Alexandria Engineering Journal 104 (2024) 587–602

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original article

The exponentiated-Weibull proportional hazard regression model with application to censored survival data

Mohamed A.S. Ishag [a,*], Anthony Wanjoya [b], Aggrey Adem [c], Rehab Alsultan [d], Abdulaziz S. Alghamdi [e], Ahmed Z. Afify [f]

[a] Department of Mathematics, Pan African University, Institute for Basic Sciences, Technology and Innovation, Nairobi, 62000-00200, Kenya
[b] Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, 62000-00200, Kenya
[c] Department of Mathematics and Physics, Technical University of Mombasa, Mombasa, Kenya
[d] Department of Mathematical Science, Faculty of Applied Science, Umm AL-Qura University, Makkah, 24382, Saudi Arabia
[e] Department of Mathematics, College of Science & Arts, King Abdulaziz University, P.O. Box 344, Rabigh, 21911, Saudi Arabia
[f] Department of Statistics, Mathematics, and Insurance, Benha University, Benha, 13511, Egypt

ARTICLE INFO

Keywords:
Survival analysis
Censored data
Proportional hazard regression model
Exponentiated-Weibull distribution
Maximum likelihood estimation
Monte Carlo simulation

ABSTRACT

The proportional hazard regression models are widely used statistical tools for analyzing survival data and estimating the effects of covariates on survival times. It is assumed that the effects of the covariates are constant across the time. In this paper, we propose a novel extension of the proportional hazard model by incorporating an exponentiated-Weibull distribution to model the baseline line hazard function. The proposed model offers more flexibility in capturing various shapes of failure rates and accommodates both monotonic and non-monotonic hazard shapes. The performance evaluation of the proposed model and comparison with other commonly used survival models including the generalized log–logistic, Weibull, Gompertz, and exponentiated exponential PH regression models are explored using simulation results. The results demonstrate the ability of the introduced model to capture the baseline hazard shapes and to estimate the effect of covariates on the hazard function accurately. Furthermore, two real survival medical data sets are analyzed to illustrate the practical importance of the proposed model to provide accurate predictions of survival outcomes for individual patients. Finally, the survival data analysis reveal that the model is a powerful tool for analyzing complex survival data.

1. Introduction

Survival analysis is a statistical approach for analyzing time-to-event data, such as the time until death, recurrence of disease, or progression of disease. Survival analysis is commonly used to understand the dynamic nature of the disease and its treatment, to identify prognostic factors, and to develop personalized treatment plans [1]. The proportional hazards (PH) model is a popular and widely used model for analyzing survival data [2]. However, the PH model assumes that the hazard rate function (HRF) is proportional across different levels of the covariates and that the baseline HRF follows a specific distribution, such as the Weibull or Gompertz distribution. These assumptions may not accurately capture the shape of the baseline HRF, which can lead to skewed regression estimations coefficients and inaccurate predictions of survival outcomes [3,4] However, various types of models, such as accelerated failure time (AFT) [5,6], the accelerated hazard (AH) [7] and the proportional odds (PO) models, have been presented [8]. The probability distributions family provides the foundation for survival models that might be parametric, semi-parametric, or non-parametric [9,10]. If a distributional assumption is correctly stated, then parametric survival models provide more efficient estimates which have lower standard errors as compared to semi-parametric and non-parametric models [11]. The first step in modeling survival data with a parametric approach is to select an appropriate baseline distribution that is capable of capturing major features of interesting observations. More information can be explored in [12–14]. Only a few statistical distributions are closed under the PH framework, and none of them are flexible enough to accommodate such type of survival data [15]. The vast majority of the distributions that closed on the assumption of PH failed in modeling unimodal and bathtub of the survival data [2,16,17]. The exponentiated-Weibull (EW) distribution is a flexible distribution that can be used to estimate the shape of the baseline hazard function more accurately [18]. The EW distribution is flexible in modeling

* Corresponding author.
E-mail address: ishag.mohamed@students.jkuat.ac.ke (M.A.S. Ishag).

https://doi.org/10.1016/j.aej.2024.08.007
Received 19 May 2024; Received in revised form 18 July 2024; Accepted 3 August 2024
1110-0168/© 2024 The Authors. Published by Elsevier B.V. on behalf of Faculty of Engineering, Alexandria University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

both monotone and non-monotone HRFs while estimating only three parameters [19].

In this paper, the PH model with EW baseline distribution is proposed to handle right-censored survival data sets and evaluate its performance. The proposed model is called the exponentiated-Weibull proportional hazard (EW-PH) regression model. The model's ability is demonstrated to capture the different shapes of the baseline HRF and to estimate the effect of covariates on the hazard function accurately. In addition, the paper contributes to the existing literature on survival analysis in cancer research by demonstrating the model's ability to handle right-censored data and to provide accurate predictions of survival outcomes for individual patients. Furthermore, the paper provides a comprehensive evaluation of the model's performance using appropriate information criteria (IC) such as the Akaike IC (AIC) and Bayesian IC (BIC), and compares its performance to other commonly regression models. The results of the analysis demonstrate that the PH model with EW baseline distribution produces better fit to the data and improves the accuracy of survival predictions compared to other commonly used models. Overall, the motivation of this paper lies in the application of a flexible and accurate model to complex and right-censored data sets, as well as in the comprehensive evaluation and comparison of the model's performance to other commonly used models. The application of this model to cancer research has the potential to improve the understanding of the dynamic nature of the disease and its treatment and to develop more accurate and efficient models for predicting survival outcomes.

The paper is outlined in the following sections: The model formulation and assumptions are discussed in Section 2. The synthesis of the EW distribution is given in Section 3. The proposed EW-PH regression model is presented in Section 4. The parameter estimation of the EW-PH model is described in Section 5. Section 6 provides a simulation study to evaluate the model's performance. In Section 7, the practical importance of the EW-PH regression model is explored by analyzing two right-censored data sets. Finally, Section 8 provides the Conclusions and future work.

2. Synthesis of the of PH model

The parametric PH models are developed by utilizing a defined baseline HRF and a link function $\varphi(\beta\, x')$ for the variables as follows:

1 - $\varphi(\beta\, x') > 0, \quad \forall_x \neq 0$.

2 - $\varphi(\beta\, x')$, is a monotone function.

3 - $\varphi(0) = 1$. The link function $\varphi(\beta x')$ is an exponential $\exp(\beta x')$ or log-linear function.

The principle of the PH assumes that the impact of the covariates of interest is proportional to increases or decreases in the HRF, that is, independent of time. Here, the PH model assumes that $\varphi(\beta x') = \exp(\beta\, x')$.

The survival function (SF) of the PH model follows as

$$S(t|x) = S_0(t)\varphi(\beta\, x') = S_0(t)\, \exp(\beta\, x'), \tag{1}$$

where $S(t|x)$ is the SF at the time t, $S_0(t)$ is the baseline SF and the $\varphi(x)$ is the link function. The HRF of the PH model follows as

$$h(t|x) = h_0(t)\, \varphi(\beta\, x') = h_0(t)\, \exp(\beta\, x') = h_0(t) e^{(\beta\, x')}, \tag{2}$$

where $h_0(t)$ is the baseline HRF. Hence, the HRF can be expressed as

$$h(t|x) = h_0(t)\, \exp(\beta_1\, x_1 + \beta_2\, x_2 + \beta_3\, x_3 + \cdots + \beta_p\, x_p). \tag{3}$$

Applying Eq. (2), it can be seen that the HRF when compared to any two covariates formulation, such as (x_i and x_j), is

$$H\,R(x_i, x_j, \beta) = \frac{h_0(t)\, \exp\,(\beta\, x_i')}{h_0(t) \exp\,(\beta\, x_j')} = \exp\left[\beta(x_i' - x_j')^T\right]. \tag{4}$$

According to Eq. (4) the baseline hazards canceled other ones out in this ratio. As a result, the hazard rate (HR) for a single person is proportional to the HR for every other person. In contrast, a constant of the proportionality is time-independent, which is the basic assumption of this framework [11]. The PH models, unlike the majority of parametric regression approaches, including AFT frameworks, do not have the point of intercept [20].

3. The EW distribution

The baseline parametric distribution is important in illustrating the different hazard shapes which can the PH model capture [7]. The EW distribution is flexible and frequently employed in analyses of survival. The EW distribution by [21] extends the Weibull distribution. Khan (2018) [22] developed the EW regression for time-to-event data, in the context of AFT models. The EW distribution has some advantages over other parametric distributions by providing a wide range of basic shapes and accommodating various types of survival patterns [23]. Many of traditional distributions closed by the PH framework like the exponential, Gompertz, and Weibull distributions, but none of which is capable to modeling unimodal and bathtub HRF shapes. As a result, it is a good idea to look for certain distributions that can handle monotone and non-monotone HRF.

The probability density function (PDF) of the EW model reduces to

$$f(t) = \rho\, \kappa\, \upsilon(\kappa\, t)^{\rho-1}(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon-1} \exp\{-(\kappa\, t)^{\rho}\}, \tag{5}$$

where $t > 0$, $\kappa > 0$ is scale parameter, and $\rho, \upsilon > 0$ are shape parameters. The HRF, cumulative distribution function (CDF), SF, and cumulative hazard function (CHF) of the EW distribution are given, respectively, by

$$h(t) = \frac{\rho\, \kappa\, \upsilon(\kappa\, t)^{\rho-1}(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon-1} \exp\{-(\kappa\, t)^{\rho}\}}{1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}}, \tag{6}$$

$$F(t) = (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}, \tag{7}$$

$$S(t) = 1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}, \tag{8}$$

and

$$H(t) = \text{-log}\{1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}\}. \tag{9}$$

Fig. 1 illustrates the shapes that HRF which accommodates constant, increasing, decreasing, bathtub, unimodal, and J-shape.

4. The proposed EW-PH regression model

The proposed PH model can be developed by incorporating covariates $\boldsymbol{x}$ into the EW (ρ, κ, υ) distribution. The PH model is defined by

$$h(t) = \frac{\rho\, \kappa\, \upsilon(\kappa\, t)^{\rho-1}(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon-1} \exp\{-(\kappa\, t)^{\rho}\}}{1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}}, \tag{10}$$

where the HRF for a person with a covariate vector x and link function $\exp(\beta x')$ is given by

$$h(t) = h_0(t)\, \exp(\beta x').$$

The HRF of the proposed model is given by

$$h_{EW-PH}(t; x) = \frac{\rho\, \kappa\, \upsilon(\kappa\, t)^{\rho-1}(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon-1} \exp\{-(\kappa\, t)^{\rho}\}}{1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}} \times \exp(\beta x'). \tag{11}$$

The CDF, SF, PDF, and CHF the proposed model are given by

$$F_{EW-PH}(t; x) = [(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}] \exp(\beta x'), \tag{12}$$

$$S_{EW-PH}(t; x) = [1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}] \exp(\beta x'), \tag{13}$$

$$f_{EW-PH}(t; x) = [\rho\, \kappa\, \upsilon(\kappa\, t)^{\rho-1}(1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon-1} \exp\{-(\kappa\, t)^{\rho}\}] \exp(\beta x'), \tag{14}$$

and

$$H_{EW-PH}(t; x) = -\log[1 - (1 - \exp\{-(\kappa\, t)^{\rho}\})^{\upsilon}] \exp(\beta x'). \tag{15}$$

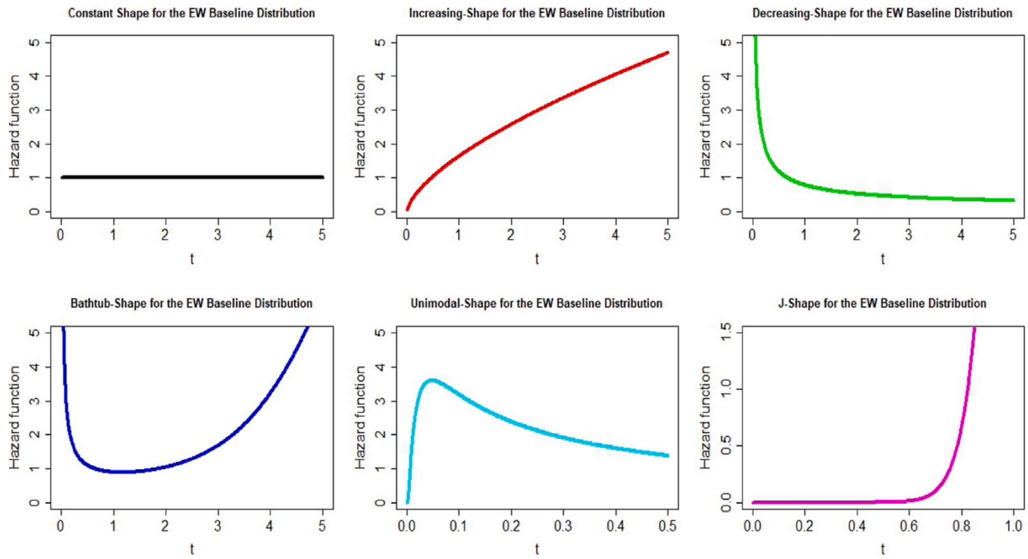

Fig. 1. Plots of the HRF of the EW distribution.

5. Estimation of the EW-PH parameters

The EW-PH model parameters are estimated via the maximum likelihood (ML) estimation technique. The ML method calculates both the regression coefficients as well as the distribution's parameters by maximizing the likelihood function (LF), which represents the probability of observing the survival times given by the covariates.

Consider the lifetimes of n individuals, say $T_1, T_2, T_3, \ldots, T_n$. If the data is subjected to the right censoring, it is found that $t_i = min(T_i, C_i)$, where $C_i > 0$ corresponds to a possible censorship time for individual i. Allowing $\delta_i = I(T_i, C_i)$ to equal 1 when $T_i \le C_i$ and 0 otherwise.

Suppose that a censored random sample of the observed data for individual i consists of $D = (t_i, \delta_i, x_i)$, $i = 1, 2, 3, \ldots, n$, where t_i is a censoring time according to whether $\delta_i = 1$ or 0, respectively, where $x_i = x_1, x_2, \ldots, x_n$ is a vector of external covariates for the ith individual. Then, the censored LF for a parametric PH model follows as

$$\begin{aligned} L(\varphi|t) &= \prod_{i=1}^{n} [f(t_i; \varphi, \beta, x)]^{\delta_i} [S(t_i; \varphi, \beta, x)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} [h(t_i|\varphi, \beta, x)\, S(t_i|\varphi, \beta, x)]^{\delta_i} [S(t_i|\varphi, \beta, x)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} [h(t_i|\varphi, \beta, x)]^{\delta_i} [S(t_i|\varphi, \beta, x)] = \prod_{i=1}^{n} [h(t_i|\varphi, \beta, x)]^{\delta_i} \\ &\quad \times \exp\left[-\int_0^t h(u)\, du\right] \\ &= \prod_{i=1}^{n} [h(t_i|\varphi) \exp(\beta x')]^{\delta_i} \exp[-H(t_i|\varphi) \exp(\beta x')], \end{aligned} \quad (16)$$

where $\varphi = (\rho, \kappa, \upsilon$ and $\beta)$ are the vector of the baseline distributional parameters and vector of the regression coefficients respectively. According to Eq. (16), the log-LF (LLF) for a parametric PH model reduces as follows:

$$\ell(\varphi|t) = \sum_{i=1}^{n} \delta_i \log[h_0(t_i|\varphi) \exp(\beta x_i^{\prime})] - \sum_{i=1}^{n} [H_0(t_i|\varphi) \exp(\beta x_i^{\prime})]. \quad (17)$$

An iterative Newton–Raphson optimization approach, can be directly used to obtain the ML estimators $\hat{\varphi}$ of φ and interval estimates of hypotheses tests of model parameters are available due to approximately normally distributed ML estimates (MLEs) [24]. By using hazard rate function in Eq. (6) and cumulative HRF as given in Eq. (9). The LLF of the EW-PH model can be rewritten as follows:

$$\ell(\varphi|t) = \sum_{i=1}^{n} \delta_i \log\left[\frac{(\rho\, \kappa\, \upsilon(\kappa\, t_i)^{\rho-1}(a_i)^{\upsilon-1} b_i)^{\rho}}{1-(a_i)^{\upsilon}} p_i\right] - \sum_{i=1}^{n}\left[(1-(a_i)^{\upsilon}) p_i\right], \quad (18)$$

where $a_i = 1 - \exp\{-(\kappa\, t_i)^{\rho}\}$, $b_i = \exp\{-(\kappa\, t_i)^{\rho}\}$, and $p_i = \exp(\beta\, x_i^{\prime})$. In order to obtain the ML estimators of $\hat{\varphi} = (\hat{\rho}, \hat{\kappa}, \hat{\upsilon}$ and $\hat{\beta})$, Eq. (18) can be maximized directly with respect to $(\rho, \kappa, \upsilon$ and $\beta)$. The 1st derivatives of the LLF are:

$$\frac{\partial \ell(\varphi)}{\partial \rho} = \log(\kappa)\left(\sum_{i=1}^{n} \delta_i\right) + \left(\sum_{i=1}^{n} \delta_i \log(t_i)\right) + \frac{1}{\rho}\sum_{i=1}^{n} \delta_i, \quad (19)$$

$$\frac{\partial \ell(\varphi)}{\partial \kappa} = \sum_{i=1}^{n} \delta_i \left(\frac{\rho}{\kappa}\right) \quad (20)$$

$$\begin{aligned} \frac{\partial \ell(\varphi)}{\partial \upsilon} &= \sum_{i=1}^{n} \delta_i \left(\frac{-2\log(a_i)\, p_i \upsilon + \delta_i a_i^{2\upsilon}}{\upsilon\left(-1 + a_i^{\upsilon}\right)^2}\right) + \sum_{i=1}^{n} \delta_i \left(\frac{\log(a_i)\, a_i^{3\upsilon} p_i \upsilon}{\upsilon\left(-1 + a_i^{\upsilon}\right)^2}\right) \\ &+ \sum_{i=1}^{n} \delta_i \left(\frac{\left(\upsilon\left(p_i - \delta_i\right)\log(a_i) - 2\delta_i\right) a_i^{\upsilon}}{\upsilon\left(-1 + a_i^{\upsilon}\right)^2}\right) + \sum_{i=1}^{n} \delta_i \left(\frac{\delta_i\left(\upsilon \log(a_i) + 1\right)}{\upsilon\left(-1 + a_i^{\upsilon}\right)^2}\right) \end{aligned} \quad (21)$$

and

$$\frac{\partial \ell(\varphi)}{\partial \beta} = \left(\sum_{i=1}^{n} \frac{\delta_i}{p_i}\right) - n - \left(\sum_{i=1}^{n} \left(-a_i^{\upsilon}\right)\right). \quad (22)$$

To maximizing the LLF, many software tools are available. In this work the R software is used.

6. Simulation analysis

In this section, a comprehensive simulation study demonstrates the proposed parametric EW-PH model's good frequentist features is explored. This study also provides the AIC and BIC values to select models that accurately describe the underlying hazard structure. The values of the parameters are selected to simulate scenarios resembling cancer population studies along with an aggressive kind of cancer, (smaller 5-year survival rate) such as lung cancer [25].

6.1. Simulation algorithm for EW-PH model

The steps for executing the proposed EW-PH model using the R software are as follow:

1- Set the initial values of the model's parameters.

2- Create the lifetime data by applying the inverse transform technique to the CHRF of the proposed model.
3- Utilize the various estimates to evaluate the true values of the parameters.
4- Analyze the inferential properties of the estimates, considering the standard error (SE), average bias (AB), mean squared error (MSE), and root mean square error (RMSE).

6.2. Generating data from EW-PH model

The inversion technique was employed to generate survival data for the EW-PH model. This technique depends on the relationship between the CHRF of survival and standard uniform random variables. When there is a closed-form formulation for the CHRF, then it can be employed inverted and simply implemented through R [26]. The censorship rate is obtained by applying administration censorship at $T_c = 5$ years, resulting in approximately 20% censoring in all sets [25]. The EW distribution (ρ, κ, υ) is employed in the simulation of survival data from the PH model using the inverse transform approach. The CHRF and the inverse of the EW model are given by Eqs. (23) and (24), respectively.

$$H_0(t; \rho, \kappa, \upsilon) = -\log[1 - (1 - \exp\{-(\kappa t)^{\rho}\})^{\upsilon}], \tag{23}$$

and

$$H_0^{-1}(t; \rho, \kappa, \upsilon) = \frac{\left(-\log\left(1 - (1 - \exp(-\kappa t))^{1/\upsilon}\right)\right)^{1/\rho}}{\kappa}. \tag{24}$$

6.3. Simulation design and analysis

The simulation study is conducted by running a series of simulations based on the PH model, with various sample sizes of $n = 1000, 5000$, and 10000 data sets. The values of the covariates are generated using a variety of uniform distributions with probability 0.25 on (30, 65), probability 0.35 on (65, 75), and probability 0.45 on (75, 85) years old. The variable “age” is simulated from continuous distribution and the binary variables “sex” and “comorbidity” are both simulated from a binomial distribution with probability 0.5, see [7]. To evaluate the properties of the studied models in all simulated scenarios, the EW-PH model is fitted and compared to the corresponding actual generating model. In addition, the AIC and BIC are obtained to select the true generating model. The actual parameter values of the generating model are used as initial points in all cases because the goal is to study the estimator’s properties instead of the optimization process. The analysis is carried out using the R program. The optimization process is carried out using 'nlminb()' of the R commands.

6.4. Measuring performance

The model flexibility for the covariates is investigated using some measures including the mean of the MLEs for the corresponding models, the SE AB, MSE, and RMSE. Furthermore, the AIC and BIC are used to compare the different fitted models. The AB, MSE, and RMSE can be computed as follows:

$$\text{AB} = \frac{\sum_{i=1}^{n}(\hat{\vartheta}_i - \vartheta)}{n}, \tag{25}$$

$$\text{MSE} = \frac{\sum_{i=1}^{n}(\hat{\vartheta}_i - \vartheta)^2}{n}, \tag{26}$$

and

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{\vartheta}_i - \vartheta)^2}{n}}. \tag{27}$$

where ϑ represents each of the parameters being considered.

6.5. Simulation results

The simulation results provide insights into the performance of the proposed model under different conditions. Tables 1–6 demonstrate the AB, SE, MSE, and RMSE of the proposed model in accurately estimating the HRF and survival probabilities under various scenarios. Moreover, it seems that both sample sizes, n = 40, 100, 200, 1000, 5000, and 10 000, and the percentage of censorship, which is about 20%, have an effect on how accurately models fit data. The model shows minimal bias in parameter estimation, with MSE and RMSE values indicating high precision for the MLEs. The averages estimates are extremely close, and the AB and MSE decrease as the sample size increases. Moreover, the different hazard rate shapes are presented in Figs. 2–10, which illustrate that the parameters estimates perform better, as the sample sizes are increased.

The results in Tables 7–12 provide comparisons for the model performance based on the AIC and BIC values with several sample sizes $n = 40$, 100, 200, 1000, 5000, 10 000, and censoring percentages 20%. The EW-PH regression model has the smallest values in terms of AIC and BIC in all scenarios, which is an indication that the proposed model performs better as compared to the generalized log–logistic PH (GLL-PH) [15], Weibull PH (W-PH) [27], Gompertz PH (G-PH) [28], and exponentiated exponential PH (EE-PH) [29] regression models. It is also noted that, as the sample size increases, under the preset of censoring, the suggested model once again provides the best fit. Finally, the proposed framework performs better than all the competitive models under all conditions, and it can successfully capture the underlying survival patterns.

7. Applications to right-censored data

In this Section, the EW-PH regression model is used to analyze two real-life censored medical data sets including the leukemia survival and larynx cancer data sets.

The first set of data is the leukemia survival data, which is available under the R package named LeukSurv [30]. The data set includes information about survival times of 1043 patients of acute myeloid leukemia. This data set was analyzed by [31]. It would be interesting to look into possible regional variations in survival after controlling for known subject-specific prognostic variables. Also, the information included the survival duration measured in years, essential state at the end of the follow-up, if 0 is right-censored, 1 is dead, age in years, sex if 0 is female, 1 is male, the townsend score (TPI), white blood cell count at diagnosis (WBC), and all patients’ specific residence addresses as well as their administrative districts (district), the entire region’s 24 districts are available in [32].

The second set of data represents the larynx cancer data, which are categorized by the stage of the cancer tumor, and it contains the lifetimes of 90 patients. The study records the time until death in months. Alvares et al. [33] and Muse et al. [16] have examined the data from different perspectives using various hazard-based regression models. More information about the original data can be found in [34]. In addition to the survival times, other factors in the data include the age of the patients at the time of diagnosis and the year of diagnosis. The study aimed to investigate the potential links between age, year of diagnosis, cancer stage, and the mortality of patients with laryngeal cancer.

The EW-PH regression model is compared to the GLL-PH, W-PH, G-PH, and EE-PH regression models. To determine which statistical distributions provides better fit to the real medical survival data, several selection criteria measurements are used, like the AIC, the consistent AIC (CAIC) [35], BIC, and Hannan-Quinn IC (HQIC) [36].

The descriptive summary of the leukemia and larynx cancer data is shown in Table 13.

Table 1
Simulation results of the EW-PH regression model for $n = 40$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
EW-PH (Increasing HRF)	β_1	0.25	0.657	0.406	0.348	0.369	0.607
	β_2	0.35	0.404	0.053	0.145	0.041	0.202
	β_3	0.45	0.677	0.227	0.333	0.256	0.506
	κ	1.00	0.988	−0.012	0.532	0.024	0.155
	ρ	1.65	1.321	−0.328	0.612	0.976	0.988
	υ	1.00	1.161	0.160	0.811	0.347	0.589
EW-PH (Decreasing HRF)	β_1	0.50	0.845	0.345	0.355	0.464	0.681
	β_2	0.40	0.488	0.088	0.156	0.079	0.281
	β_3	0.30	0.516	0.216	0.340	0.177	0.421
	κ	0.50	0.261	−0.238	0.863	0.182	0.427
	ρ	0.50	0.231	−0.268	0.150	0.196	0.443
	υ	0.25	0.526	0.276	0.406	0.215	0.464
EW-PH (Bathtub HRF)	β_1	0.95	1.226	0.276	0.342	0.601	0.775
	β_2	0.85	0.838	−0.011	0.181	0.020	0.141
	β_3	0.804	0.867	0.117	0.336	0.084	0.436
	κ	3.00	4.070	1.070	3.226	7.569	2.751
	ρ	4.00	1.780	−2.219	2.217	12.830	3.582
	υ	0.005	0.100	0.050	0.125	0.008	0.089
EW-PH (Unimodel HRF)	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.11	0.110	<0.0001	<0.0001	<0.0001	<0.0001

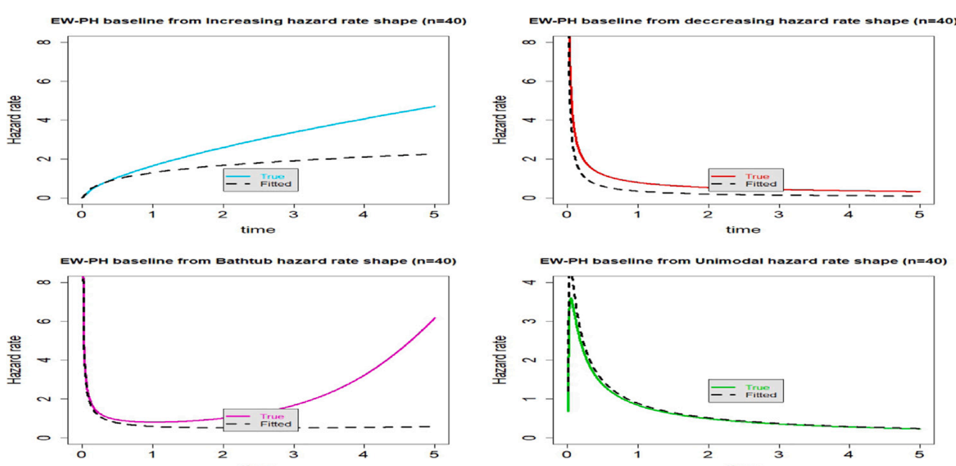

Fig. 2. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 40$ and 20% censoring.

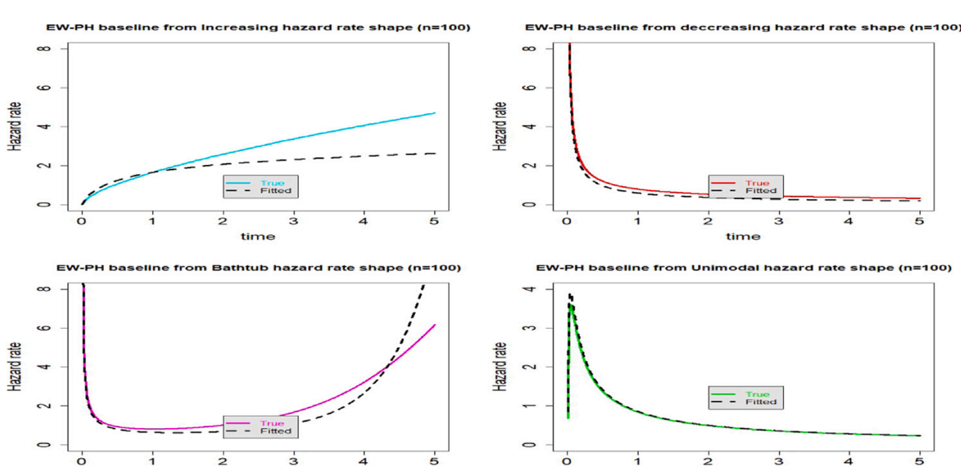

Fig. 3. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 100$ and 20% censoring.

Table 2
Simulation results of the EW-PH regression model for $n = 100$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
	β_1	0.25	0.154	−0.095	0.207	0.039	0.197
EW-PH	β_2	0.35	0.246	−0.103	0.080	0.062	0.249
(Increasing HRF)	β_3	0.45	0.580	0.130	0.211	0.134	0.366
	κ	1.00	0.771	−0.229	0.304	0.406	0.637
	ρ	1.65	1.256	−0.393	0.405	1.145	1.070
	υ	1.00	1.350	0.350	0.730	0.823	0.907
	β_1	0.50	0.382	−0.117	0.210	0.104	0.322
EW-PH	β_2	0.40	0.294	−0.105	0.082	0.073	0.270
(Decreasing HRF)	β_3	0.30	0.452	0.152	0.209	0.115	0.339
	κ	0.50	0.238	−0.261	0.371	0.193	0.439
	ρ	0.50	0.335	−0.164	0.141	0.137	0.370
	υ	0.25	0.346	0.096	0.173	0.057	0.239
	β_1	0.95	0.857	−0.093	0.202	0.168	0.410
EW-PH	β_2	0.85	0.727	−0.123	0.100	0.195	0.442
(Bathtub HRF)	β_3	0.804	0.916	0.166	0.201	0.278	0.527
	κ	3.00	3.952	0.952	0.649	6.625	2.574
	ρ	4.00	7.935	3.935	4.906	46.979	6.854
	υ	0.005	0.022	−0.027	0.014	0.002	0.045
	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
EW-PH	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
(Unimodel HRF)	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.11	0.110	<0.0001	<0.0001	<0.0001	<0.0001

Table 3
Simulation results of the EW-PH regression model for $n = 200$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
	β_1	0.25	0.253	0.003	0.146	0.002	0.045
EW-PH	β_2	0.35	0.300	−0.049	0.057	0.032	0.179
(Increasing HRF)	β_3	0.45	0.454	0.003	0.145	0.003	0.055
	κ	1.00	0.845	−0.154	0.178	0.286	0.535
	ρ	1.65	1.398	−0.251	0.256	0.767	0.876
	υ	1.00	1.271	0.271	0.382	0.616	0.785
	β_1	0.50	0.480	−0.019	0.149	0.019	0.138
EW-PH	β_2	0.40	0.322	−0.077	0.059	0.056	0.237
(Decreasing HRF)	β_3	0.30	0.296	−0.003	0.144	0.002	0.045
	κ	0.50	0.225	−0.274	0.218	0.199	0.446
	ρ	0.50	0.382	−0.118	0.108	0.104	0.322
	υ	0.25	0.334	0.083	0.111	0.049	0.221
	β_1	0.95	0.993	0.043	0.150	0.085	0.292
EW-PH	β_2	0.85	0.759	−0.091	0.076	0.146	0.382
(Bathtub HRF)	β_3	0.804	0.781	0.031	0.142	0.048	0.219
	κ	3.00	2.903	−0.096	0.755	0.570	0.755
	ρ	4.00	3.0518	−0.948	1.429	6.686	2.586
	υ	0.005	0.065	0.015	0.032	0.002	0.045
	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
EW-PH	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
(Unimodel HRF)	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.11	0.110	<0.0001	<0.0001	<0.0001	<0.0001

7.1. Assumption of proportionality

In this subsection, the standardized Schoenfeld residuals (SSR) are applied to determine the Cox model's PH assumption for each covariate data included in the model. According to Figs. 10 and 11 with 5% level of significance of the threshold, there is no enough evidence for the rejection of the assumption of the PH for both data sets.

Tables 14 and 15 provide the MLEs, standard errors (SE), z-value, *p*-value, and confidence intervals (CI) of estimated parameters, as well as the values of AIC, BIC, CAIC, and HQIC measures for the EW-PH regression model and other regression models for the two survival data sets, respectively. The EW-PH regression model provides the smallest values of selection criteria for both data sets. Hence, the proposed EW-PH regression model provides the best fit for the two analyzed data sets and performs better as compared to the GLL-PH, W-PH, G-PH, and EE-PH regression models.

Furthermore, the estimated HRF of the EW-HP model based on leukemia and larynx cancer data are presented in Figs. 12 and 13. The plots support the TTT plots in Figs. 8 and 9, showing that the estimated HRF of the EW-PH model is decreasing for leukemia data and increasing for larynx cancer data.

8. Conclusions

In this article, we propose a proportional hazard regression model based on the exponentiated-Weibull distribution called the exponentiated-Weibull proportional hazard (EW-PH) regression model. The EW-PH regression model offers a flexible and accurate approach for survival analysis and analyzing right-censored cancer data. The model accurately estimates the regression coefficients and distribution parameters, captures the shapes of the baseline hazard function more accurately than other commonly used regression models and provides

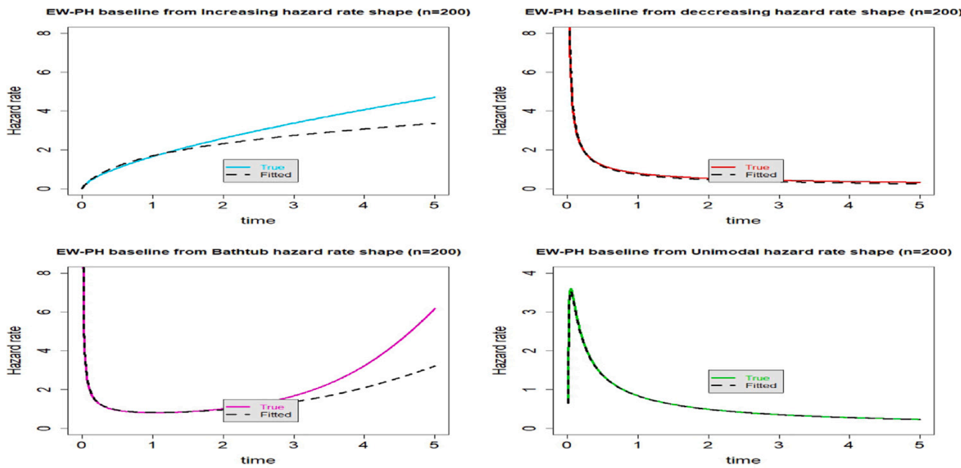

Fig. 4. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 200$ and 20% censoring.

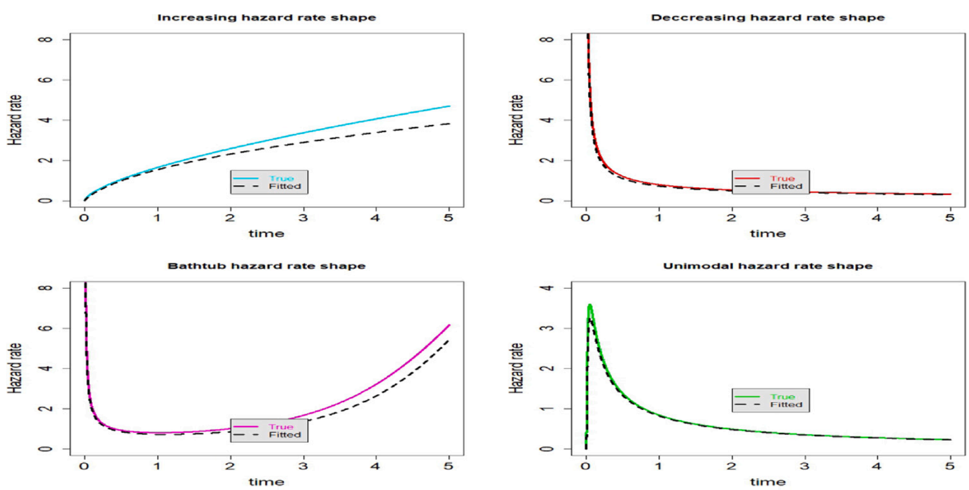

Fig. 5. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 1000$ and 20% censoring.

Table 4
Simulation results of the EW-PH regression model for $n = 1000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
EW-PH (Increasing HRF)	β_1	0.25	0.255	0.005	0.064	0.002	0.045
	β_2	0.35	0.318	−0.032	0.025	0.021	0.145
	β_3	0.45	0.502	0.052	0.065	0.049	0.221
	κ	1.00	0.979	−0.020	0.091	0.040	0.200
	ρ	1.65	1.544	−0.106	0.140	0.340	0.583
	υ	1.00	1.189	0.139	0.165	0.298	0.546
EW-PH (Decreasing HRF)	β_1	0.50	0.518	0.018	0.066	0.018	0.134
	β_2	0.40	0.377	−0.023	0.027	0.018	0.134
	β_3	0.30	0.368	0.068	0.065	0.045	0.212
	κ	0.50	0.709	0.209	0.265	0.252	0.502
	ρ	0.50	0.521	0.021	0.081	0.022	0.148
	υ	0.25	0.249	−0.001	0.043	0.000	0.000
EW-PH (Bathtub HRF)	β_1	0.95	0.975	0.025	0.063	0.048	0.219
	β_2	0.85	0.824	−0.026	0.033	0.043	0.207
	β_3	0.804	0.750	0.054	0.063	0.084	0.292
	κ	3.00	3.340	0.340	0.247	2.155	1.468
	ρ	4.00	4.475	0.475	0.726	4.028	2.007
	υ	0.005	0.047	−0.003	0.008	0.000	0.000
EW-PH (Unimodel HRF)	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.11	0.110	<0.0001	<0.0001	<0.0001	<0.0001

Table 5
Simulation results of the EW-PH regression model for $n = 5000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
EW-PH (Increasing HRF)	β_1	0.25	0.249	−0.001	0.029	0.001	0.001
	β_2	0.35	0.367	0.017	0.011	0.012	0.110
	β_3	0.45	0.416	−0.034	0.011	0.029	0.170
	κ	1.00	0.974	−0.026	0.036	0.051	0.226
	ρ	1.65	1.612	−0.038	0.061	0.124	0.352
	υ	1.00	1.040	0.040	0.061	0.082	0.286
EW-PH (Decreasing HRF)	β_1	0.50	0.495	−0.005	0.029	0.005	0.071
	β_2	0.40	0.418	0.018	0.012	0.015	0.122
	β_3	0.30	0.259	−0.040	0.029	0.023	0.152
	κ	0.50	0.458	−0.042	0.076	0.040	0.200
	ρ	0.50	0.494	−0.006	0.031	0.006	0.077
	υ	0.25	0.253	0.003	0.018	0.002	0.045
EW-PH (Bathtub HRF)	β_1	0.95	0.949	−0.001	0.028	0.002	0.044
	β_2	0.85	0.871	0.021	0.015	0.036	0.191
	β_3	0.75	0.711	−0.039	0.028	0.057	0.238
	κ	3.00	2.933	−0.067	0.112	0.397	0.632
	ρ	4.00	3.956	−0.044	0.267	0.350	0.592
	υ	0.05	0.051	0.001	0.003	0.001	0.007
EW-PH (Unimodel HRF)	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.11	0.110	<0.0001	<0.0001	<0.0001	<0.0001

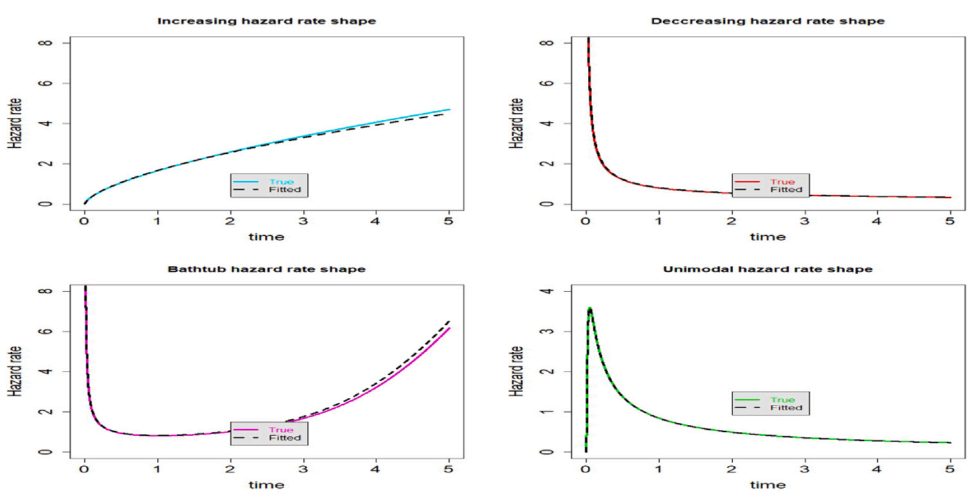

Fig. 6. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 5000$ and 20% censoring.

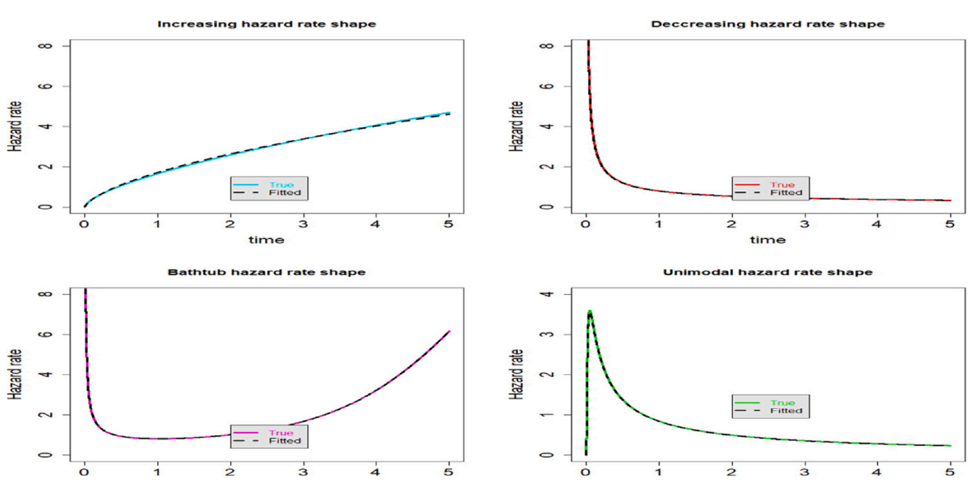

Fig. 7. The true and fitted increasing, decreasing, bathtub, and unimodel HRF shapes of the EW-PH model for $n = 10000$ and 20% censoring.

Table 6
Simulation results of the EW-PH regression model for $n = 10000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	RMSE
EW-PH (Increasing HRF)	β_1	0.25	0.229	−0.021	0.020	0.010	0.100
	β_2	0.35	0.353	0.003	0.008	0.002	0.045
	β_3	0.45	0.422	−0.028	0.020	0.024	0.155
	κ	1.00	0.973	−0.027	0.026	0.053	0.230
	ρ	1.65	1.629	−0.021	0.043	0.068	0.261
	υ	1.00	1.028	0.029	0.045	0.058	0.241
EW-PH (Decreasing HRF)	β_1	0.50	0.529	0.029	0.021	0.030	0.173
	β_2	0.40	0.397	−0.002	0.008	0.002	0.045
	β_3	0.30	0.304	0.004	0.020	0.003	0.055
	κ	0.50	0.546	0.047	0.065	0.049	0.221
	ρ	0.50	0.512	0.013	0.024	0.013	0.114
	υ	0.25	0.246	−0.004	0.013	0.002	0.045
EW-PH (Bathtub HRF)	β_1	0.95	0.950	0.000	0.020	0.000	0.000
	β_2	0.85	0.849	−0.001	0.011	0.000	0.000
	β_3	0.75	0.749	−0.001	0.020	0.002	0.004
	κ	3.00	2.999	−0.001	0.076	0.000	0.000
	ρ	4.00	4.003	0.003	0.190	0.000	0.000
	υ	0.05	0.05	0.000	0.002	0.000	0.000
EW-PH (Unimodel HRF)	β_1	0.25	0.250	<0.0001	<0.0001	<0.0001	<0.0001
	β_2	0.35	0.350	<0.0001	<0.0001	<0.0001	<0.0001
	β_3	0.45	0.450	<0.0001	<0.0001	<0.0001	<0.0001
	κ	0.000005	0.000005	<0.0001	<0.0001	<0.0001	<0.0001
	ρ	0.15	0.150	<0.0001	<0.0001	<0.0001	<0.0001
	υ	0.110	0.110	<0.0001	<0.0001	<0.0001	<0.0001

Table 7
Simulation results for the proposed EW-PH model and competing models for $n = 40$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
EW-PH	β_1	0.25	0.657	0.407	0.348	0.369	48.026	58.159
	β_2	0.35	0.404	0.054	0.145	0.041		
	β_3	0.45	0.677	0.227	0.333	0.256		
	κ	1.00	0.988	−0.012	0.532	0.024		
	ρ	1.65	1.321	−0.329	0.612	0.977		
	υ	1.00	1.161	0.161	0.811	0.348		
GLL-PH	β_1	0.25	0.120	−0.370	0.346	0.048	52.686	62.819
	β_2	0.35	0.328	−0.022	0.126	0.015		
	β_3	0.45	0.309	−0.141	0.346	0.107		
	κ	1.00	1.465	0.465	0.303	1.146		
	ρ	1.65	1.473	−0.177	0.401	0.553		
	υ	1.00	0.332	−0.668	0.724	0.890		
W-PH	β_1	0.25	0.279	−0.529	0.358	0.015	77.181	85.625
	β_2	0.35	0.060	−0.410	0.148	0.119		
	β_3	0.45	0.551	−1.001	0.351	0.101		
	κ	1.00	0.928	−0.072	0.173	0.139		
	ρ	1.65	1.766	0.116	0.223	0.396		
G-PH	β_1	0.25	0.210	−0.460	0.360	0.018	78.018	86.463
	β_2	0.35	0.058	−0.408	0.151	0.119		
	β_3	0.45	0.635	−1.085	0.365	0.201		
	κ	1.00	0.978	−0.022	0.243	0.044		
	ρ	1.65	0.613	−1.037	0.239	2.347		
EE-PH	β_1	0.25	0.482	0.232	0.354	0.170	77.587	86.031
	β_2	0.35	0.014	−0.364	0.120	0.122		
	β_3	0.45	0.304	−0.146	0.346	0.110		
	κ	1.00	1.388	0.388	0.214	0.927		
	ρ	1.65	1.774	0.124	0.228	0.425		

accurate predictions of survival outcomes for individual patients. A complete simulation study shows that the EW-PH regression model provides good results with a lower absolute bias and mean square error for the regression coefficients. The last investigation concentrated on two right-censored real-world medical data applications, which include the leukemia survival times and larynx cancer data sets. In conclusion, The proposed EW-PH regression model outperforms other competing PH regression models. Furthermore, the significance of the distribution parameters and regression coefficients is provided.

The proposed EW-PH regression model can be extended to accommodate time-varying covariates and competing risks, in light of the work of Rehman et al. [14], which are often encountered in survival analysis. Developing methodologies to handle these complex scenarios within the EW-PH framework would enhance its applicability in diverse research areas. Conducting robustness checks and sensitivity analyses to assess the performance of the EW-PH regression model under different assumptions and data conditions would provide insights into its robustness and reliability. Furthermore, conducting more case studies and applications in diverse applied areas such as healthcare, engineering, and finance would provide further evidence of the effectiveness of the EW-PH model and broaden its practical impact. Various

Table 8
Simulation results for the proposed EW-PH model and competing models for $n = 100$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
	β_1	0.25	0.154	−0.096	0.207	0.039		
EW-PH	β_2	0.35	0.246	−0.104	0.080	0.062		
	β_3	0.45	0.580	0.130	0.211	0.134	117.615	133.246
	κ	1.00	0.771	−0.229	0.304	0.406		
	ρ	1.65	1.256	−0.394	0.405	1.145		
	υ	1.00	1.350	0.350	0.730	0.823		
	β_1	0.25	0.152	−0.098	0.208	0.039		
GLL-PH	β_2	0.35	0.273	−0.077	0.078	0.048		
	β_3	0.45	0.260	−0.190	0.213	0.135	121.112	136.743
	κ	1.00	1.478	0.478	0.164	1.184		
	ρ	1.65	1.187	−0.463	0.158	1.314		
	υ	1.00	0.122	−0.878	0.368	0.985		
	β_1	0.25	0.031	−0.281	0.203	0.062		
W-PH	β_2	0.35	0.106	−0.244	0.074	0.111		
	β_3	0.45	0.101	−0.551	0.202	0.192	171.114	184.1394
	κ	1.00	1.058	0.058	0.103	0.119		
	ρ	1.65	1.722	0.072	0.137	0.243		
	β_1	0.25	0.389	−0.639	0.225	0.089		
G-PH	β_2	0.35	0.147	−0.203	0.075	0.101		
	β_3	0.45	0.162	−0.612	0.219	0.176	173.770	186.796
	κ	1.00	1.069	0.069	0.167	0.143		
	ρ	1.65	0.595	−1.055	0.128	2.368		
	β_1	0.25	0.267	0.017	0.214	0.009		
EE-PH	β_2	0.35	0.105	−0.245	0.067	0.111		
	β_3	0.45	0.382	−0.832	0.215	0.057	166.108	179.133
	κ	1.00	1.047	0.047	0.097	00.096		
	ρ	1.65	1.781	0.131	0.142	0.449		

Table 9
Simulation results for the proposed EW-PH model and competing models for $n = 200$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
	β_1	0.25	0.253	0.003	0.146	0.002		
EW-PH	β_2	0.35	0.300	−0.050	0.057	0.032		
	β_3	0.45	0.454	0.004	0.145	0.004	227.487	247.277
	κ	1.00	0.845	−0.155	0.178	0.286		
	ρ	1.65	1.398	−0.252	0.256	0.768		
	υ	1.00	1.271	0.271	0.382	0.615		
	β_1	0.25	0.416	0.166	0.147	0.111		
GLL-PH	β_2	0.35	0.295	−0.055	0.052	0.035		
	β_3	0.45	0.549	0.099	0.148	0.099	228.448	248.238
	κ	1.00	1.619	0.619	0.117	1.621		
	ρ	1.65	0.956	−0.694	0.087	1.809		
	υ	1.00	0.107	−0.893	0.224	0.989		
	β_1	0.25	0.187	−0.063	0.146	0.028		
W-PH	β_2	0.35	0.073	−0.277	0.054	0.117		
	β_3	0.45	0.110	−0.340	0.143	0.190	310.445	326.936
	κ	1.00	1.119	0.119	0.082	0.252		
	ρ	1.65	1.685	0.035	0.093	0.117		
	β_1	0.25	0.570	0.320	0.150	0.262		
G-PH	β_2	0.35	0.154	0.023	0.057	0.099		
	β_3	0.45	0.013	−0.437	0.146	0.202	323.242	339.734
	κ	1.00	1.017	0.017	0.119	0.034		
	ρ	1.65	0.473	−1.275	0.067	2.582		
	β_1	0.25	0.230	−0.020	0.143	0.010		
EE-PH	β_2	0.35	0.081	−0.269	0.049	0.116		
	β_3	0.45	0.097	−0.353	0.142	0.193	308.462	324.953
	κ	1.00	1.128	0.128	0.083	0.272		
	ρ	1.65	1.693	0.043	0.094	0.144		

Table 10
Simulation results for the proposed EW-PH model and competing models for $n = 1000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
	β_1	0.25	0.255	0.064	0.005	0.003		
EW-PH	β_2	0.35	0.318	0.025	−0.032	0.021		
	β_3	0.45	0.502	0.065	0.052	0.050	1162	1192
	κ	1.00	0.980	0.140	−0.106	0.339		
	ρ	1.65	1.544	0.091	−0.020	0.040		
	υ	1.00	1.139	0.165	0.139	0.297		
	β_1	0.25	0.208	−0.042	0.065	0.019		
GLL-PH	β_2	0.35	0.285	−0.065	0.025	0.041		
	β_3	0.45	0.369	−0.081	0.065	0.066	1163	1193
	κ	1.00	1.695	0.695	0.064	1.873		
	ρ	1.65	1.005	−0.645	0.043	1.712		
	υ	1.00	0.128	−0.872	0.141	0.984		
	β_1	0.25	0.011	0.063	−0.239	0.062		
W-PH	β_2	0.35	0.031	0.023	−0.319	0.122		
	β_3	0.45	0.070	0.063	−0.380	0.198	1418	1443
	κ	1.00	0.987	0.033	−0.013	0.026		
	ρ	1.65	1.624	0.040	−0.026	0.085		
	β_1	0.25	0.011	0.064	−0.239	0.062		
G-PH	β_2	0.35	0.032	0.023	−0.318	0.121		
	β_3	0.45	0.076	0.064	−0.374	0.197	1478	1502
	κ	1.00	0.860	0.051	−0.140	0.260		
	ρ	1.65	0.597	0.042	−1.053	2.366		
	β_1	0.25	0.250	0.000	0.000	0.000		
EE-PH	β_2	0.35	0.350	0.000	0.000	0.000		
	β_3	0.45	0.450	0.000	0.000	0.000	1629	1654
	κ	1.00	1.233	0.023	0.233	0.520		
	ρ	1.65	1.659	0.041	0.009	0.030		

Table 11
Simulation results for the proposed EW-PH model and competing models for $n = 5000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
	β_1	0.25	0.249	0.029	−0.001	0.000		
EW-PH	β_2	0.35	0.367	0.011	0.017	0.012		
	β_3	0.45	0.416	0.029	−0.034	0.029	5564	5603
	κ	1.00	0.974	0.036	−0.026	0.051		
	ρ	1.65	1.612	0.061	−0.038	0.124		
	υ	1.00	1.040	0.061	0.040	0.082		
	β_1	0.25	0.225	−0.025	0.029	0.012		
GLL-PH	β_2	0.35	0.347	−0.003	0.011	0.002		
	β_3	0.45	0.422	−0.028	0.029	0.024	5598	5637
	κ	1.00	1.645	0.645	0.024	1.706		
	ρ	1.65	1.006	−0.644	0.018	1.710		
	υ	1.00	0.052	−0.948	0.071	0.997		
	β_1	0.25	−0.001	0.028	−0.251	0.062		
W-PH	β_2	0.35	−0.017	0.010	−0.367	0.122		
	β_3	0.45	0.041	0.028	−0.409	0.201	7225	7258
	κ	1.00	1.006	0.015	0.006	0.012		
	ρ	1.65	1.646	0.018	−0.004	0.013		
	β_1	0.25	−0.007	0.028	−0.257	0.062		
G-PH	β_2	0.35	−0.018	0.010	−0.368	0.122		
	β_3	0.45	0.054	0.028	−0.396	0.200	7563	7595
	κ	1.00	0.855	0.022	−0.145	0.269		
	ρ	1.65	0.583	0.018	−1.067	2.383		
	β_1	0.25	0.250	0.000	0.000	0.000		
EE-PH	β_2	0.35	0.350	0.000	0.000	0.000		
	β_3	0.45	0.450	0.000	0.000	0.000	8420	8453
	κ	1.00	2.102	0.039	1.102	3.418		
	ρ	1.65	1.070	0.018	−0.580	1.578		

Table 12
Simulation results for the proposed EW-PH model and competing models for $n = 10000$ and approximately 20% censored observations.

Model	Parameter	True value	MLE	AB	SE	MSE	AIC	BIC
	β_1	0.25	0.271	0.020	0.021	0.011		
EW-PH	β_2	0.35	0.346	0.008	−0.004	0.003		
	β_3	0.45	0.445	0.020	−0.005	0.004	11 029	11 072
	κ	1.00	0.977	0.026	−0.023	0.045		
	ρ	1.65	1.599	0.043	−0.051	0.166		
	υ	1.00	1.060	0.045	0.060	0.124		
	β_1	0.25	0.243	−0.007	0.020	0.003		
GLL-PH	β_2	0.35	0.359	0.009	0.008	0.006		
	β_3	0.45	0.492	0.042	0.020	0.040	11 065	11 108
	κ	1.00	1.645	0.645	0.014	1.706		
	ρ	1.65	0.994	−0.656	0.011	1.734		
	υ	1.00	0.023	−0.977	0.048	0.999		
	β_1	0.25	0.011	0.020	−0.239	0.062		
W-PH	β_2	0.35	0.003	0.007	−0.347	0.122		
	β_3	0.45	0.015	0.020	−0.435	0.202	14 417	14 454
	κ	1.00	1.003	0.011	0.003	0.006		
	ρ	1.65	1.657	0.013	0.007	0.023		
	β_1	0.25	0.019	0.020	−0.231	0.062		
G-PH	β_2	0.35	0.004	0.007	−0.346	0.122		
	β_3	0.45	0.009	0.020	−0.441	0.202	15 117	15 153
	κ	1.00	0.862	0.016	−0.138	0.257		
	ρ	1.65	0.585	0.013	−1.065	2.380		
	β_1	0.25	0.250	0.000	0.000	0.000		
EE-PH	β_2	0.35	0.350	0.000	0.000	0.000		
	β_3	0.45	0.450	0.000	0.000	0.000	16 891	16 927
	κ	1.00	2.127	0.028	1.127	3.524		
	ρ	1.65	1.076	0.013	−0.574	1.565		

Table 13
Statistical description of the leukemia and larynx cancer data sets.

Data	Minimum	Q1	Mean	Q2	Variance	SD	Q3	Maximum
Data I	0.003	0.112	1.459	0.507	6.044	2.458	1.467	13.626
Data II	0.100	2	4.198	4	6.901	2.627	6.200	10.700

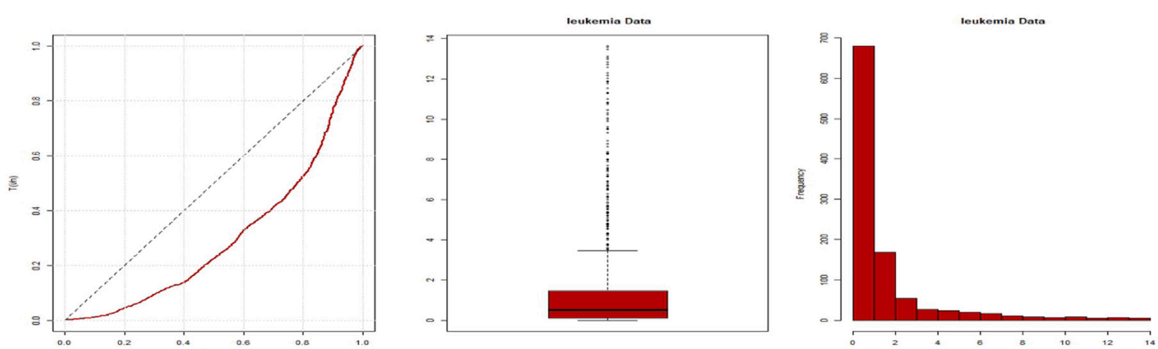

Fig. 8. The TTT, box, and histogram plots for the survival times of leukemia data.

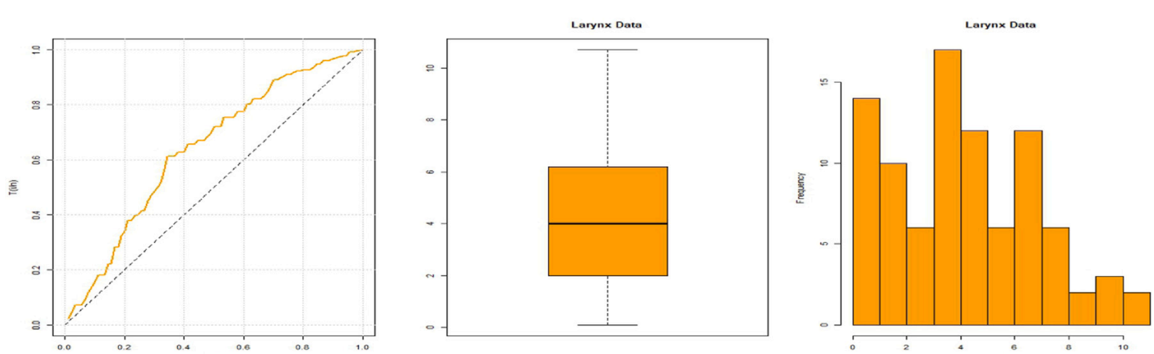

Fig. 9. The TTT, box, and histogram plots for the survival times of the larynx cancer data.

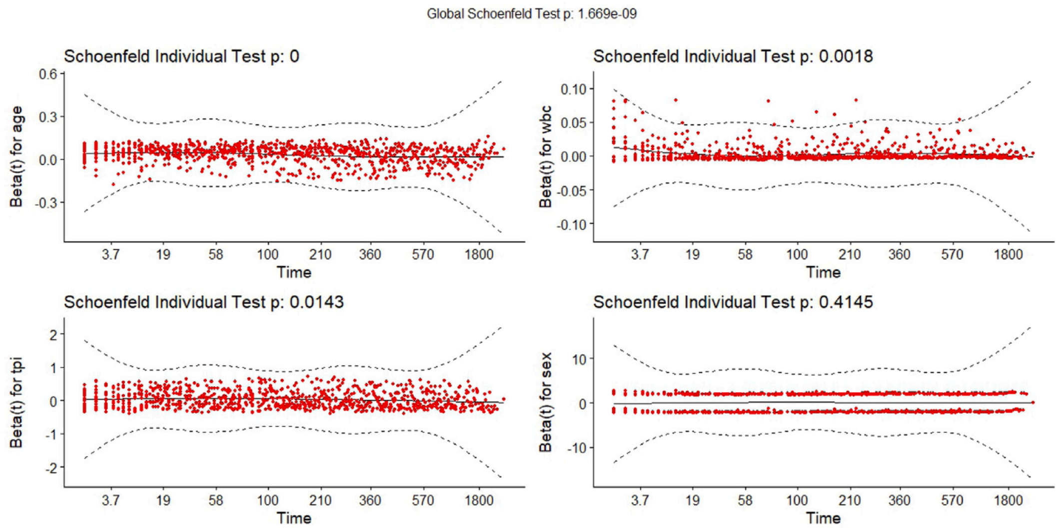

Fig. 10. The SSR for leukemia data considering the *p*-value test for each covariate.

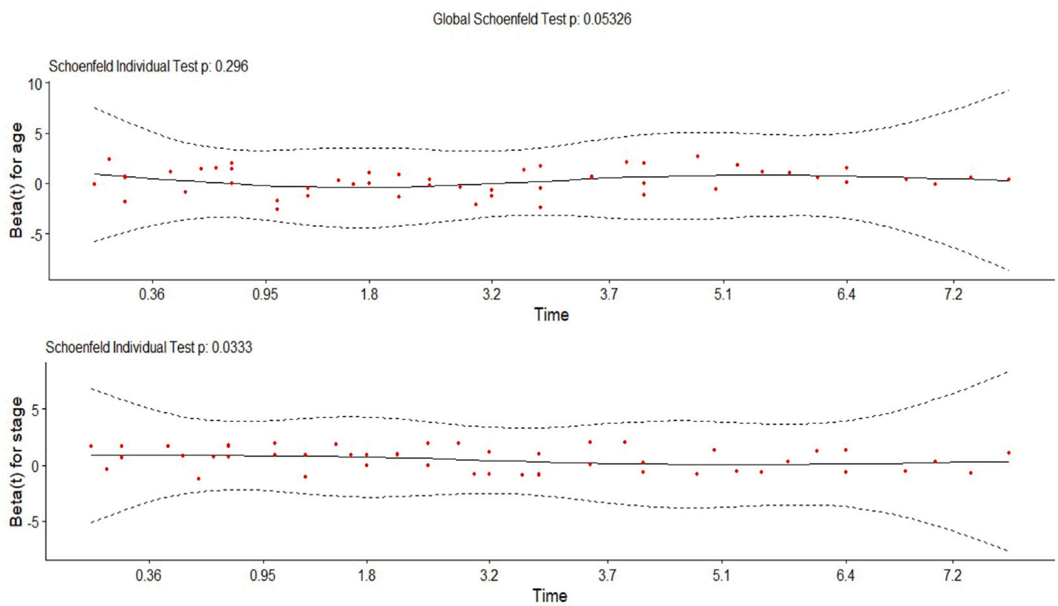

Fig. 11. The SSR for larynx cancer data considering the *p*-value test for each covariate.

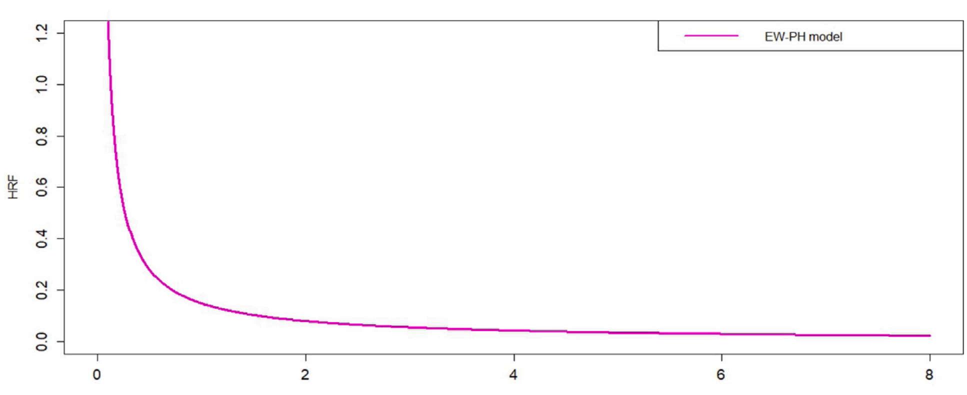

Fig. 12. The estimated HRF of the EW-PH model based on leukemia data set.

Table 14
The MLEs, SE, z-value, CI, and p-value at the level of significance 5% for each model, and analytical measures for different survival models.

Model	Parameter	MLE	SE	z-value	CI 95%	p-value	AIC	BIC	CAIC	HQIC
	β_1	0.029	0.002	13.888	(0.025, 0.033)	0.000	1581	1616	1623	1595
EW-PH	β_2	0.003	0.000	6.625	(0.002, 0.004)	0.000				
	β_3	0.027	0.009	2.950	(0.009, 0.044)	0.003				
	β_4	0.056	0.068	0.826	(−0.077, 0.188)	0.040				
	κ	0.045	0.054	0.843	(−0.060, 0.151)	0.039				
	ρ	0.146	0.016	8.962	(0.114, 0.178)	0.000				
	υ	8.305	1.842	4.509	(4.695, 11.915)	0.000				
	β_1	0.028	0.002	14.021	(0.025, 0.033)	0.000	1582	1617	1625	1596
GLL-PH	β_2	0.003	0.000	6.852	(0.002, 0.004)	0.000				
	β_3	0.028	0.009	3.060	(0.010,0.045)	0.003				
	β_4	0.056	0.068	0.821	(−0.078, 0.189)	0.412				
	κ	0.764	0.034	22.501	(0.698, 0.831)	0.000				
	ρ	0.142	0.038	3.755	(0.068, 0.217)	0.000				
	υ	1.047	0.339	3.089	(0.383, 1.712)	0.002				
	β_1	0.004	0.001	3.136	(0.002, 0.007)	0.002	1874	1904	1910	1886
W-PH	β_2	−0.001	0.001	2.079	(−0.003, 0.000)	0.380				
	β_3	0.033	0.009	3.807	(0.016, 0.050)	0.000				
	β_4	0.111	0.064	1.726	(−0.015, 0.236)	0.840				
	κ	1.753	0.246	7.115	(1.270, 2.236)	0.000				
	ρ	0.532	0.015	36.468	(0.504, 0.561)	0.000				
	β_1	0.023	0.001	21.695	(0.021, 0.026)	0.000	4010	4039	4046	4022
G-PH	β_2	0.004	0.000	7.807	(0.003, 0.004)	0.000				
	β_3	−0.058	0.007	7.689	(−0.072, −0.043)	0.000				
	β_4	2.833	0.071	40.056	(2.694, 2.971)	0.000				
	κ	0.012	0.009	1.319	(−0.006, 0.029)	0.187				
	ρ	0.028	0.001	25.653	(0.026, 0.030)	0.000				
	β_1	−0.024	0.001	18.014	(−0.026, −0.021)	0.000	2428	2458	2464	2439
EE-PH	β_2	−0.008	0.001	8.124	(−0.010, −0.006)	0.000				
	β_3	−0.048	0.009	5.468	(−0.065, −0.031)	0.000				
	β_4	2.721	0.071	38.462	(2.582, 2.859)	0.000				
	κ	1.348	0.045	29.639	(1.259, 1.437)	0.000				
	ρ	1.310	0.054	24.050	(1.203, 1.416)	0.000				

Table 15
The MLEs, SE, z-value, CI, and p-value at the level of significance 5% for each model, and analytical measures for different survival models.

Model	Parameter	MLE	SE	z-value	CI 95%	p-value	AIC	BIC	CAIC	HQIC
	β_1	0.257	0.155	1.652	(−0.048, 0.561)	0.090	291.381	301.880	305.880	295.421
EW-PH	β_2	0.514	0.139	3.699	(0.242, 0.787)	0.000				
	κ	20.043	33.413	0.600	(−45.444, 85.531)	0.045				
	ρ	1.015	2.575	0.394	(−4.031, 6.061)	0.069				
	υ	1.093	3.244	0.337	(−5.266, 7.452)	0.037				
	β_1	0.299	0.157	1.908	(−0.008, 0.607)	0.056	291.905	302.404	306.404	295.945
GLL-PH	β_2	0.543	0.134	4.038	(0.279, 0.806)	0.000				
	κ	1.167	0.276	4.233	(0.627, 1.708)	0.000				
	ρ	0.054	0.039	1.376	(−0.023, 0.131)	0.169				
	υ	0.065	0.223	0.292	(−0.372, 0.502)	0.771				
	β_1	0.256	0.152	1.692	(−0.041, 0.554)	0.091	293.718	303.717	307.717	297.751
W-PH	β_2	0.464	0.126	3.666	(0.216,0.712)	0.000				
	κ	17.323	4.798	3.610	(7.919, 26.728)	0.000				
	ρ	1.123	0.130	8.612	(0.867, 1.378)	0.000				
	β_1	0.247	0.156	1.581	(−0.059, 0.553)	0.114	293.571	303.570	307.570	297.603
G-PH	β_2	0.472	0.132	3.576	(0.213, 0.730)	0.000				
	κ	0.025	0.065	0.382	(−0.103, 0.153)	0.702				
	ρ	0.042	0.017	2.429	(0.008, 0.076)	0.015				
	β_1	0.153	0.131	1.163	(−0.105, 0.411)	0.245	292.571	301.570	306.570	296.603
EE-PH	β_2	−0.198	0.058	3.429	(−0.312, −0.085)	0.001				
	κ	47.430	6.708	7.071	(34.283, 60.576)	0.000				
	ρ	13.716	12.480	1.099	(−10.745, 38.177)	0.272				

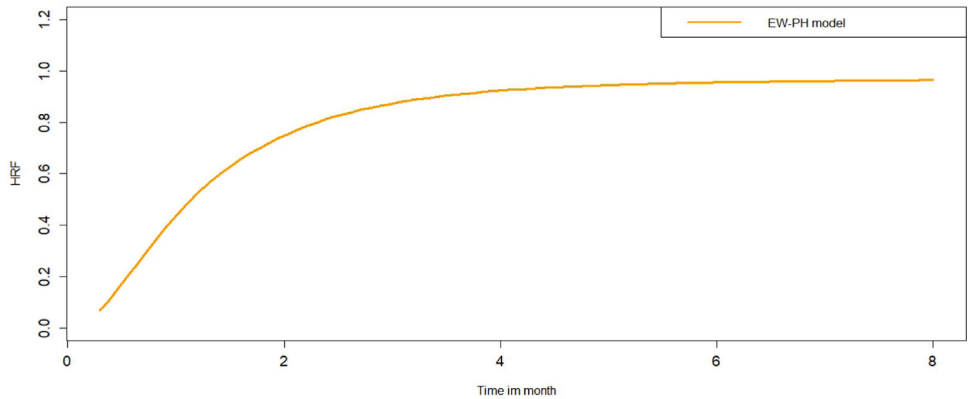

Fig. 13. The estimated HRF of the EW-PH model based on larynx cancer data.

censorship and truncated observations, including, left, interval, and double censoring, could be also employed.

CRediT authorship contribution statement

Mohamed A.S. Ishag: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Anthony Wanjoya:** Writing – review & editing, Supervision, Software, Resources, Methodology, Investigation, Conceptualization. **Aggrey Adem:** Writing – review & editing, Supervision, Software, Resources, Methodology, Investigation, Conceptualization. **Rehab Alsultan:** Writing – review & editing, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Abdulaziz S. Alghamdi:** Writing – review & editing, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition. **Ahmed Z. Afify:** Writing – review & editing, Validation, Supervision, Software, Resources, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data sets are mentioned along in the paper.

Acknowledgments

The first author would like to thank the Pan African University, Institute for Basic Sciences, Technology, and Innovation (PAUSTI). All authors have read and agreed to the published version of the manuscript.

References

[1] S. Guo, Survival Analysis, Oxford University Press, 2010.
[2] P.W. Withana Gamage, C.S. McMahan, L. Wang, A flexible parametric approach for analyzing arbitrarily censored data that are potentially subject to left truncation under the proportional hazards model, Lifetime Data Anal. 29 (1) (2023) 188–212.
[3] M. Cockeran, S.G. Meintanis, J.S. Allison, Goodness-of-fit tests in the Cox proportional hazards model, Comm. Statist. Simulation Comput. 50 (12) (2021) 4132–4143.
[4] J. Guo, W. Kang, Y. Wang, Multi-perspective option price forecasting combining parametric and non-parametric pricing models with a new dynamic ensemble framework, Technol. Forecast. Soc. Change 204 (2024) 123429.
[5] A.B.S. Mastor, A.S. Alghamdi, O. Ngesa, J. Mung'atu, C. Chesneau, A.Z. Afify, The extended exponential-Weibull accelerated failure time model with application to Sudan COVID-19 Data, Mathematics 11 (2) (2023) 460.
[6] A.H. Muse, S. Mwalili, O. Ngesa, H.M. Alshanbari, S.K. Khosa, E. Hussam, Bayesian and frequentist approach for the generalized log-logistic accelerated failure time model with applications to larynx-cancer patients, Alex. Eng. J. 61 (10) (2022) 7953–7978.
[7] F.J. Rubio, L. Remontet, N.P. Jewell, A. Belot, On a general structure for hazard-based regression models: an application to population-based cancer research, Stat. Methods Med. Res. 28 (8) (2019) 2404–2417.
[8] X. Wang, Y.R. Yue, J.J. Faraway, Bayesian Regression Modeling with INLA, CRC Press, 2018.
[9] M.H. Tahir, G.M. Cordeiro, Compounding of distributions: a survey and new generalized classes, J. Stat. Distrib. Appl. 3 (2016) 1–35.
[10] S. Lee, P.C. Lambert, M.J. Sweeting, N.R. Latimer, M.J. Rutherford, Evaluation of flexible parametric relative survival approaches for enforcing long-term constraints when extrapolating all-cause survival, Value Health 27 (1) (2024) 51–60.
[11] D. Collett, Modelling Survival Data in Medical Research, CRC press, 2015.
[12] K. Burke, M. Jones, A. Noufaily, A flexible parametric modelling framework for survival analysis, J. R. Stat. Soc. Ser. C. Appl. Stat. 69 (2) (2020) 429–457.
[13] N. Taketomi, K. Yamamoto, C. Chesneau, T. Emura, Parametric distributions for survival and reliability analyses, a review and historical sketch, Mathematics 10 (20) (2022) 3907.
[14] H. Rehman, N. Chandra, F.S. Hosseini-Baharanchi, A.R. Baghestani, M.A. Pourhoseingholi, Cause-specific hazard regression estimation for modified Weibull distribution under a class of non-informative priors, J. Appl. Stat. 49 (7) (2022) 1784–1801.
[15] S.A. Khan, S.K. Khosa, Generalized log-logistic proportional hazard model with applications in survival analysis, J. Stat. Distrib. Appl. 3 (2016) 1–18.
[16] A.H. Muse, C. Chesneau, O. Ngesa, S. Mwalili, Flexible parametric accelerated hazard model: Simulation and application to censored lifetime data with crossing survival curves, Math. Comput. Appl. 27 (6) (2022) 104.
[17] S.K. Khosa, Parametric Proportional Hazard Models with Applications in Survival Analysis (Ph.D. thesis), University of Saskatchewan, 2019.
[18] G.S. Mudholkar, D.K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure-rate data, IEEE Trans. Reliab. 42 (2) (1993) 299–302.
[19] J. Abubakar, M.A.A. Abdullah, O.R. Olaniran, Variational Bayesian inference for exponentiated Weibull right censored survival data, Stat. Optim. Inf. Comput. 11 (4) (2023) 1027–1040.
[20] C. Robert, R. Christensen, W. Johnson, A. Branscum, T. Hanson, Bayesian Ideas and Data Analysis, Taylor & Francis, 2012.
[21] G.S. Mudholkar, D.K. Srivastava, M. Freimer, The exponentiated Weibull family: A reanalysis of the bus-motor-failure data, Technometrics 37 (4) (1995) 436–445.
[22] S.A. Khan, Exponentiated Weibull regression for time-to-event data, Lifetime Data Anal. 24 (2) (2018) 328–354.
[23] S. Nadarajah, G.M. Cordeiro, E.M. Ortega, The exponentiated Weibull distribution: a survey, Statist. Papers 54 (2013) 839–877.
[24] J.F. Lawless, Statistical Models and Methods for Lifetime Data, John Wiley & Sons, 2011.
[25] F.J. Rubio, B. Rachet, R. Giorgi, C. Maringe, A. Belot, On models for the estimation of the excess mortality hazard in case of insufficiently stratified life tables, Biostatistics 22 (1) (2021) 51–67.
[26] R. Ancuceanu, M.V. Hovanet, A.I. Anghel, F. Furtunescu, M. Neagu, C. Constantin, M. Dinu, Computational models using multiple machine learning algorithms for predicting drug hepatotoxicity with the DILIrank dataset, Int. J. Mol. Sci. 21 (6) (2020) 2114.
[27] Z. Zhang, Parametric regression model for survival data: Weibull regression model as an example, Ann. Transl. Med. 4 (24) (2016).

[28] S.R. Haile, J.-H. Jeong, X. Chen, Y. Cheng, A 3-parameter Gompertz distribution for survival data with competing risks, with an application to breast cancer data, J. Appl. Stat. 43 (12) (2016) 2239–2253.
[29] H.S. Bakouch, M.M. Ristić, A. Asgharzadeh, L. Esmaily, B.M. Al-Zahrani, An exponentiated exponential binomial distribution with application, Statist. Probab. Lett. 82 (6) (2012) 1067–1081.
[30] A. Muse, S. Mwalili, O. Ngesa, M. Kilai, AHSurv: An R package for flexible parametric accelerated hazards (AH) regression models, 2022.
[31] R. Henderson, S. Shimakura, D. Gorst, Modeling spatial variation in leukemia survival data, J. Amer. Statist. Assoc. 97 (460) (2002) 965–972.
[32] F.J. Rubio, R. Drikvandi, MEGH: A parametric class of general hazard models for clustered survival data, Stat. Methods Med. Res. 31 (8) (2022) 1603–1616.
[33] D. Alvares, E. Lázaro, V. Gómez-Rubio, C. Armero, Bayesian survival analysis with BUGS, Stat. Med. 40 (12) (2021) 2975–3020.
[34] O. Kardaun, Statistical survival analysis of male larynx-cancer patients-a case study, Stat. Neerl. 37 (3) (1983) 103–125.
[35] H. Bozdogan, Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions, Psychometrika 52 (3) (1987) 345–370.
[36] E.J. Hannan, B.G. Quinn, The determination of the order of an autoregression, J. R. Stat. Soc. Ser. B Stat. Methodol. 41 (2) (1979) 190–195.