	
	ρ	0.15	0.744	0.015	0.595	0.532		0.772	0.016	0.623	0.575	

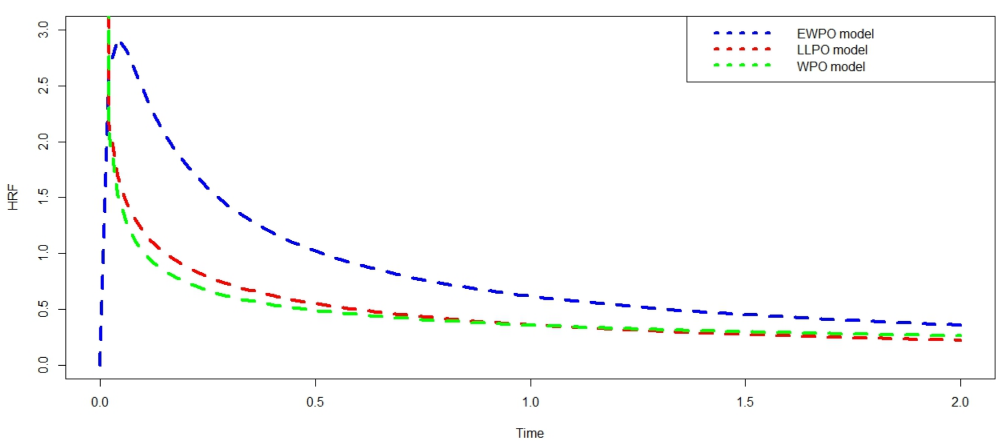


Figure 5: Estimated HRFs of Scenario 4 for the competing baseline HRFs.

7 Survival Analysis to Right-Censored Data

This section provides a comprehensive analysis of two right-censored clinical trial datasets to demonstrate the effectiveness and practicality of the proposed fully parametric EW-PO regression model for survival analysis. The EWPO regression model is compared with other PO regression models, including the generalized log-logistic PO (GLLPO), Weibull PO (WPO), and log-logistic PO (LLPO) models. The performance of these models is evaluated using two information criteria: the AIC and the Bayesian information criterion (BIC).

7.1 Dataset 1: IPASS data set

To highlight the significance of the proposed EWPO regression model, we analyze the IPASS dataset from a randomized clinical trial. This study compared gefitinib and carboplatin-paclitaxel in terms of progression-free survival for patients with advanced pulmonary adenocarcinoma.

The IPASS dataset is reconstructed and re-published by Mok et al. [17], and it is now available freely in the AHSurv R package [20]. The reconstructed data set still contains all the features mentioned in the references, and it is accessible for the clinical trial's results. The database contains data from March 2006 through April 2008. The trial aims to evaluate the effect of gefitinib compared to carboplatin-paclitaxel doublet chemotherapy on the progression-free survival (measured in months) of patients diagnosed with non-small cell lung cancer. According to the trial protocol, $1,207$ individuals from East Asia with advanced lung adenocarcinoma-who were either nonsmokers or former light smokers and had not received prior treatment-are randomly assigned to two groups. The first group comprised 608 patients who are given carboplatin + paclitaxel, while the second group included 609 patients who were given gefitinib. The dataset shows that the event of interest occurred 965 times (79.3%), with 449 occurrences (73.7%) in patients who received gefitinib and 516 (84.9%) in patients who received carboplatin + paclitaxel. The proposed fully parametric EWPO regression model will be applied to the reconstructed IPASS data to accurately assess the data and estimate the regression coefficients.

Figure 6 displays a concavity pattern in the total time on the test (TTT) plot, indicating the increasing HR shape of the data. This shows that the TTT, including the histogram, is appropriate for analyzing this dataset.

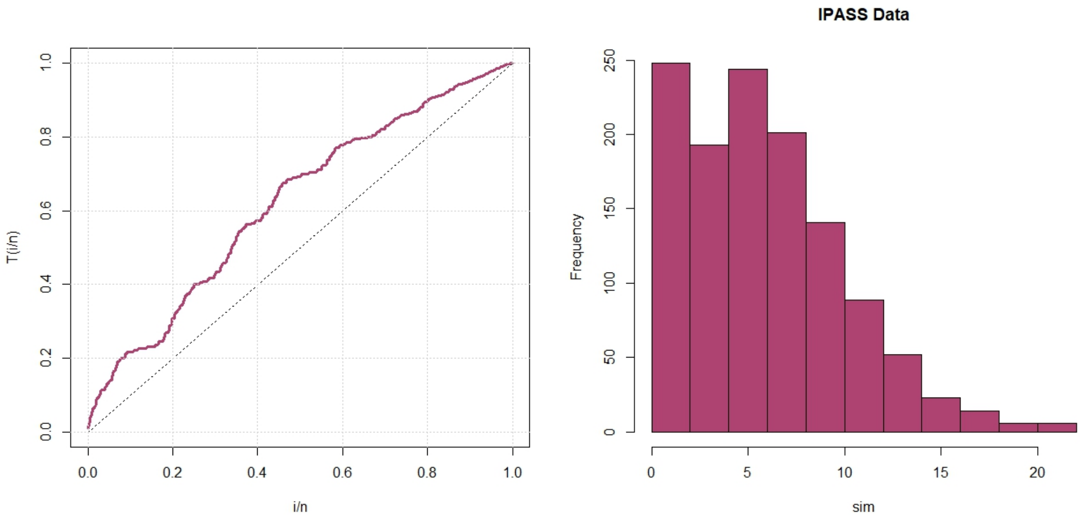


Figure 6: The TTT and histogram plots for the IPASS clinical trial dataset.

Table 10 displays the results of the PO regression models using the IPASS clinical trial data. Table 10 presents the model parameters, mean, SE, z-values, p-values, and information criterion values. Figure 7 complements Table 10 by demonstrating the fitted estimate HRFs for different models. Figure 7 and Table 10 indicate that the EWPO model outperforms all other models in terms of providing the best fit to the IPASS clinical trial data. This is demonstrated by the lowest information criterion values for the EWPO model. Additionally, the EWPO model parameters are significant at the 5% significance level. Figure 7 shows that the proposed model provides better fits to the IPASS clinical trial data over time.

Table 10: Findings of each model for IPASS dataset, along with analytical measures for various models.

Model	Parameter	Mean	SE	z-value	L 95%	U 95%	p-value	AIC	BIC
EWPO	β	-0.005	0.105	0.049	-0.211	0.200	0.090	5707.532	5727.949
	λ	5.941	1.410	4.213	3.177	8.706	0.000		
	ρ	1.077	0.184	5.842	0.716	1.439	0.000		
	υ	1.433	0.418	3.429	0.614	2.253	0.000		
GLLPO	β	-0.139	0.103	-0.184	-0.222	0.183	0.085	5710.571	5730.987
	λ	0.145	0.007	19.161	0.126	0.154	0.000		
	ρ	1.395	0.064	22.278	1.298	1.549	0.000		
	υ	0.035	0.020	1.728	-0.005	0.074	0.084		
WPO	β	0.575	0.103	5.597	0.374	0.777	0.000	6888.596	6903.909
	λ	10.641	0.342	31.159	9.972	11.311	0.000		
	ρ	1.045	0.029	35.734	0.989	1.103	0.000		
LLPO	β	0.615	0.102	6.028	0.415	0.815	0.000	6913.954	6929.267
	λ	7.440	0.265	28.112	6.922	7.960	0.000		
	ρ	1.306	0.036	36.499	1.236	1.377	0.000		

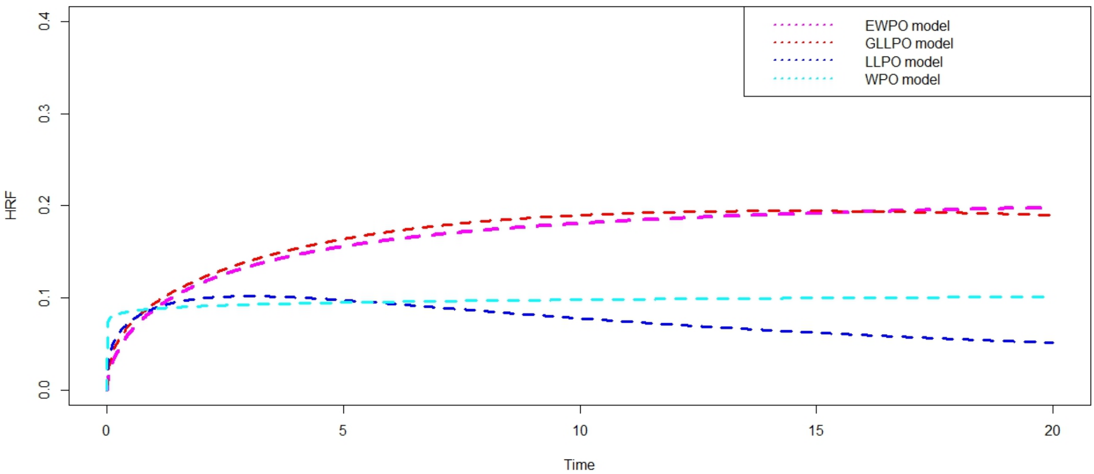
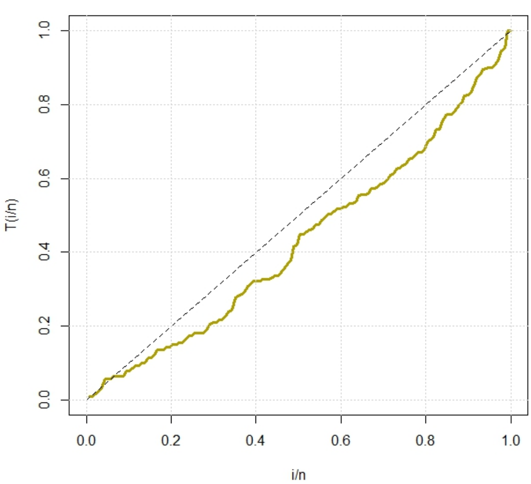


Figure 7: The estimated HRFs for the competing models based on IPASS data set.

7.2 Data Set 2: Lung cancer dataset

This section analyzes the data set from a clinical investigation as discussed in [14]. This dataset can be found in the R package survival. The study followed up on 137 lung cancer patients who are Veterans Administration. The censorship rate in this study is approximately 6.5%, indicating that nine out of 137 observations were censored. The response variable in this clinical trial is the time until death (measured in days), while the exploratory factors include the number of months from diagnosis to study enrollment, age (in years), and treatment type (Treat). The TTT plot shows that the HRFin Figure 8 has a decreasing shape. We use the EW distribution, which can accommodate different HR shapes. Figure 8 displays the histogram and TTT plots.



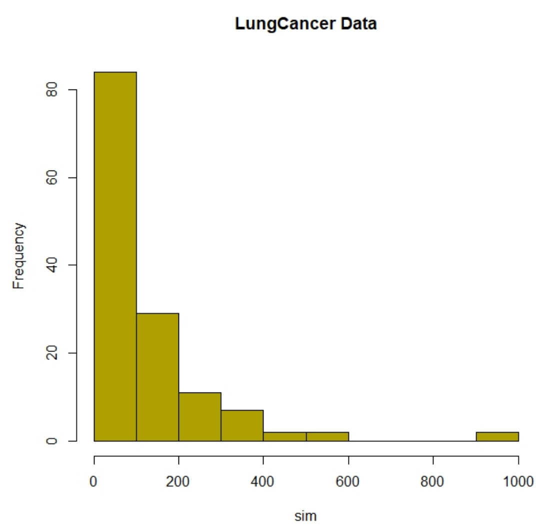


Figure 8: The TTT and histogram plots for survival times of lung cancer dataset.

Table 11 presents the results of the EWPO regression model alongside other regression models for the lung cancer dataset. The table displays the model parameters, mean, SE, z-values, p-values, and information criterion values. Figure 9 and Table11 indicate that the EWPO model is the most effective in fitting the lung cancer dataset, outperforming all other models. The lowest values for information criteria show that the proposed model provides the best fit. Additionally, all parameters of the EWPO model are significant at the 5% significance level. Figure 9 shows that the proposed model gives better fits to the lung cancer dataset over time.

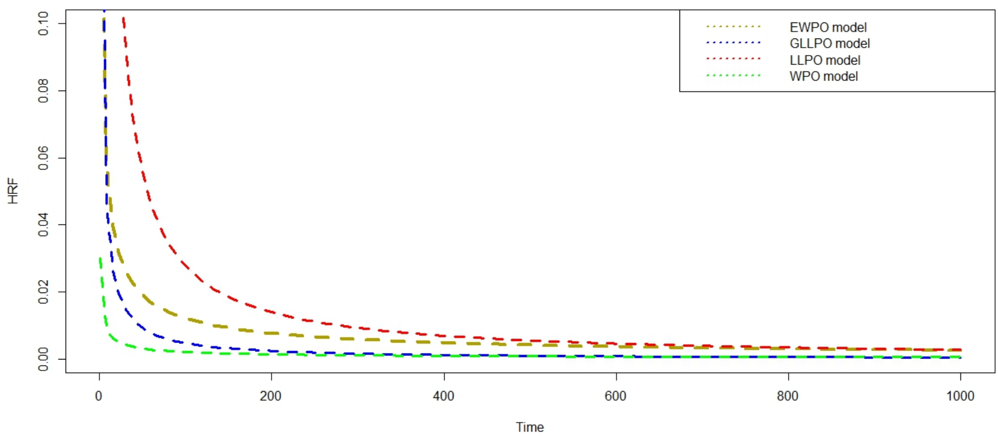


Figure 9: The estimated HRF shapes of the competing models for lung cancer dataset.

Table 11: Findings of each model for lung cancer data, along with analytical measures for various models.

Model	Parameter	Mean	SE	z-value	L 95 %	U 95 %	p-value	AIC	BIC
EWPO	β_1	0.009	0.013	0.761	-0.016	0.035	0.040	1524.993	1542.513
	β_2	-1.066	0.299	3.561	-1.653	-0.479	0.000		
	β_3	-0.050	0.008	6.377	-0.066	-0.035	0.000		
	λ	1.432	0.957	1.497	-0.444	3.308	0.011		
	ρ	0.364	0.035	10.407	0.296	0.433	0.000		
	υ	1.313	0.831	1.581	-0.315	2.942	0.013		
GLLPO	β_1	-0.064	0.007	9.851	-0.077	-0.052	0.000	1542.570	1560.09
	β_2	0.456	0.294	1.551	-0.120	1.034	0.012		
	β_3	-0.011	0.014	0.804	-0.040	0.017	0.042		
	λ	0.614	0.056	10.988	0.505	0.724	0.000		
	ρ	0.747	0.046	16.395	0.659	0.837	0.000		
	υ	0.356	0.121	2.956	0.120	0.593	0.003		
WPO	β_1	0.069	0.012	5.998	0.047	0.093	0.000	1617.568	1632.168
	β_2	0.823	0.302	2.723	0.231	1.416	0.019		
	β_3	-0.034	0.007	4.669	-0.049	-0.020	0.018		
	λ	666.642	147.153	4.530	378.229	955.057	0.000		
	ρ	1.121	0.093	12.114	0.940	1.303	0.000		
LLPO	β_1	-0.064	0.010	6.520	-0.084	-0.045	0.000	2171.097	2185.697
	β_2	2.667	0.339	7.866	2.003	3.332	0.070		
	β_3	-0.043	0.007	6.219	-0.057	-0.030	0.030		
	λ	1.172	0.729	1.609	-0.256	2.600	0.000		
	ρ	0.076	0.006	12.688	0.065	0.088	0.000		

8 Concluding Remarks

This paper introduces a flexible, fully parametric model for proportional odds regression that integrates the fundamental shapes of the failure rate through the EW distribution. The proposed model is referred to as the exponentiated-Weibull proportional odds (EWPO) regression model. A Monte Carlo simulation study is conducted to assess the model's performance, and it is applied to two censored survival datasets. The results indicate that this model outperforms existing proportional odds models, such as the GLL, Weibull, and log-logistic models, in accurately representing both monotonic and nonmonotonic HR functions. The EWPO model also demonstrates good performance, as indicated by the SE, AB, SE, MSE, and RMSE values. The study then applied the model to two real-world right-censored survival datasets, namely the IPASS dataset and data from lung cancer patients. The results showed that the EWPO model performes better than other competing PO models, indicating significant distributional parameters and regression coefficients. However, the EWPO model has some limitations. It is not suitable for modeling survival data with crossing survival curves. Additionally, the complexity of the model may lead to overfitting, particularly when the number of parameters is too high relative to the sample size.

In the future, this approach could be expanded to address other event scenarios, including multi-state and competing risk models. The model could also be adapted for use within Bayesian frameworks and excess hazard models. Furthermore, future research could explore various censoring strategies, such as left censoring, interval censoring, middle censoring, and double censoring.

Acknowledgement The authors would like to thank the Editorial Board and the three reviewers for their constructive comments, which greatly improved the final version of the paper.

Conflicts of Interest The authors declare no conflict of interest.

References

[1] G. Y. I. Abdallh & Z. Y. Algamal (2023). An investigational modeling approach for improving gene selection using regularized Cox regression model. *Mathematical Biology and Bioinformatics*, *18*(2), 282–293. https://doi.org/10.17537/2023.18.282.

[2] S. Afroja, M. R. Kabir & M. A. Islam (2020). Analysis of determinants of severity levels of childhood anemia in Bangladesh using a proportional odds model. *Clinical Epidemiology and Global Health*, *8*(1), 175–180. https://doi.org/10.1016/j.cegh.2019.06.009.

[3] N. M. Alharbi, J. Arasan, N. H. Abd.Rahman & W. S. Y. Ling (2021). Inference for the generalized exponential distributions with covariate and right-censored data. *Malaysian Journal of Mathematical Sciences*, *15*(3), 411–423.

[4] B. Aljawadi (2020). Nonparametric estimation of a survival function with interval censored data. *Malaysian Journal of Mathematical Sciences*, *14*(2), 285–293.

[5] S. Bashir, B. Masood, L. A. Al-Essa, A. Sanaullah & I. Saleem (2024). Properties, quantile regression, and application of bounded exponentiated Weibull distribution to COVID-19 data of mortality and survival rates. *Scientific Reports*, *14*(1), Article ID: 14353. https://doi.org/10.1038/s41598-024-65057-6.

[6] S. Bennett (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine*, *2*(2), 273–277. https://doi.org/10.1002/sim.4780020223.

[7] K. Boonlha (2024). Weighted likelihood estimator for exponentiated Weibull distribution. *Interdisciplinary Research Review*, *19*(5), 1–10. https://ph02.tci-thaijo.org/index.php/jtir/article/view/247829.

[8] D. Collett (2023). *Modelling Survival Data in Medical Research*. Chapman and Hall/CRC, New York. https://doi.org/10.1201/9781003282525.

[9] A. A. Ewees, M. A. Al-Qaness, L. Abualigah, D. Oliva, Z. Y. Algamal, A. M. Anter, R. Ali Ibrahim, R. M. Ghoniem & M. Abd Elaziz (2021). Boosting arithmetic optimization algorithm with genetic algorithm operators for feature selection: Case study on Cox proportional hazards model. *Mathematics*, *9*(18), Article ID: 2321. https://doi.org/10.3390/math9182321.

[10] J. J. Hsieh & Y. J. Chen (2022). Proportional hazard model and proportional odds model under dependent truncated data. *Axioms*, *11*(10), Article ID: 521. https://doi.org/10.3390/axioms11100521.

[11] X. Huang, C. Xiong, T. Jiang, J. Lu & J. Xu (2022). Efficient estimation and inference in the proportional odds model for survival data. *Mathematics*, *10*(18), Article ID: 3362. https://doi.org/10.3390/math10183362.

[12] M. A. S. Ishag, A. Wanjoya, A. Adem & A. Z. Afify (2023). Exponentiated Weibull mixture cure model to handle right-censored data set. In *Mathematical Methods for Engineering Applications*, volume 439 of *Springer Proceedings in Mathematics & Statistics* pp. 241–251. Springer Nature Switzerland, Cham. https://doi.org/10.1007/978-3-031-49218-1_17.

[13] M. A. S. Ishag, A. Wanjoya, A. Adem, R. Alsultan, A. S. Alghamdi & A. Z. Afify (2024). The exponentiated-Weibull proportional hazard regression model with application to censored survival data. *Alexandria Engineering Journal*, *104*, 587–602. https://doi.org/10.1016/j.aej.2024.08.007.

[14] J. D. Kalbfleisch & R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, United States. https://doi.org/10.1002/9781118032985.

[15] S. A. Khan (2018). Exponentiated Weibull regression for time-to-event data. *Lifetime Data Analysis*, *24*(2), 328–354. https://doi.org/10.1007/s10985-017-9394-3.

[16] K. K. Mahanta & J. Hazarika (2023). A multivariate proportional odds frailty model with Weibull hazard under Bayesian mechanism. *Indian Journal of Science and Technology*, *16*(Special Issue 2), 30–37. https://doi.org/10.17485/IJST/v16iSP2.6447.

[17] T. S. Mok, Y. L. Wu, S. Thongprasert, C. H. Yang, D. T. Chu, N. Saijo, P. Sunpaweravong, B. Han, B. Margono, Y. Ichinose, Y. Nishiwaki, Y. Ohe, J. J. Yang, B. Chewaskulyong, H. Jiang, E. L. Duffield, C. L. Watkins, A. A. Armour & M. Fukuoka (2009). Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *New England Journal of Medicine*, *361*(10), 947–957. https://doi.org/10.1056/nejmoa0810699.

[18] G. S. Mudholkar & D. K. Srivastava (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, *42*(2), 299–302. https://doi.org/10.1109/24.229504.

[19] G. S. Mudholkar, D. K. Srivastava & M. Freimer (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, *37*(4), 436–445. https://doi.org/10.1080/00401706.1995.10484376.

[20] A. H. Muse, S. Mwalili, O. Ngesa, C. Chesneau, A. Al-Bossly & M. El-Morshedy (2022). Bayesian and frequentist approaches for a tractable parametric general class of hazard-based regression models: An application to oncology data. *Mathematics*, *10*(20), Article ID: 3813. https://doi.org/10.3390/math10203813.

[21] A. H. Muse, S. Mwalili, O. Ngesa, C. Chesneau, H. M. Alshanbari & A.-A. H. El-Bagoury (2022). Amoud class for hazard-based and odds-based regression models: Application to oncology studies. *Axioms*, *11*(11), Article ID: 606. https://doi.org/10.3390/axioms11110606.

[22] A. H. Muse, O. Ngesa, S. Mwalili, H. M. Alshanbari & A.-A. H. El-Bagoury (2022). A flexible Bayesian parametric proportional hazard model: Simulation and applications to right-censored healthcare data. *Journal of Healthcare Engineering*, *2022*(1), Article ID: 2051642. https://doi.org/10.1155/2022/2051642.

[23] S. Nadarajah, G. M. Cordeiro & E. M. M. Ortega (2013). The exponentiated Weibull distribution: A survey. *Statistical Papers*, *54*, 839–877. https://doi.org/10.1007/s00362-012-0466-x.

[24] P. Royston & M. K. B. Parmar (2002). Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, *21*(15), 2175–2197. https://doi.org/10.1002/sim.1203.

[25] N. Taketomi, K. Yamamoto, C. Chesneau & T. Emura (2022). Parametric distributions for survival and reliability analyses, a review and historical sketch. *Mathematics*, *10*(20), Article ID: 3907. https://doi.org/10.3390/math10203907.

[26] M. G. F. Vieira, M. R. P. Cardial, R. Matsushita & E. Y. Nakano (2023). Proportional odds hazard model for discrete time-to-event data. *Axioms*, *12*(12), Article ID: 1102. https://doi.org/10.3390/axioms12121102.

[27] L. Wang & L. Wang (2022). An EM algorithm for analyzing right-censored survival data under the semiparametric proportional odds model. *Communications in Statistics-Theory and Methods*, *51*(15), 5284–5297. https://doi.org/10.1080/03610926.2020.1837879.

[28] P. W. Withana Gamage, C. S. McMahan & L. Wang (2023). A flexible parametric approach for analyzing arbitrarily censored data that are potentially subject to left truncation under the proportional hazards model. *Lifetime Data Analysis*, *29*(1), 188–212. https://doi.org/10.1007/s10985-022-09579-z.

[29] S. Yang & R. L. Prentice (1999). Semiparametric inference in the proportional odds regression model. *Journal of the American Statistical Association*, *94*(445), 125–136. https://doi.org/10.1080/01621459.1999.10473829.

[30] B. Zhao, S. Wang & C. Wang (2024). Variable selection in proportional odds model with informatively interval-censored data. *Statistical Papers*, *65*(4), 2461–2488. https://doi.org/10.1007/s00362-023-01486-1.

[31] L. Zhu, X. Tong, D. Cai, Y. Li, R. Sun, D. K. Srivastava & M. M. Hudson (2021). Maximum likelihood estimation for the proportional odds model with mixed interval-censored failure time data. *Journal of Applied Statistics*, *48*(8), 1496–1512. https://doi.org/10.1080/02664763.2020.1789077.