

Research Article

Censoring Balancing Functions for Undetected Probably Significant Effects in Cox Regression

Ildephonse Nizeyimana ^{1,2} **George Otieno Orwa** ³ **Michael Arthur Ofori** ¹
and Samuel Musili Mwalili ³

¹Pan African University Institute for Basic Sciences, Technology and Innovation, Nairobi, Kenya

²University of Rwanda, Kigali, Rwanda

³Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Correspondence should be addressed to Ildephonse Nizeyimana; ildenze01@gmail.com

Received 23 August 2023; Revised 11 September 2023; Accepted 14 September 2023; Published 26 September 2023

Academic Editor: Niansheng Tang

Copyright © 2023 Ildephonse Nizeyimana et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Weighted Cox regression models were proposed as an alternative to the standard Cox proportional hazards models where consistent estimators can be obtained with more relative strength compared to unweighted cases. We proposed censoring balancing functions which can be built in a way that allows us to obtain the maximum possible significant treatment effects that may have gone undetected due to censoring. The harm caused by this is compensated and new weighted parameter estimates are found. These functions are constructed to be monotonic because even the hazard ratios are not exactly constant as in the ideal case, but are violated by monotonic deviations in time. For more than one covariate, even the interaction between covariates in addition to censoring can lead to the loss of significance for some covariates' effects. Undetected significant effects of one covariate can be obtained, still keeping the significance and approximate size of the remaining one(s). This is performed by keeping the consistency of the parameter estimates. The results from both the simulated datasets and their application to real datasets supported the importance of the suggested censoring balancing functions in both one covariate and more than one covariate cases.

1. Introduction

1.1. Background. Survival data analysis has many applications in the real world such as engineering like testing the lifetime of life bulbs and medicine like testing the efficiency of different treatments and it even found its role in social sciences. Cox proportional hazard models have been hugely exploited since their introduction by Cox [1] and many modifications have been made with the aim of improving the models' accuracy. In its nature, a Cox model is built based on the hazard (rate) of a given event of interest in a specific group and the two hazard models corresponding to the two groups under comparison are used to find the hazard ratio. This is also referred to as the risk ratio, and in Cox proportional hazards models, it is assumed to be proportional over the whole time of the study.

The hazard rate is sometimes assumed to be associated with covariates which can be continuous or categorical and their effects are deducted from the computed regression parameters. Recently, weighted Cox models attracted the attention of many, and the weights used are the same as those employed in weighted logrank tests. We even recall that the weighted logrank tests are derived from the partial likelihood of the weighted Cox models [2]. In our recent work, censoring balancing functions were introduced to balance the negative effects of censoring on the power of such tests and they were found to be of some remarkable improvement. In this work, such functions will be employed to investigate covariate effects. Yu et al. [3] used what they referred to as "censoring correction" in weighted Cox regression by using the inverse probability distribution of censoring, $G(t)^{-1}$ combined with the selected weight where $G(t)$ is the Kaplan-Meier estimate of

censoring. In their work, they also explored the average hazards ratio and the same technique had been used by Schemper et al. [4].

1.2. Parameter Estimation in Cox Regression. The general form of a Cox proportional hazard model as introduced by Cox [1] and widely used in many works such as [5], [6], and [7] is expressed as

$$\lambda(t|X) = \lambda_0(t) \exp(X\beta), \quad (1)$$

where X is a vector of covariates and β is the Cox regression parameter for the covariates and $\lambda_0(t)$ is the baseline hazard function which can be estimated from any parametric distribution. If the vector of covariates X is made of k covariates, such that $X = (X_1, X_2, \dots, X_k)$, then β will also be of the same size since for each component X_i , there must be the corresponding regression coefficient β_i .

If one is interested in the hazard ratio, it can be immediately deducted from the two hazards with the same baseline hazard without knowing or estimating this since it may be canceled out in both quantities. The covariates can be constant or time-dependent, and in the latter case, they are written as $X(t)$. The regression coefficient β can also be time-dependent. We highlight that when exploring the Cox proportional hazards models, the main interest is in the failure time, let us say, T . Thus, the hazard $\lambda(t|X)$ can be understood as the rate of failure for the objects that still survive till then.

If t is a failure time, then the likelihood that it is the specific subject i that fails at that time is given by

$$L_i(t, \beta) = \frac{\exp(X_i\beta)}{\sum_{j \in R_i} \exp(X_j\beta)}, \quad (2)$$

where R_i is the set of all subjects at risk at time t .

The resulting partial likelihood is obtained by multiplying these individual probabilities. That is,

$$L(t, \beta) = \prod_{i=1}^M \frac{\exp(X_i\beta)}{\sum_{j \in R_i} \exp(X_j\beta)}. \quad (3)$$

When we take a more general case where there may be ties in occurrences of events, M event times which were observed, will be reduced to m tied event times, where we will denote by S_i , i.e., the sum of covariates' X_i of all individuals who underwent the event of interest at time t_i and equation (3) will then be rewritten as

$$L(t, \beta) = \prod_{i=1}^m \frac{\exp(S_i\beta)}{\sum_{j \in R_i} \exp(X_j\beta)}. \quad (4)$$

In our next deductions and computations, we will refer to the notation in equation (3) since it is the one generally used in other research works. Survival data are practically characterized by censoring and the likelihood expression can be rewritten otherwise once the times considered are not specified as the times of occurrence for the event of interest.

This is so because they may observe either the event of interest (such as death, failure, or any other) or the censoring. So, if the censoring times are also recorded, being known that the likelihood is considered and estimated only at time points of occurrence of the event of interest, Lin [5] and Fisher and Lin [8], among others, rewrote equation (3) as

$$L(t, \beta) = \prod_{i=1}^N \left(\frac{\exp(X_i\beta)}{\sum_{j \in R_i} \exp(X_j\beta)} \right)^{\delta_i}, \quad (5)$$

where δ_i is the indicator function and it is such that $\delta_i = 1$ if it is the event of interest that occurred at time t_i and 0 in the case of censoring.

The aim is to find the regression parameter β which maximizes the partial log-likelihood function and this is computed as follows:

$$l(t, \beta) = \log(L(t, \beta)) \\ = \sum_{i=1}^N \delta_i \left[X_i\beta - \log \left(\sum_{j \in R_i} \exp(X_j\beta) \right) \right]. \quad (6)$$

Note that, M is the number of distinct times of occurrence for the event of interest [9] while N is the total number of times including the censoring ones. Any logic from either equation (6) or (7) is understandable because at the censoring time, the corresponding quantity in equation (5) is raised to power 0 and it gives 1, which, once multiplied by other quantities, brings no change. In the same way, for the log partial likelihood, the sum of logarithms will remain the same since the quantities corresponding to censoring are multiplied by 0.

In the next transformations, we will continue with the form of partial log-likelihood which does not contain δ_i and we get

$$l(t, \beta) = \sum_{i=1}^M \left[X_i\beta - \log \left(\sum_{j \in R_i} \exp(X_j\beta) \right) \right]. \quad (7)$$

To maximize $l(t, \beta)$ in equation (7), we need to find $\hat{\beta}$ which is a solution to the following equation:

$$\frac{\partial l(t, \beta)}{\partial \beta} = 0. \quad (8)$$

Equation (8) is called the score statistic from the partial log-likelihood, sometimes referred to as $U(t, \beta)$ and is written as

$$U(t, \beta) = \sum_{i=1}^M \left[X_i - \frac{\sum_{j \in R_i} \exp(X_j\beta) X_j}{\sum_{j \in R_i} \exp(X_j\beta)} \right]. \quad (9)$$

From equation (9), the concept of using weight functions to obtain weighted covariate effects was introduced and worked on by many authors including Fisher and Lin [8], León et al. [10], Lin and Wei [11], Lin [5], Murphy et al. [12], and Sasieni [7].

The weighted score function is obtained from equation (9) where the same weight functions as those used in weighted logrank tests are used. Moreover, the weight function which gives more power to the weighted logrank test is the one also preferred to obtain the regression parameter $\hat{\beta}_{\max}$ once $U_w(t, \beta) = 0$ is solved for β . The weighted score function is expressed as follows:

$$U_w(t, \beta) = \sum_{i=1}^M w(t) \left[X_i - \frac{\sum_{j \in R_i} \exp(X_j \beta) X_j}{\sum_{j \in R_i} \exp(X_j \beta)} \right]. \quad (10)$$

1.3. Censoring in Cox Regression. Censoring is known as the special characteristic of survival data which is also a major threat at the same time. It is always reported to negatively affect the results of both weighted logrank tests and Cox regression. Weighted Cox regression was adopted to supplement the usual Cox regression and its parameter estimates are still consistent just as for the case of nonweighted Cox regression provided that the proportionality of hazard holds. Treatment effects can be underestimated or overestimated in the presence of censoring and the deviation increases as the censoring level goes higher.

In Cox regression, hazard rates and survival times of different (two or more groups) groups are considered. As stated by Persson and Khamis [13], censoring distribution plays a big role in the negative effects being observed. In their work, the types of censoring under consideration were early, late, or random censoring. All types of censoring heavily affect the estimates when their proportions are relatively higher and the worst case is observed for early censoring when the proportionality of hazards is observed [13]. Sasieni [7] highlighted that allocating lower weights to latter failures through artificial censoring enhances the relative influence of the outlier and this leads to a higher bias. In our present work, our analysis was performed under the consideration of random censoring. Dunkler et al. [9] stated that the unweighted Cox estimates are relatively more efficient than the weighted ones even though they overtake when the proportional hazards scenario is not met. Moreover, in contrast to the unweighted estimates, the weighted ones remain fairly constant irrespective of the size and type of censoring. This relative strength is still kept while calculating the concordance probability because the unweighted estimate tends to underestimate it. Again, even though the unweighted estimate performs well in the PH case, the weighted one remains acceptable since it does not introduce bias.

The main problem being addressed in this work is the following: "Aren't there cases where negative or positive prognoses are found not significant because of censoring while they probably or really are?" If so, what might they be? If they were reported as significant, are they the real ones? What may be their possible maximum? Can't low effects be interfered by other covariates in addition to censoring? If yes, can their real strength be optimally detected without affecting other covariates?

2. Methods

2.1. Inverse Probability of Censoring Weighting (IPCW) Method. When other conditions are fixed, the higher the censoring rate, the higher the loss of power of the analysis, or simply the higher the bias of the estimates. Schemper et al. [4] used the Fleming–Harrington (FH (ρ, γ)) family of weights and the weight which gave higher power in the weighted logrank tests was the one to be used for the computation of β in Cox regression. To compensate the harm from censoring, Schemper et al. [4] attached on the weight the inverse of the censoring distribution function, often denoted by $G(t)$, and the resulting weight was $w(t)G(t)^{-1}$. $G(t)$ is found the same way as the usual Kaplan–Meier estimates $S(t)$ are found but with the reversed status variable. It is understood that $G(t)$ is decreasing as censoring is observed and hence $G(t)^{-1}$ is increasing and is at least equal to 1.

In Yu et al.'s study [3], just as in Schemper et al.'s study [4], the same quantity $G(t)^{-1}$ was investigated and used while referred to as "censoring correction." It was still applied in León et al. [10] and Dunkler et al.'s study [9] where it was annexed to different weights that were explored to find the maximum of treatment effects. It had also been employed separately as an independent weight function by Xu and O'Quigley [14]. All those transformations were performed with the aim of maximizing the efforts in restoring the ability to capture the treatment effects, which in most cases is referred to as the hazard ratio. The concept of weighting in Cox regression led to the name "average hazards ratio" or "average regression effects" where e^β is like the weighted mean. Weighting by $G(t)^{-1}$ was also explored in logrank tests where it is referred to as the inverse of the probability of censoring weight (IPCW) and it is used to address some biases due to censoring in different estimations such as survival probabilities. In general, IPCW has been and is still employed as a response to noncompliance from censoring as found in the studies of Kvamme and Borgan [15], Matsouaka and Atem [16], Satten and Datta [17], Wakounig et al. [18], and Robins and Finkelstein [19]. Dong et al. [20] found that an estimate of treatment effects obtained through the IPCW-adjusted win ratio statistic is unbiased. However, Howe et al. [21] demonstrated its limitations in the presence of a strong selection bias or in the case of unmeasured common predictors, sample size, or model misspecification.

2.2. Censoring-Balanced Weight Function. Censoring balancing functions were proposed to balance the negative effects of censoring on the respective results. In this work, they will be implied in the computation of the Cox regression coefficients which help to measure the covariates' effects or even hazards ratio. The proposed censoring balancing function is defined as

$$f(c(t)) = (1 + (t))^r, \quad (11)$$

where $c(t)$ indicates the censoring rate until event time t . The parameter “ r ” can vary and take on any value in \mathbb{Z} depending on the researcher’s judgment and aim. The aim of this study as mentioned before is finding the ideal maximum possible value (significance) of the treatment effects for the univariate (multivariate) case. Negative values reduce the size of the original weight and hence will act like penalties whereas the positive ones improve it which makes them act like compensations.

As the IPCW through $G(t)^{-1}$ on the existing weights brings some improvement, the newly proposed censoring balancing functions are built in a more flexible way which gives them the ability to be modified differently to have the desired nature and size. This will make their use more friendly and adaptive.

The general censoring-balanced weight function will be of the following form:

$$W_{CB}(t) = (1 + c(t))^r w(t). \quad (12)$$

The index CB is used to indicate that the weight under consideration is censoring-balanced. $w(t)$ is the initial weight function and in our case, we will focus on $w(t) = 1$ as the initial weight and hence will apply different levels of compensations by manipulating the censoring balancing function as given above through varying the parameter r . Our main focus is on the average hazard ratio obtained under a specific weight function. The negative impact of censoring will be dealt with to see if there is any improvement brought by the censoring balancing function (CBFs) on the quantity $\exp(\beta)$ which stands for the average hazard ratio. Under random censoring, it is more likely that the treatment effects are underestimated. The proposed form of censoring balancing weight function will help to compensate for that loss of sensitivity. However, the loss can be overcompensated since in real applications, the true treatment effects are not known. So, we will compute the expected maximum possible treatment effects which may even exceed the true treatment effects. To understand this, a simulated dataset will be used to describe the situation. Since the monotone deviations from the proportionality are the most encountered, the proposed functions will be combined with $w(t) = 1$ because they themselves will act like weights. So, we do not need to combine them with another weight now, but it can be explored afterward. Therefore, the weighted Cox regression that will be dealt with will be from the weight function of the following form:

$$W_{CB}(t) = (1 + c(t))^r. \quad (13)$$

2.3. Application of the Goodness-of-Fit Test. The assumption of proportional hazards can be tested against linear monotone departures where the regression coefficient β will be replaced by a linear monotone function of time, let us say, $a(t)$. Hence, the hypothesis which will be tested is $H_0: \lambda(t, X) = \lambda_0(t) \exp(X(t)\beta)$ against $H_1: \lambda(t, X) = \lambda_0(t) \exp(X(t)a(t))$.

$a(t)$ is assumed to be monotone because the monotone departures from the proportional hazards assumptions are the most generally encountered ones in real scenarios of model misspecification. From this, monotone weight functions are the most appropriate since they are more sensitive to such monotone departures. The weighted Cox regression was found to be efficient when some deviation from the proportionality assumption was observed. Lin [5] proposed a test for goodness-of-fit analysis which remains consistent even under model misspecification and helps to detect the difference between the two Cox regression parameters which are β and β_w .

So, the same method can also be employed to test the following hypothesis:

$$H_0: \beta_w = \beta \text{ against } H_1: \beta_w \neq \beta. \quad (14)$$

They stated that when the proportionality assumption is held, the two parameters do not differ significantly in terms of magnitude or absolute value. We recall that in this work, we will consider different levels of censoring balancing functions. The proposed test is of the following form:

$$Q_w = n(\widehat{\beta}_w - \widehat{\beta})' D_w(\widehat{\beta})^{-1} (\widehat{\beta}_w - \widehat{\beta}), \quad (15)$$

where

- (i) $\widehat{\beta}_w$ is the weighted regression parameter estimate,
- (ii) $\widehat{\beta}$ is the unweighted regression parameter estimate,
- (iii) $D_w(\widehat{\beta})$ is the difference of the covariance matrices. i.e., $D_w(\widehat{\beta}) = C(\widehat{\beta}_w) - C(\widehat{\beta})$.

$C(\widehat{\beta}_w)$ and $C(\widehat{\beta})$ are the covariance matrices from $\widehat{\beta}_w$ and $\widehat{\beta}$, respectively, and they are computed as below. The basic expression to be define is

$$S^{(r)}(\beta, t) = \frac{1}{n} \sum_{i=1}^n \exp(X_i(t)\beta) \cdot X_i(t)^{\otimes r}, \quad (16)$$

where $r = 0, 1, 2$ and $X_i(t)$ is the covariates vector for the individual i whose event time is t (it does not mean that the covariates are time-dependent).

The following quantities are deduced from $S^{(r)}(\beta, t)$ according to r :

$$S^{(0)}(\beta, t) = \frac{1}{n} \sum_{i=1}^n \exp(X_i(t)\beta),$$

$$S^{(1)}(\beta, t) = \frac{1}{n} \sum_{i=1}^n \exp(X_i(t)\beta) \cdot X_i(t), \quad (17)$$

$$S^{(2)}(\beta, t) = \frac{1}{n} \sum_{i=1}^n \exp(X_i(t)\beta) \cdot X_i(t)^{\otimes 2}.$$

With $E(\beta, t)$ defined as $E(\beta, t) = S^{(1)}(\beta, t) / S^{(0)}(\beta, t)$, the score function defined in equation (10) can be rewritten as

$$U_w(t, \beta) = \sum_{i=1}^M w(t) [X_i(t) - E(\beta, t)]. \quad (18)$$

The resulting information matrix from which we will get the standard errors for the respective components of β is obtained by the second derivative of the partial log-likelihood function and in terms of $S^{(r)}$'s, we have

$$V(\beta, t) = \left[\frac{S^{(2)}(\beta, t)}{S^{(0)}(\beta, t)} - \frac{S^{(1)}(\beta, t)^{\otimes 2}}{(S^{(0)}(\beta, t))^2} \right]. \tag{19}$$

By using the $E(\beta, t)$ defined previously, $V(\beta, t)$ can be written as

$$V(\beta, t) = \left[\frac{S^{(2)}(\beta, t)}{S^{(0)}(\beta, t)} - E(\beta, t)^{\otimes 2} \right]. \tag{20}$$

To make the notation more logically understandable to the reader, we will write $V(\beta, X_i)$ instead of $V(\beta, t)$. This is because the summation is made over event times and the subscript i will make the summation more logical. The variance of the weighted regression estimator is computed as

$$C_w(\hat{\beta}) = A_w(\hat{\beta})^{-1} B_w(\hat{\beta}) A_w(\hat{\beta})^{-1}, \tag{21}$$

with

$$A_w(\hat{\beta}) = \frac{1}{M} \sum_{i=1}^M w(X_i) V(\beta, X_i), \tag{22}$$

$$B_w(\hat{\beta}) = \frac{1}{M} \sum_{i=1}^M w(X_i)^2 V(\beta, X_i).$$

It is easily noticed that $A_w(\hat{\beta})$ and $B_w(\hat{\beta})$ are different only in the sense that they are computed with weights $w(X_i)$ and $w(X_i)^2$, respectively.

2.4. Optimization of the Parameter. Equation (10) will be solved for many options to obtain the optimum parameter. For a single covariate, the effects can be found significant or not and, in that case, we will find their probable maximum value since censoring is expected to have affected it. For two or more covariates, we try to maintain the parameter component which was found to be significant and we try to find the solution for equation (10) under different values of r where the optimal solution will be the one with a small deviation from the already significant component ($\beta_r[j] - \beta[j]$) or with nearly the same norm as the initial parameter ($|\beta_r| = |\beta|$) but with the significance of effects for targeted covariates, $\beta[i]$. In short, the optimization is performed as follows.

For a single covariate, the optimization is performed as

$$\beta_{\max} = \max_r \{ \beta_r : U_{w_r}(t, \beta) = 0 \}. \tag{23}$$

For two or more covariates, the optimization is performed as

$$\beta_{\text{opt}} = \max_r \{ |\beta_r| : U_{w_r}(t, \beta) = 0 \text{ and } \beta_r[i] \text{ is significant} \}, \tag{24}$$

or

$$\beta_{\text{opt}} = \min_r \{ \beta_r[j] - \beta[j] : U_{w_r}(t, \beta) = 0 \text{ and } \beta_r[i] \text{ is significant} \}. \tag{25}$$

Note that, we avoid the notation β_{\max} but used β_{opt} for the case of more than one covariate because the aim is not the maximum in size but the significance of effects for as many as possible covariates. León et al. [10] computed β_{\max} from a set of weights of the Fleming–Harrington ($G^{p,\gamma}$) family, but in this study, it will be found by varying r . For each dataset or case, the optimal r is reported together with the censoring level where it gave the intended optimal solution which was also reported and compared with the initially obtained parameter.

3. Data Description and Analysis

We dealt with two simulated datasets from a Weibull distribution and showed the proportional hazards nature and two real ones. For each of the simulated datasets, the total sample size is 200 with a 1:1 allocation ratio. We hence assumed three levels of censoring which are 20%, 40%, and 60%. The censoring times were assumed randomly from the previously simulated time points. For each censoring level, the regression parameter was estimated. In the case of one covariate, the level of censoring balancing functions which yields the absolute maximum value of the parameter was obtained and this was compared with the one obtained under censoring. For two covariate cases, the alternative regression parameter was estimated with the aim of small deviation from the significant component targeting the parameter whose all components are significant. Once the targeted weighted parameter was obtained, it was compared with the one previously obtained under censoring using the stated test to see if the obtained parameter is still consistent.

For real data applications, two datasets were used. Those are the colon cancer dataset and the lung cancer dataset which are freely accessible online and can be imported to different data analysis software. The colon dataset has a total sample size of 1858 with an overall censoring rate of 50.48%. The covariate of interest was sex where there were 968 males (sex = 1) and 890 females (sex = 0) who were subjected to three treatments and the variable of interest was the time of the event of interest (death) or recurrence. Since the two groups (from the sex variable) show the PH nature, we applied the Cox regression to investigate the sex covariate effects. The lung cancer dataset is made of an overall sample size of 228 with an overall censoring rate of 27.63% where 90 of them are females (sex = 1 in our analysis) and 138 are males. After performing the analysis with this single covariate, we also analyzed the same dataset with two covariates: “sex” and “age.”

For each dataset (simulated or real one), $\hat{\beta}$ and $\hat{\beta}_w$ were computed, respectively. The parameter from weighted Cox regression $\hat{\beta}_w$ was then compared with the initial parameter from unweighted Cox regression. This is why in the following tables, $D_w(\hat{\beta})$, Q_w , and the corresponding p values are in the row of weighted regression where $r \neq 0$ in the column (c, r). c represents the censoring rate and r is the level of the censoring balancing functions that resulted in the maximum or optimum covariates' effects. The test statistic

TABLE 1: Results from simulation.

c, r	β_w	95% confidence interval	$D_w(\hat{\beta})$	Q_w	P value
<i>For simulated dataset 1 with one covariate "treatment"</i>					
0, 0	0.7411	[0.4434, 1.0388]	—	—	—
20, 0	0.6948	[0.3714, 1.0181]	—	—	—
20, 136	1.2296	[1.2294, 1.2298]	199.51	0.2868	0.5923
40, 0	0.6426	[0.2714, 1.0138]	—	—	—
40, 64	0.8503	[0.8502, 0.8505]	345773.86	0.000025	0.996
60, 0	0.7574	[0.2881, 1.2268]	—	—	—
60, 11	0.7833	[0.6900, 0.8766]	583.14	0.00023	0.988
<i>For simulated dataset 2 with two covariates "sex" and "treatment"</i>					
0, 0	$\begin{bmatrix} -0.1095 \\ 0.6829 \end{bmatrix}$	$\begin{bmatrix} -0.4009, 0.1818 \\ 0.3972, 0.9687 \end{bmatrix}$	—	—	—
20, 0	$\begin{bmatrix} -0.1155 \\ 0.7177 \end{bmatrix}$	$\begin{bmatrix} -0.432, 0.201 \\ 0.406, 1.029 \end{bmatrix}$	—	—	—
20, 16	$\begin{bmatrix} -0.1634 \\ 0.54012 \end{bmatrix}$	$\begin{bmatrix} -0.3195, -0.007 \\ 0.385, 0.695 \end{bmatrix}$	$\begin{bmatrix} 32.991 & -2.324 \\ -2.324 & 34.34 \end{bmatrix}$	0.2962	0.8623
40, 0	$\begin{bmatrix} -0.0421 \\ 0.9634 \end{bmatrix}$	$\begin{bmatrix} -0.407, 0.323 \\ 0.601, 1.326 \end{bmatrix}$	—	—	—
40, 13	$\begin{bmatrix} -0.149 \\ 0.675 \end{bmatrix}$	$\begin{bmatrix} -0.245, -0.053 \\ 0.581, 0.769 \end{bmatrix}$	$\begin{bmatrix} 21.1143 & -1.705 \\ -1.705 & 21.692 \end{bmatrix}$	0.5546	0.7578
60, 0	$\begin{bmatrix} 0.1443 \\ 1.0276 \end{bmatrix}$	$\begin{bmatrix} -0.318, 0.607 \\ 0.5813, 1.474 \end{bmatrix}$	—	—	—
60, 20	$\begin{bmatrix} 0.1312 \\ 0.6843 \end{bmatrix}$	$\begin{bmatrix} 0.1189, 0.1435 \\ 0.6728, 0.6958 \end{bmatrix}$	$\begin{bmatrix} 5.1475 & -1.4403 \\ -1.4403 & 7.9741 \end{bmatrix}$	1.2751	0.5286

Q_w has a central chi-square distribution with degrees of freedom equal to the dimension of the regression parameter. The data analysis task was performed in Python.

4. Results and Discussion

4.1. Simulation Results. 4.2. Application to Real Data. 4.3. Discussion of the Results. For the simulated datasets, the estimates in the noncensoring cases are obtained. As censoring occurs, the deviation from the true coefficient increases and the smaller components in terms of absolute value can even change the sign. For more than one covariate case, while investigating the effects' significance for one covariate, the magnitude of effects for the remaining one(s) may be slightly affected. Low censoring ($\leq 20\%$) gives estimates still close to the real ones as seen in Table 1 and the application of censoring balancing functions may mislead since they tend to overcompensate the small harm from the low censoring. This can be seen in estimates in the cases of 20 and 16 with 0.54012 vs 0.6829 where obtaining a significant component for the failing one will cause a noticeable decrease in the preexisting significant component, and again, on 20 and 136, we have 1.2296 vs 0.7411 to show a higher deviation in the case of one covariate (Table 1). This is supported by the real data application where the lung dataset with sex covariate is treated separately as shown in Table 2, and the coefficient increase was approximately 66% (from -0.5311 to -0.8816), while for the simulated dataset 1, it was nearly 77.5% (from 0.7411 to 1.2296).

The censoring level of 40% gives estimates which are relatively far from the real ones. However, when searching for the possible maximum, it does not go as far as the 20% censoring level. This means that it is the level where the negative impacts of censoring are unavoidable and the censoring balancing functions will play their role more appropriately (refer to rows 40 and 64 with 0.8503 vs 0.7411 and 40 and 13 with 0.675 vs 0.6829 in Table 1).

The covariates with small coefficients are more likely to change the sign when it comes to searching for their significance when the censoring is too high ($\geq 50\%$). This can be checked on row 60 and 20 with 0.1312 vs -0.1095 in Table 1 and is supported by the real application on row 50.48 and 7 with 0.1642 vs -0.0336 in Table 2. The change of sign is logical and can be understood from the fact that non-significant effects are represented by a parameter estimate whose confidence interval includes 0 and hence, it can be either positive or negative while for hypothetical significance, it must be on one side and statistically different from 0.

So, for colon cancer, if covariate "sex" has undetected significant effects, they are more likely positive and they may be of $\beta_w = 0.1642$ giving the hazard rate of $e^{0.1642} \approx 1.18 \in (1.05, 1.32)$. In the same way, for the lung dataset, if the effects of the covariate "sex" treated separately were affected by censoring, whatever they might be, they are likely less than $\beta_{\max} = -0.8816$ corresponding to the hazards ratio of 0.41 from 0.60. Still, if the covariate age treated together with covariate "sex" has undetected significant

TABLE 2: Results from real data application.

c, r	β_w	95% confidence interval	$D_w(\hat{\beta})$	Q_w	P value
<i>For colon cancer dataset with covariate "sex"</i>					
50.48, 0	-0.0336	[-0.1629, 0.0957]	—	—	—
50.48, 7	0.1642	[0.0470, 0.2814]	218.55	0.3327	0.5641
<i>For lung cancer dataset with covariate "sex" alone</i>					
27.63, 0	-0.5311	[-0.859, -0.2034]	—	—	—
27.63, -409	-0.8816	[-1.763, -0.00007]	39.02	0.7177	0.3969
<i>For lung cancer dataset with covariates "sex" and "age"</i>					
27.63, 0	$\begin{bmatrix} -0.5133 \\ 0.0171 \end{bmatrix}$	[-0.8415, -0.1851] [-0.0010, 0.0351]	—	—	—
27.63, 1	$\begin{bmatrix} -0.4932 \\ 0.01728 \end{bmatrix}$	[-0.8042, -0.1822] [0.000018, 0.0346]	$\begin{bmatrix} 29.475 & -45.17 \\ -45.17 & 12037.7 \end{bmatrix}$	0.00314	0.9984

effects due to censoring, they might be just around 0.01728 meaning that the person is at a higher risk with a hazard ratio of $e^{0.01728} \approx 1.02$ but it is still less than 1.04 compared to the one with the same sex to whom they are one year older. In that case, females are at a lower risk than males of the same age with a hazard ratio of $e^{-0.4932} \approx 0.61$.

As we highlighted that r can be positive or negative, the optimal weight obtained in a row (27.63, 1) in Table 2 was also obtained with $r = -409$ and the difference observed was only in the corresponding 95% CIs for the respective components where the 95% CIs at $r = -409$ are $[-0.5086, -0.4778]$ and $[0.0164, 0.0181]$ which are too much narrower and hence better. Note that, this level of $r = -409$ was the one which yielded the maximum effects when covariate "sex" was treated separately, but when treated together with age, the same level yielded the significance of age covariate effects with a minimum deviation from the existing sex covariate effects. In addition, by referring to the magnitudes of the two parameters, we obtain $|\beta_w| = 0.4935$ and $|\beta| = 0.5136$ with a difference of 0.02 which is relatively small.

In general, by looking at the obtained test statistic in the two tables of the results either from the simulation or real data application, we fail to reject the null hypothesis, and since the p values are relatively higher, it categorically discredits the assumed model. The same results were obtained on the three datasets analyzed in our reference work of [5]. So, there is no significant difference between the effects of the covariates under censoring (normally obtained) and those obtained under the use of censoring balancing functions.

5. Conclusion

Censoring generally negatively affects the results of the Cox regression. The significance of treatment effects can be detected if lost due to censoring but their magnitude (or size) must be controlled with due attention for them to remain realistic. For one covariate case, the possible maximum effects can be estimated. If censoring is very high, for example up to 40 % or above, some covariates can be judged as nonsignificant while they really are and in some cases, they may even be in the opposite direction. This means that the

effects were found to be positive (negative) but not significant under censoring while they were negative (positive) and significant. The employed chi-square statistic test helps to know if the newly obtained regression parameter estimate is still consistent by comparing it to the initial one. When the difference is detected, Lin [5] proposed to perform a component-wise comparison of the two parameters. This is why the detection of the probably undetected significant effects is performed while controlling the deviation in the components which is already significant. In this work, the significance of some covariates' effects was detected through the employment of censoring balancing functions without experiencing a significant difference between the parameters' magnitudes.

For low censoring (for example, below 20%), the censoring does not have stronger negative effects that the probable existing significance might be lost unless probably worsened by smaller sample sizes. For this, censoring balancing functions can mislead the researcher by overestimating the coefficient while the harm from censoring is not really that high. The reliance on compensation depends on the harm caused. However, the general remark is that the covariates' effects obtained under the use of censoring balancing functions are also consistent as proved by Lin [5] for the general weighted Cox regression. The use of weight functions defined in the same context as censoring balancing functions (proposed in this work) will help researchers in clinical trials and pharmaceutical studies to maximize the strength of their statistical tests and make sure that no covariates with really significant effects are reported as nonsignificant due to censoring. The proposed censoring balancing functions can, therefore, be recommended for use while aiming at the investigation of significant covariates' effects which were undetected in the presence of censoring.

Data Availability

The real datasets used in this study are available and freely accessible online and especially in R where both the lung cancer data and colon cancer data are accessed through the "survival" package.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B*, vol. 34, no. 2, pp. 187–202, 1972.
- [2] R. S. Lin and L. F. León, "Estimation of treatment effects in weighted log-rank tests," *Contemporary clinical trials communications*, vol. 8, pp. 147–155, 2017.
- [3] C. Yu, X. Huang, H. Nian, and P. He, "A weighted log-rank test and associated effect estimator for cancer trials with delayed treatment effect," *Pharmaceutical Statistics*, vol. 20, no. 3, pp. 528–550, 2021.
- [4] M. Schemper, S. Wakounig, and G. Heinze, "The estimation of average hazard ratios by weighted cox regression," *Statistics in Medicine*, vol. 28, no. 19, pp. 2473–2489, 2009.
- [5] D. Lin, "Goodness-of-fit analysis for the cox regression model based on a class of parameter estimators," *Journal of the American Statistical Association*, vol. 86, no. 415, pp. 725–728, 1991.
- [6] M. Schemper, "Cox analysis of survival data with non-proportional hazard functions," *The Statistician*, vol. 41, no. 4, pp. 455–465, 1992.
- [7] P. Sasieni, "Maximum weighted partial likelihood estimators for the cox model," *Journal of the American Statistical Association*, vol. 88, no. 421, pp. 144–152, 1993.
- [8] L. D. Fisher and D. Y. Lin, "Time-dependent covariates in the cox proportional-hazards regression model," *Annual Review of Public Health*, vol. 20, no. 1, pp. 145–157, 1999.
- [9] D. Dunkler, M. Ploner, M. Schemper, and G. Heinze, "Weighted cox regression using the r package coxphw," *Journal of Statistical Software*, vol. 84, no. 2, pp. 1–26, 2018.
- [10] L. F. León, R. Lin, and K. M. Anderson, "On weighted log-rank combination tests and companion cox model estimators," *Statistics in Biosciences*, vol. 12, no. 2, pp. 225–245, 2020.
- [11] D. Y. Lin and L.-J. Wei, "The robust inference for the cox proportional hazards model," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 1074–1078, 1989.
- [12] S. Murphy, A. Rossini, and A. W. van der Vaart, "Maximum likelihood estimation in the proportional odds model," *Journal of the American Statistical Association*, vol. 92, no. 439, pp. 968–976, 1997.
- [13] I. Persson and H. Khamis, "Bias of the cox model hazard ratio," *Journal of Modern Applied Statistical Methods*, vol. 4, no. 1, pp. 90–99, 2005.
- [14] R. Xu and J. O'Quigley, "Estimating average regression effect under non-proportional hazards," *Biostatistics*, vol. 1, no. 4, pp. 423–439, 2000.
- [15] H. Kvamme and Ø. Borgan, "The brier score under administrative censoring: problems and a solution," *Journal of Machine Learning Research*, vol. 24, no. 2, pp. 1–26, 2023.
- [16] R. A. Matsouaka and F. D. Atem, "Regression with a right-censored predictor using inverse probability weighting methods," *Statistics in Medicine*, vol. 39, no. 27, pp. 4001–4015, 2020.
- [17] G. A. Satten and S. Datta, "The kaplan–meier estimator as an inverse-probability-of-censoring weighted average," *The American Statistician*, vol. 55, no. 3, pp. 207–210, 2001.
- [18] S. Wakounig, G. Heinze, and M. Schemper, "Non-parametric estimation of relative risk in survival and associated tests," *Statistical Methods in Medical Research*, vol. 24, no. 6, pp. 856–870, 2015.
- [19] J. M. Robins and D. M. Finkelstein, "Correcting for non-compliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests," *Biometrics*, vol. 56, no. 3, pp. 779–788, 2000.
- [20] G. Dong, L. Mao, B. Huang et al., "The inverse-probability-of-censoring weighting (ipcw) adjusted win ratio statistic: an unbiased estimator in the presence of independent censoring," *Journal of Biopharmaceutical Statistics*, vol. 30, no. 5, pp. 882–899, 2020.
- [21] C. J. Howe, S. R. Cole, J. S. Chmiel, and A. Munoz, "Limitation of inverse probability-of-censoring weights in estimating survival in the presence of strong selection bias," *American Journal of Epidemiology*, vol. 173, no. 5, pp. 569–577, 2011.