STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **6**, September 2018, pp 305–334.
Published online in International Academic Press (www.IAPress.org)

Proportional Odds under Conway-Maxwell-Poisson Cure Rate Model and Associated Likelihood Inference

N. Balakrishnan *, T. Feng

Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada.

Abstract Cure rate models are useful while modelling lifetime data involving long time survivors. In this work, we discuss a flexible cure rate model by assuming the number of competing causes for the event of interest to follow the Conway-Maxwell Poisson distribution and the lifetimes of the non-cured individuals to follow a proportional odds model. The baseline distribution is considered to be either Weibull or log-logistic distribution. Under right censoring, we develop the maximum likelihood estimators using EM algorithm. Model discrimination among some well-known special cases are discussed under both likelihood- and information-based criteria. An extensive simulation study is carried out to examine the performance of the proposed model and the inferential methods. Finally, a cutaneous melanoma dataset is analyzed for illustrative purpose.

Keywords Cure rate model, Long-term survivor, Conway-Maxwell-Poisson distribution, Weibull distribution, Log-logistic distribution, Maximum likelihood estimator, Expectation-Maximization algorithm, Profile likelihood, Akaike information criterion, Bayesian information criterion, Proportional odds model, Model discrimination

AMS 2010 subject classifications 62H10, 62F15

DOI: 10.19139/soic.v6i3.573

1. Introduction

Statistical models accommodating a surviving fraction are known as cure rate models. The cure rate model was first introduced by [8] and [7], and have been subsequently studied by many authors. Applications of cure rate models are not limited to biomedical studies, and can also be seen in industrial reliability, finance, manufacturing, demography, and criminology. The basic cure rate model can be seen as a two-component mixture model. If $S_p(t)$ is the population survival function, it can be expressed as

$$S_p(t) = p_0 + (1 - p_0)S_s(t), \tag{1}$$

where p_0 is the probability of cure and $S_s(t)$ is the survival function of the non-cured or susceptible individuals in the population. Primarily, cure data have been analyzed in the literature by the structure of the underling survival model of the non-cured individuals $S_s(t)$ as proportional hazards (PH) mixture cure model [20] [15] [24], accelerated failure time (AFT) mixture cure rate model [25] [17] [14], accelerated hazards (AH) mixture cure rate model [26], and proportional odds (PO) mixture cure rate model [10] [18]. More generally, a cure model can be approached through a competing risks set up as follows. Suppose M is an unobservable random variable denoting the number of competing causes related to the occurrence of an event of interest. Let W_j, $j = 1, \ldots, m$, be the random variables denoting the time-to-event for the jth competing cause. Given $M = m$, $W_1, \ldots, W_m$ are assumed to be independent and identically distributed (i.i.d.) with a common cumulative distribution function

*Correspondence to: N. Balakrishnan (Email: bala@mcmaster.ca). Department of Mathematics and Statistics, McMaster University, 1280 Main St W, Hamilton, ON,Canada, L8S 4L8.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

(c.d.f) $F(w) = 1 - S(w)$. Then, the population time-to-event or lifetime is given by

$$Y = \min\{W_0, W_1, \ldots, W_m\}, \tag{2}$$

where W_0 is corresponding to the individual who are not susceptible to the event occurrence (namely, with infinite lifetime). This leads to a proportion of the cured group, known as cure rate. The survival function for the entire population is then [22]

$$S_p(y) = \sum_{m=0}^{\infty} P(M = m)[S(y)]^m = A_M(S(y)), \tag{3}$$

where $A_M(.)$ is the probability generating function (p.g.f.) of M.

In the present work, we assume a proportional odds model for the distribution of W_j, with a parametric assumption on the baseline odds function. To be more specific, the odds function of W_j is taken as

$$O(w; \mathbf{x}) = \theta O_0(w), \tag{4}$$

where $O(w) = S(w)/F(w)$ is the odds of survival up to w, the term θ is linked to covariates as $e^{\mathbf{x}'\boldsymbol{\gamma}_2}$ with $\mathbf{x} = (x_1, \ldots, x_p)'$ being a vector of p covariates, $\boldsymbol{\gamma}_2 = (\gamma_{21}, \ldots, \gamma_{2p})'$ is the proportional odds regression coefficients, and $O_0(w)$ is the baseline odds function. We can further obtain the survival function of W_j as

$$S(w) = [1 + e^{-\mathbf{x}'\boldsymbol{\gamma}_2}(S_0(w)^{-1} - 1)]^{-1}, \tag{5}$$

with the corresponding probability density function (p.d.f.)

$$f(w) = f_0(w)e^{-\mathbf{x}'\boldsymbol{\gamma}_2}[(1 - S_0(w))e^{-\mathbf{x}'\boldsymbol{\gamma}_2} + S_0(w)]^{-2}. \tag{6}$$

The rest of this paper proceeds as follows. In Section 2, we present briefly the COM-Poisson cure model and its three special cases. Section 3 describes the data and the likelihood, while the estimation of the cure rate and associated inferential issues are discussed in Section 4. In Section 5, an extensive simulation study is carried out. In Section 6, we discuss model discrimination using information- and likelihood-based methods. A data on cutaneous melanoma is analyzed in Section 7 for illustrative purpose. Some concluding comments are finally made in Section 8.

2. The COM-Poisson cure rate model

Suppose the number of the competing causes M follows a COM-Poisson distribution [9]. The probability mass function (p.m.f.) of M is given by

$$P(M = m; \eta, \phi) = \frac{1}{Z(\eta, \phi)} \frac{\eta^m}{(m!)^\phi}, \; m = 0, 1, 2, \ldots, \tag{7}$$

where the normalization constant is given by

$$Z(\eta, \phi) = \sum_{j=0}^{\infty} \frac{\eta^j}{(j!)^\phi}, \tag{8}$$

with $\phi \geq 0$ and $\eta > 0$. Cure rate is the probability

$$p_0 = P(M = 0; \eta, \phi) = (Z(\eta, \phi))^{-1}. \tag{9}$$

As a weighted Poisson random varaiable (r.v.), M leads to a Poisson r.v. with mean equal to η when $\phi = 1$, and M leads to a under- or over-dispersion if $\phi > 1$ or $\phi < 1$ (see [23] and [12].) For example, M approaches the

Bernoulli r.v. with parameter $\frac{1}{1+\eta}$ when $\phi \to \infty$ and $Z(\eta,\phi) \to 1+\eta$, and M reduces to a Geometric r.v. with parameter $1-\eta$ if $\phi = 0, \eta < 1$ and $Z(\eta,\phi) = \frac{1}{1-\eta}$. Note that M is undefined for $\eta \geq 1$ and $\phi = 0$. The population survival function and density function of the time-to-event Y is then

$$S_p(y) = \frac{Z(\eta S(y),\phi)}{Z(\eta,\phi)}, \tag{10}$$

$$f_p(y) = \frac{1}{Z(\eta,\phi)} \frac{f(y)}{S(y)} \sum_{j=1}^{\infty} \frac{j(\eta S(y))^j}{(j!)^\phi}. \tag{11}$$

Note that as $y \to \infty$, $S_p(y) \to p_0 > 0$. Hence, $S_p(y)$ is not a proper survival function. Suppose we have a indicator variable of I such that $I = 0$ if the subject is immune (belongs to set I_0) with probability p_0 and $I = 1$ if the subject is susceptible (belongs to set I_1) with probability $1-p_0$. The cumulative distribution and survival function of the overall population can be viewed as a mixture of two populations,

$$F_p(y) = P[Y \leq y | I = 0]P(I = 0) + P[Y \leq y | I = 1]P(I = 1) = F_s(y)(1-p_0) \tag{12}$$

For a detailed discussion, interested readers may refer to [2], [3], [4], [6], [1], and [21].

3. Data and the likelihood

Suppose the time-to-event is not completely observed and is subject to non-informative right censoring, which means that the data above a certain value are not observed. Therefore, the observation time T_i, for the ith subject, would be the minimum of the censoring time C_i and actual lifetime Y_i, i.e.,

$$T_i = \min\{Y_i, C_i\}, i = 1, \ldots, n. \tag{13}$$

We define an indicator function $\delta_i = I(Y_i \leq C_i)$ for the i-th subject such that $\delta_i = 0$ if the lifetime is observed while $\delta_i = 1$ if the lifetime is right censored, Δ_0 and Δ_1 are sets with all the i's equal to 0 and 1, respectively, and set Δ^* contains all the i's. It is to be noted that the cure rate $p_0 = Z(\eta,\phi)^{-1}$ is purely a function of η for a fixed value of ϕ. The range of $1/p_0$ is from 1 to infinity and it is monotone in η. Therefore, it is natural to use a logistic regression model $H_\phi(\eta) = 1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$ to link the covariate $x_1, \ldots, x_p$ to the cured proportion p_{0i}, i.e.,

$$p_{0i} = p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) = Z(\eta,\phi)^{-1} = H_\phi(\eta)^{-1} = (1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}, \tag{14}$$

where p_{0i} is the cured proportion for the ith category, $\boldsymbol{x}_i = (1, \boldsymbol{x}_{ic}')' = (1, x_{i1}, \ldots, x_{ip})'$ is a vector of $p+1$ covariates, and $\boldsymbol{\beta}$ is the vector of regression coefficients. Under this link, η would equal $H^{-1}(1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$, i.e., η can be calculated from the inverse function of $H_\phi(.)$ analytically for the Geometric, Poisson and Bernoulli distributions or by using numerical method for the general COM-Poisson distribution.

For *n pairs of* observations $(\boldsymbol{t}, \boldsymbol{\delta})$=$\{(t_1, \delta_1), \ldots, (t_n, \delta_n)\}$, the observed data likelihood function under the non-informative censoring is then given by

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} \{f_p(t_i; \boldsymbol{\theta})\}^{\delta_i} \{S_p(t_i; \boldsymbol{\theta})\}^{1-\delta_i}, \tag{15}$$

where $\boldsymbol{\theta}$ is the set of parameters $(\phi, \boldsymbol{\beta}', \boldsymbol{\gamma}')$, which is equivalent to

$$L(\boldsymbol{\theta}; \boldsymbol{t}, \boldsymbol{\delta}) \propto \prod_{i \in \Delta_1} f_p(t_i; \boldsymbol{\theta}) \prod_{i \in \Delta_0} \{p_0 + (1-p_0) S_s(t_i; \boldsymbol{\theta})\}. \tag{16}$$

Here, we consider two baseline distributions for the proportional odds survival model corresponding to the time-to-event random variable, namely, Weibull and log-logistic distributions. It should also be noted that log-logistic

distribution in fact processes the proportional odds property, while the Weibull distribution does not. The survival function and p.d.f. of W under a Weibull baseline, for example, are

$$S(w,\boldsymbol{\gamma}) = [1+e^{-\mathbf{x}'\boldsymbol{\gamma}_2}(e^{(\gamma_1 w)^{1/\gamma_0}}-1)]^{-1}, w>0, \tag{17}$$

$$f(w,\boldsymbol{\gamma}) = (\gamma_1 w)^{1/\gamma_0} e^{\mathbf{x}'\boldsymbol{\gamma}_2-(\gamma_1 w)^{1/\gamma_0}}[e^{-(\gamma_1 w)^{1/\gamma_0}}(e^{\mathbf{x}'\boldsymbol{\gamma}_2}-1)+1]^{-2}/(\gamma_0 w), w>0, \tag{18}$$

where $\gamma_0>0$ and $\gamma_1>0$ are the shape and scale parameters of the baseline Weibull distribution, respectively. If we assume the baseline distribution to be a log-logistic distribution with $\gamma_0>0$ and $\gamma_1>0$ as the scale and shape parameters, respectively, then the corresponding odds function of W_i is given by

$$O(w,;\boldsymbol{x}_c',\boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1}}{w^{\gamma_1}} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2} = O_0(w,;\gamma_0,\gamma_1)e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}. \tag{19}$$

We observe that W_i still follows a two-parameter log-logistic distribution ($\gamma_0,\gamma_1>0$) with shape parameter γ_1 and scale parameter $\gamma_0 e^{-\boldsymbol{x}_c',\boldsymbol{\gamma}_2/\gamma_1}$, and with corresponding survival function

$$S(w,\boldsymbol{\gamma}) = \frac{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}}{\gamma_0^{\gamma_1} e^{\boldsymbol{x}_c'\boldsymbol{\gamma}_2}+w^{\gamma_1}}, w>0. \tag{20}$$

Note that the mean does not exist if $\gamma_1<1$ and the variance does not exist if $\gamma_1<2$.

4. Estimation of parameters

We propose an Expectation-Maximization (EM) algorithm for obtaining the MLE of $\boldsymbol{\theta}$, and a profile likelihood approach for the estimation of the dispersion patameter ϕ. It is well-known that EM is an effective technique for finding the MLEs of unknown parameters of a model involving unobserved variables (for further discussion, refer to [19]). In our model, the random variable I_i's are observed for i in the set Δ_1 and unobserved for i in the set Δ_0, where $I_i=1$ if the individual is susceptible and $I_i=0$ if the individual is cured. Let us denote the set of complete data by $(\boldsymbol{t},\boldsymbol{\delta},\boldsymbol{x},\boldsymbol{I})$=$\{(t_1,\delta_1,\boldsymbol{x}_1,I_1),\ldots,(t_n,\delta_n,\boldsymbol{x}_n,I_n)\}$. The complete data likelihood function is then

$$L_c(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{x},\boldsymbol{\delta},\boldsymbol{I}) \propto \prod_{i\in\Delta_1} f_p(t_i,\boldsymbol{x}_i,\boldsymbol{\theta}) \prod_{i\in\Delta_0} p_0(\boldsymbol{\beta},\boldsymbol{x}_i)^{1-I_i}[(1-p_0(\boldsymbol{\beta},\boldsymbol{x}_i))S_s(t_i,\boldsymbol{x}_{ic};\boldsymbol{\theta})]^{I_i}, \tag{21}$$

where $\mathbf{I}=(I_1,\ldots,I_n)'$, $\mathbf{x}_{ic}=(x_{i1},\ldots,x_{ip})'$ and $\mathbf{x}_i=(1,\mathbf{x}_{ic}')'$. The corresponding complete log-likelihood function

$$\begin{aligned} l_c(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{x},\boldsymbol{\delta},\boldsymbol{I}) = \text{constant} &+ \sum_{i\in\Delta_1}\log f_p(t_i,\boldsymbol{x}_i,\boldsymbol{\theta}) + \sum_{i\in\Delta_0}(1-I_i)\log p_0(\boldsymbol{\beta},\boldsymbol{x}_i) \\ &+ \sum_{i\in\Delta_0} I_i\log[1-p_0(\boldsymbol{\beta},\boldsymbol{x}_i)] + \sum_{i\in\Delta_0} I_i\log S_s(t_i,\boldsymbol{x}_{ic};\boldsymbol{\theta}). \end{aligned} \tag{22}$$

4.1. E-step

The expectation step is achieved by calculating the expected value of the complete data log-likelihood function with respect to the conditional distribution of the unobserved I_i's ($i\in\Delta_0$), given the observed data $\boldsymbol{O}=\{(t_i,\delta_i,\boldsymbol{x}_i), i\in\Delta_1\}$ and the current estimates of the parameters $\boldsymbol{\theta}^{(k)}=(\boldsymbol{\beta}',\boldsymbol{\gamma}')'$ for a fixed value of ϕ. Let us denote the function as

$$Q(\boldsymbol{\theta}^*,\boldsymbol{\pi}^{(k)}) = E(l_c(\boldsymbol{\theta};\boldsymbol{t},\boldsymbol{x},\boldsymbol{\delta},\boldsymbol{I})|\boldsymbol{O},\boldsymbol{\theta}^{(k)}), \tag{23}$$

at the k-th iteration step. In our model, I_i's are Bernoulli random variables and we can easily find the conditional expectation if the ith individual being susceptible is

$$\boldsymbol{\pi}_i^{(k)} = \mathrm{E}(I_i|\boldsymbol{O},\boldsymbol{\theta}^{(k)}) = P(I_i=1|T>t) = \frac{(1-p_0(\boldsymbol{\beta}^{(k)},\boldsymbol{x}_i))S_s(t_i,\boldsymbol{x}_{ic};\boldsymbol{\theta}^{(k)})}{S_p(t_i,\boldsymbol{x}_i;\boldsymbol{\theta}^{(k)})}|_{\theta=\theta^{(k)}}. \tag{24}$$

Now, for a fixed value of ϕ, the Q function is given by

$$Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)}) = \sum_{I\in\Delta_1} \log f_p(t_i, \boldsymbol{x}_i, \boldsymbol{\theta}) + \sum_{i\in\Delta_0} (1-\boldsymbol{\pi}_i^{(k)}) \log p_0(\boldsymbol{\beta}, \boldsymbol{x}_i) + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \log[1-p_0(\boldsymbol{\beta}, \boldsymbol{x}_i)] + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \log S_s(t_i, \boldsymbol{x}_{ic}; \boldsymbol{\theta}) \tag{25}$$

which can be further simplified as

$$Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)}) = -\sum_{i\in\Delta*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_1} \log f(t_i, \boldsymbol{\gamma}) - \sum_{i\in\Delta_1} \log S(t_i, \boldsymbol{\gamma}) + \sum_{i\in\Delta_1} \log z_{2,i} + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \log z_{1,i}, \tag{26}$$

where

$$z_{1,i} = z_1(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{\{\eta_i S(t_i; \boldsymbol{\gamma})\}^j}{(j!)^\phi}, \tag{27}$$

$$z_{2,i} = z_2(\boldsymbol{\theta}; \boldsymbol{x}_i, t_i) = \sum_{j=1}^{\infty} \frac{j\{\eta_i S(t_i; \boldsymbol{\gamma})\}^j}{(j!)^\phi}. \tag{28}$$

4.2. M-step

The M-step is achieved by maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)})$ function in (26) in order to obtain the improved estimate of $\boldsymbol{\theta}$, i.e.,

$$\boldsymbol{\theta}^{*(k+1)} = \arg\max Q(\boldsymbol{\theta}, \boldsymbol{\pi}^{(k)}). \tag{29}$$

The MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not have explicit expressions. In this paper, the numerical maximization is carried out by Newton-Raphson method.

For a fixed value of ϕ, the E-step and M-step are alternated until the parameter estimate converges to a desired level of accuracy. The parameter ϕ is determined by using the profile likelihood technique. We consider a range of ϕ with small increment, and then for each value of ϕ, the MLEs of other parameters are found, and the estimates with the largest likelihood is chosen as the final estimate. The following subsections present explicit forms of the first- and second-order derivatives of the Q function as well as update function for the case of COM-Poisson distribution.

4.3. Results for the COM-Poisson cure rate model

The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for fixed values of ϕ, are as follows:

$$\frac{\partial Q}{\partial \beta_l} = -\sum_{i\in\Delta*} x_{il}\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}}{z_{2,i}z_{01,i}}x_{il} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{2,i}}{z_{1,i}z_{01,i}}x_{il},$$

$$\frac{\partial Q}{\partial \gamma_h} = \sum_{i\in\Delta_1}\frac{\partial \log f(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{2,i}}{z_{1,i}}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h},$$

$$\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} = -\sum_{i\in\Delta*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2} + \sum_{i\in\Delta_1}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(z_{2,i}z_{01,i})^2}[z_{2,i}(z_{21,i}z_{01,i}+z_{31,i}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) - z_{21,i}[z_{21,i}+\frac{z_{2,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(z_{1,i}z_{01,i})^2}[z_{1,i}(z_{21,i}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}+z_{2,i}z_{01,i}) - z_{2,i}[z_{2,i}+\frac{z_{1,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}_i'\boldsymbol{\beta}}],$$

$$\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_h} = \sum_{i\in\Delta_1}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2z_{01,i}}\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}z_{1,i}-z_{2,i}^2}{z_{1,i}^2z_{01,i}}\frac{\partial \log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h},$$

$$\frac{\partial^2 Q}{\partial\gamma_h\partial\gamma_{h'}} = \sum_{i\in\Delta_1}\frac{\partial \log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} + \sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} + \sum_{i\in\Delta_1}(\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2})\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{2,i}}{z_{1,i}}\frac{\partial \log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{z_{21,i}z_{1,i}-z_{2,i}^2}{z_{1,i}^2}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}},$$

for $l, l' = 0,\ldots,p$, $x_{i0}\equiv 1$, $h, h' = 0, 1, j^*$, $j* = 21, 22, \ldots, 2p$, $i = 1,\ldots,n$, where

$$z_{21,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^2\{\eta_i S(t_i;\boldsymbol{\gamma})\}^j}{(j!)^\phi}, \quad z_{31,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^3\{\eta_i S(t_i;\boldsymbol{\gamma})\}^j}{(j!)^\phi},$$
$$z_{01,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j\eta_i^j}{(j!)^\phi}, \quad z_{02,i} = z_2(\boldsymbol{\theta};\boldsymbol{x}_i,t_i) = \sum_{j=1}^{\infty}\frac{j^2\eta_i^j}{(j!)^\phi},$$

See $z_{1,i}$ and $z_{2,i}$ in (27) and (28).

4.4. Standard errors and asymptotic confidence intervals

We may approximate the asymptotic variance-covariance matrix of the MLEs $(\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}')'$ by inverting the observed Fisher information matrix of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, for a fixed value of ϕ. The components of the observed Fisher information matrix can be calculated from the negative of the second-order derivatives of the complete data likelihood function with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ (for detailed information, refer to [16]). Thus, we can obtain the standard errors of the estimates and then construct corresponding asymptotic confidence intervals for the parameters.

4.5. Estimation of the cure rate and its standard error

Suppose $\hat{\boldsymbol{\beta}}$ is the MLE of the regression coefficient $\boldsymbol{\beta}$. The estimated cure rate for the corresponding group i is then $\hat{p}_{0i} = (1+e^{\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}})^{-1}$, for $i = 1, ..., \tau$. The standard error of $\hat{p}_{0i}$ can be found through delta method as

$$\text{sd}(\hat{p_{0i}}) = \sqrt{(\frac{\partial \hat{p_{0i}}}{\partial\hat{\beta}_0 ... \partial\hat{\beta}_p})\widehat{\text{var}\boldsymbol{\beta}}(\frac{\partial \hat{p_{0i}}}{\partial\hat{\beta}_0 ... \partial\hat{\beta}_p})'}. \tag{30}$$

4.6. Equivalent models

Proposition: The population survival function under Geometric and Bernoulli cure rate model are equivalent through re-parametrization if baseline odds follow log-logistic distribution.

Proof: Cure rates p_0 under Geometric and Bernoulli cure rate models are $1-\eta_1$, and $\frac{1}{1+\eta_3}$, respectively. If we equate these cure rates, we obtain a relationship between η_1 and η_3 as

$$\eta_3 = \frac{\eta_1}{1-\eta_1}. \tag{31}$$

Suppose $S_1 = \frac{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}}}{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} + t^{\gamma_1}}$ and $S_3 = \frac{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}}}{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} + t^{\gamma_1}}$ are the survival functions of the susceptible group under Geometric and Bernoulli cure rate models, respectively. Then, the population survival function for the under the Bernoulli case is

$$S_p = \frac{1+\eta_3 S_3}{1+\eta_3}, \tag{32}$$

If we fix the relationship between γ_{30}, γ_{10}, γ_{32} and γ_{12} as

$$\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} = \gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} p_0 = \gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}} (1-\eta_1), \tag{33}$$

then we obtain from (32) that

$$S_p = 1-\eta_1+\eta_1 S_3 = 1-\eta_1 \frac{t^{\gamma_1}}{\gamma_{30}^{\gamma_1} e^{\boldsymbol{x}\gamma_{32}} + t^{\gamma_1}} = 1 - \frac{t^{\gamma_1}\eta_1}{\gamma_{10}^{\gamma_1} e^{\boldsymbol{x}\gamma_{12}}(1-\eta_1) + t^{\gamma_1}} = \frac{1-\eta_1}{1-\eta_1 S_1}, \tag{34}$$

which is the population survival function for the Geometric cure rate model under proportional odds assumption. Thus, these two models in this case are re-parametrizations of each other.

5. Simulation study

An extensive Monte Carlo simulation study is carried out in this paper to evaluate the performance of the proposed methodology by varying the sample size, cure rate, censoring proportion, lifetime parameter, and underling lifetime distribution. We try to mimic the cutaneous melanoma data, and considered 4 possible categories for the individuals, namely, $x = 0, 1, 2, 3$. Three different sample sizes are considered in the study: $n = 200$ $(50, 42, 53, 55)$, $n = 400$ $(95, 102, 97, 106)$, $n = 800$ $(200, 168, 212, 220)$ to reflect small, medium and large sample sizes. Moreover, if we assume that $\boldsymbol{\beta} = (\beta_0, \beta_1)$ has two parameters, fixing the cure rates for the first and fourth categories would be enough to cover all cases as the cure rates for the second and third categories can then be obtained from $\boldsymbol{\beta}$. Here, we take $(p_{00}, p_{03}) = (0.4, 0.2)$ and $(p_{00}, p_{03}) = (0.6, 0.25)$ with respect to categories one and four for low and high cure rates, respectively. Also, the cure rate would be in a decreasing order in this way. The $\boldsymbol{\beta}$'s are

$$\beta_0 = \ln(1/p_{00} - 1) \ , \ \beta_1 = (\ln(1/p_{03} - 1) - \beta_0)/3. \tag{35}$$

We thus obtain the true value of $\boldsymbol{\beta}$ as (0.405,0.321) and (-0.405,0.501), respectively. In addition, we consider light and heavy censored data in the simulation. The light and heavy censoring rates are $(0.52, 0.45, 0.37, 0.3)$ with $(0.65, 0.49, 0.4, 0.35)$ as the corresponding cure rates, $(0.7, 0.57, 0.45, 0.34)$ and $(0.8, 0.64, 0.5, 0.38)$ for the corresponding cure rates. It is natural to assume that the probability of censored population for the susceptible group equal to the difference between the probability of getting censored and cured; i.e.,

$$P(Y \geq C_x \cap M \geq 1 | X = x) = c_x - p_{0x}. \tag{36}$$

If we assume the censoring time C_x follows an exponential distribution with rate λ_x on $x = 0, 1, 2, 3$, Eq. (36) can be re-written as

$$\lambda_x \int_0^\infty S_p(C_x) e^{-\lambda_x c_x} dc_x = c_x. \tag{37}$$

The choice of (γ_0, γ_1) in the underlying distribution of the proportional odds survival model are $(0.571, 0.307)$ and $(1.75, 3.25)$ for Weibull and log-logistic distributions, respectively. The odds parameter is specified by $\gamma_2 = -0.75$ to ensure a decreasing lifetime for the four nodule categories. We consider an inverse transform sampling method to simulate the actual survival lifetime Y_i for each individual under different competing risks, i.e., $w_i = \frac{1}{\gamma_1}[\log(1 + (\frac{1}{u} - 1)e^{x_i\gamma_2})]^{\gamma_0}$ and $w_i = \gamma_0(\frac{u}{1-u}e^{x_i\gamma_2})^{1/\gamma_1}, i = 1, \ldots, n$, under the proportional odds model with Weibull and log-logistic baseline distribution, respectively, where u follows an uniform distribution with range 0 to 1.

Under the above setting, the procedure to generate the data form different cure rate models is as follows. **Geometric cure rate model**: For each individual, we simulate the number of competing risk M_i from Geometric distributions with probability $P(M_i = 0) = p_{0x}$; and we simulate the censoring time C_x from exponential distribution with rate λ_x, the parameter λ_x can be found from (37). If M_i does not equal zero, we simulate M_i number of actual lifetimes $\{Y_i, .., Y_{M_i}\}$ from proportional odds survival model, and the observed lifetime T_i is taken as the minimum of all the actual lifetime and the censoring time, i.e., $\min\{Y_i, .., Y_{M_i}, C_i\}$. If $Y_i = C_i$, we make the censoring indicator $\delta_i = 1$, otherwise $\delta_i = 0$. On the other hand, $M_i = 0$ means the individual is censored, we assign C_i to the actual lifetime T_i, and the censoring indicator is taken to be $\delta_i = 0$. **Poisson cure rate model**: In this case, the procedure is the same as the Geometric cure rate model except that M_i is simulated from Poisson distribution with parameter $-log(p_{0x})$. **Bernoulli cure rate model**: There are two ways to do the data generation in this case. One is the same as Geometric cure rate model except that M_i is simulated from Bernoulli distribution with probability of success as $1 - p_{0x}$. Another way is a little bit simpler since M_i can only be taken as 0 or 1 in this case. For each individual, we simulate the censoring time C_x from exponential distribution with rate λ_x. Then, we simulate an uniform random variable U_i and if $U_i \leq p_{0x}$, the observed lifetime T_i is set to C_x; otherwise, we generate the observed lifetime T_i from the proportional odds survival model.

In our simulation study, 1000 Monte Carlo runs were considered in each scenario. The estimates were calculated through EM method. We stopped our estimation if the absolute difference between two consecutive estimates was less than 10^{-5}. We calculated the empirical Bias, standard errors(SE), root Mean Square Error (RMSE), and 95% coverage probabilities (CPs) for the estimates of the parameters. In addition, we computed the cure rate, SE and 95% CPs. Here, the initial values of the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ were taken from a grid of parameters with a range from 80% to 120% of the true value, and those estimates having the maximum likelihood were chosen as the initial value.

Tables 1-6 present the bias, SE, RMSE, and coverage probabilities for the three special cases. We can see that the estimates are quite accurate under different cure rate models. The Bias, standard error along with RMSE get reduced as the sample size increases. The same follows when the censoring is light or the cure rate is high. The standard errors and RMSE of β_0 are always larger than other parameters. The coverage probabilities of the confidence intervals based on the asymptotic normality of the MLEs are quite close to the nominal level in most of the cases. To summarize, a larger sample size, smaller censoring proportion, and lower cure rate would result in more accurate estimates.

6. Model discrimination

The COM-Poisson distribution contains many commonly used discrete distributions under different selection of ϕ. It would be of interest to select the suitable ϕ and make full use of the COM-Poisson distribution to get the best fit for the data. So, we focus here on a model discrimination among the three special cases of the COM-Poisson distribution.

We simulated 1000 random samples from the following five choice of ϕ from COM-Poisson distributions: $\phi = 0$ (Geometric), $\phi = 0.5$, $\phi = 1$ (Poisson), $\phi = 2$, $\phi \to \infty$ (Bernoulli). Two different sample sizes were considered: $n = 400$ (95, 102, 97, 106) and $n = 800$ (200, 168, 212, 220) for medium and large sample sizes. The light and heavy censoring rates considered were (0.52, 0.45, 0.37, 0.3) and (0.65,0.49,0.4,0.35) with cure rates (0.4,0.2) if the lifetime follows proportional odds model under Weibull baseline with parameter $\gamma_0 = 0.571$, $\gamma_1 = 0.307$, $\gamma_2 = -0.75$. The light and heavy censoring rates were taken as (0.7, 0.57,0.45,0.34) and (0.8, 0.64, 0.5, 0.38) with

Table 1. Bias, SE, RMSE, and CP for the estimates of the parameters of the Geometric cure rate model under proportional odds with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.003	0.208	0.212	94.7	-0.405	-0.005	0.205	0.208	94.5
β_1	0.327	0.003	0.116	0.121	94.5	0.501	0.007	0.111	0.111	94
γ_0	1.75	0.01	0.147	0.151	93.4	1.75	-0.001	0.154	0.155	93.8
γ_1	3.25	0.021	0.186	0.195	93.6	3.25	0.033	0.205	0.202	95.6
γ_2	-0.75	-0.003	0.149	0.153	94.7	-0.75	-0.001	0.156	0.154	95.2
n=400, HC										
β_0	0.405	0.012	0.25	0.248	96.1	-0.405	-0.002	0.257	0.265	95.1
β_1	0.327	0.002	0.134	0.133	95.6	0.501	0.004	0.131	0.133	95.3
γ_0	1.75	0.019	0.18	0.182	95.3	1.75	0.014	0.198	0.197	95.3
γ_1	3.25	0.026	0.199	0.195	96	3.25	0.045	0.225	0.236	94.6
γ_2	-0.75	-0.011	0.17	0.165	95.8	-0.75	-0.015	0.18	0.178	95.3
n=800, LC										
β_0	0.405	0.004	0.146	0.147	94.3	-0.405	-0.002	0.145	0.146	94.4
β_1	0.327	0.001	0.08	0.078	95.2	0.501	0.001	0.077	0.076	95.7
γ_0	1.75	0.01	0.104	0.106	95.1	1.75	0	0.11	0.111	95.4
γ_1	3.25	0.005	0.13	0.132	95	3.25	0.019	0.144	0.144	94.9
γ_2	-0.75	-0.006	0.103	0.102	95.6	-0.75	-0.003	0.109	0.111	93.8
n=800, HC										
β_0	0.405	0.007	0.177	0.175	95.6	-0.405	-0.003	0.183	0.184	95.7
β_1	0.327	-0.002	0.092	0.091	95.3	0.501	0.003	0.091	0.093	95.5
γ_0	1.75	0.012	0.127	0.126	95.5	1.75	0	0.14	0.144	94.4
γ_1	3.25	0.011	0.14	0.139	94.7	3.25	0.017	0.157	0.155	95.1
γ_2	-0.75	-0.007	0.118	0.115	95.5	-0.75	-0.001	0.126	0.128	95.2

Table 2. Bias, SE, RMSE, and CP for the estimates of the parameters of the Poisson cure rate model under proportional odds model with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.004	0.207	0.202	95.8	-0.405	-0.007	0.205	0.204	95.2
β_1	0.327	0.003	0.115	0.113	96.1	0.501	0.004	0.111	0.108	95.4
γ_0	1.75	0.001	0.12	0.115	95	1.75	-0.005	0.14	0.141	94.8
γ_1	3.25	0.031	0.189	0.195	95.5	3.25	0.039	0.208	0.208	96
γ_2	-0.75	-0.004	0.12	0.119	95.1	-0.75	-0.003	0.134	0.137	94.9
n=400, HC										
β_0	0.405	0.009	0.247	0.248	94	-0.405	-0.004	0.256	0.258	95.7
β_1	0.327	0	0.131	0.133	93.6	0.501	0.007	0.13	0.129	95.5
γ_0	1.75	0.004	0.141	0.136	95.7	1.75	0.006	0.173	0.168	95.2
γ_1	3.25	0.048	0.203	0.213	94.4	3.25	0.041	0.227	0.24	94.2
γ_2	-0.75	-0.013	0.133	0.132	95.3	-0.75	-0.009	0.152	0.159	94
n=800, LC										
β_0	0.405	0.001	0.146	0.141	95.2	-0.405	-0.006	0.145	0.15	94.2
β_1	0.327	0	0.079	0.077	95.6	0.501	0.001	0.077	0.078	95.3
γ_0	1.75	0	0.086	0.084	96	1.75	0.002	0.101	0.101	95.3
γ_1	3.25	0.014	0.132	0.131	94.8	3.25	0.007	0.145	0.138	95.8
γ_2	-0.75	-0.001	0.084	0.082	95.4	-0.75	-0.001	0.094	0.094	94.8
n=800, HC										
β_0	0.405	0.008	0.175	0.176	94.7	-0.405	-0.007	0.182	0.182	95.3
β_1	0.327	-0.002	0.091	0.094	94	0.501	0.003	0.09	0.09	95.1
γ_0	1.75	0.006	0.102	0.1	95.3	1.75	0.001	0.124	0.125	95
γ_1	3.25	0.018	0.142	0.148	93.7	3.25	0.019	0.159	0.163	94.7
γ_2	-0.75	-0.008	0.093	0.096	94.6	-0.75	-0.001	0.107	0.111	93.7

Table 3. Bias, SE, RMSE, and CP for the estimates of the parameters of the Bernoulli cure rate model under proportional odds model with log-logistic baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.002	0.208	0.214	94.6	-0.405	-0.02	0.205	0.207	94.8
β_1	0.327	0.002	0.116	0.117	95.7	0.501	0.011	0.111	0.114	94.2
γ_0	1.75	0.003	0.112	0.114	94.9	1.75	0.002	0.136	0.137	94.1
γ_1	3.25	0.025	0.186	0.193	95.3	3.25	0.032	0.206	0.204	95.2
γ_2	-0.75	-0.004	0.11	0.111	94.6	-0.75	-0.006	0.126	0.127	94.9
n=400, HC										
β_0	0.405	-0.009	0.25	0.257	94	-0.405	0.004	0.256	0.244	96.3
β_1	0.327	0.009	0.134	0.138	94.9	0.501	-0.002	0.13	0.127	96.2
γ_0	1.75	0.011	0.128	0.126	95.6	1.75	0.01	0.162	0.164	94.4
γ_1	3.25	0.028	0.2	0.212	94	3.25	0.037	0.225	0.236	94.2
γ_2	-0.75	-0.012	0.12	0.114	96.4	-0.75	-0.013	0.14	0.146	93.8
n=800, LC										
β_0	0.405	0	0.146	0.146	94.5	-0.405	-0.005	0.145	0.146	95.1
β_1	0.327	0.002	0.08	0.079	95.5	0.501	0.005	0.077	0.078	94.6
γ_0	1.75	0.001	0.079	0.076	95.6	1.75	0.001	0.097	0.097	94.9
γ_1	3.25	0.015	0.131	0.135	94.6	3.25	0.018	0.144	0.148	94.7
γ_2	-0.75	-0.002	0.077	0.076	94.8	-0.75	-0.002	0.088	0.087	95.2
n=800, HC										
β_0	0.405	-0.005	0.177	0.181	94	-0.405	-0.001	0.182	0.17	96.6
β_1	0.327	0.004	0.092	0.092	95.4	0.501	0	0.09	0.085	96.7
γ_0	1.75	0.006	0.091	0.089	94.8	1.75	0.003	0.116	0.117	94.9
γ_1	3.25	0.011	0.14	0.136	95.9	3.25	0.016	0.157	0.156	94.7
γ_2	-0.75	-0.005	0.084	0.083	95.3	-0.75	-0.004	0.098	0.1	94.1

Table 4. Bias, SE, RMSE, and CP for the estimates of the parameters of the Geometric cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	-0.002	0.208	0.209	94.589	-0.405	-0.011	0.197	0.199	94.6
β_1	0.327	0.004	0.117	0.119	94.6	0.501	0.002	0.107	0.108	95.6
γ_0	0.571	-0.001	0.033	0.034	94.399	0.571	0.011	0.037	0.044	95.76
γ_1	0.307	0.005	0.037	0.035	95.884	0.307	0.007	0.037	0.04	94.333
γ_2	-0.75	-0.003	0.153	0.155	93.265	-0.75	0.004	0.157	0.15	94.425
n=400, HC										
β_0	0.405	0.007	0.274	0.282	94.874	-0.405	-0.028	0.26	0.279	94.668
β_1	0.327	0.001	0.143	0.144	95.3	0.501	0.005	0.129	0.133	94.684
γ_0	0.571	0	0.036	0.037	95.19	0.571	0.012	0.046	0.046	93.89
γ_1	0.307	0.005	0.052	0.054	94.874	0.307	0.015	0.057	0.058	93.443
γ_2	-0.75	-0.002	0.183	0.184	95.33	-0.75	0.005	0.189	0.181	93.552
n=800, LC										
β_0	0.405	-0.004	0.146	0.148	94.45	-0.405	-0.015	0.139	0.144	93.3
β_1	0.327	0.002	0.081	0.081	94.289	0.501	0.005	0.075	0.074	95.2
γ_0	0.571	0.001	0.023	0.025	92.893	0.571	0.009	0.026	0.035	93.318
γ_1	0.307	0.002	0.026	0.026	94.864	0.307	0.007	0.026	0.029	93.251
γ_2	-0.75	0	0.106	0.108	94.914	-0.75	0.001	0.109	0.106	94.49
n=800, HC										
β_0	0.405	0.002	0.191	0.191	94.271	-0.405	-0.014	0.183	0.196	93.493
β_1	0.327	0.002	0.098	0.096	95.395	0.501	0.001	0.09	0.091	95.495
γ_0	0.571	0.002	0.026	0.027	93.594	0.571	0.01	0.031	0.037	88.934
γ_1	0.307	0.002	0.037	0.037	94.456	0.307	0.009	0.039	0.041	91.247
γ_2	-0.75	0	0.127	0.125	94.586	-0.75	-0.005	0.131	0.125	94.153

Table 5. Bias, SE, RMSE, and CP for the estimates of the parameters of the Poisson cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.005	0.21	0.207	96	-0.405	0.001	0.207	0.212	94.3
β_1	0.327	0.002	0.118	0.115	95.7	0.501	0.006	0.113	0.118	94.4
γ_0	0.571	-0.003	0.034	0.035	94.1	0.571	-0.005	0.037	0.038	94.5
γ_1	0.307	0	0.028	0.028	95.4	0.307	0.002	0.031	0.032	93.8
γ_2	-0.75	-0.009	0.12	0.119	95.1	-0.75	-0.009	0.132	0.132	95.3
n=400, HC										
β_0	0.405	0.007	0.262	0.257	95.2	-0.405	0.011	0.273	0.267	95.7
β_1	0.327	0.003	0.139	0.141	94.6	0.501	-0.001	0.136	0.134	95.9
γ_0	0.571	-0.003	0.036	0.036	94.3	0.571	-0.005	0.041	0.041	93.9
γ_1	0.307	0.004	0.036	0.037	94.2	0.307	0.005	0.044	0.046	94.5
γ_2	-0.75	-0.005	0.136	0.14	93.7	-0.75	-0.011	0.153	0.157	95.4
n=800, LC										
β_0	0.405	0.003	0.148	0.145	95.7	-0.405	0	0.147	0.146	95.3
β_1	0.327	0.001	0.081	0.08	95.4	0.501	0.001	0.078	0.073	96.2
γ_0	0.571	-0.001	0.024	0.025	93.3	0.571	-0.002	0.027	0.027	94.8
γ_1	0.307	-0.001	0.02	0.02	95.5	0.307	0.001	0.022	0.023	95.2
γ_2	-0.75	-0.008	0.083	0.082	95.3	-0.75	-0.004	0.092	0.093	94.8
n=800, HC										
β_0	0.405	0.007	0.185	0.187	94.8	-0.405	-0.004	0.193	0.197	95
β_1	0.327	-0.003	0.096	0.098	95.3	0.501	0.001	0.095	0.095	94.7
γ_0	0.571	-0.001	0.026	0.025	95.4	0.571	-0.003	0.029	0.028	94.7
γ_1	0.307	0.003	0.026	0.026	95	0.307	0.003	0.032	0.033	94
γ_2	-0.75	-0.001	0.095	0.097	95	-0.75	-0.006	0.107	0.11	95

Table 6. Bias, SE, RMSE, and CP for the estimates of the parameters of the Bernoulli cure rate model under proportional odds model with Weibull baseline.

	low cure rate					high cure rate				
Param	True	Bias	SE	RMSE	CP(95%)	True	Bias	SE	RMSE	CP(95%)
n=400, LC										
β_0	0.405	0.011	0.209	0.207	96.2	-0.405	-0.008	0.207	0.206	96.2
β_1	0.327	0.003	0.118	0.118	95.6	0.501	0.009	0.113	0.112	95.8
γ_0	0.571	-0.001	0.034	0.034	95	0.571	-0.004	0.038	0.038	94.1
γ_1	0.307	0.002	0.023	0.023	94.9	0.307	0.003	0.028	0.028	95.6
γ_2	-0.75	-0.004	0.107	0.103	95.5	-0.75	-0.006	0.12	0.125	94.2
n=400, HC										
β_0	0.405	-0.001	0.255	0.266	93.6	-0.405	-0.004	0.266	0.266	95.9
β_1	0.327	0.007	0.138	0.142	93.9	0.501	0.003	0.135	0.135	94.7
γ_0	0.571	-0.003	0.036	0.036	94.7	0.571	-0.005	0.041	0.04	94.7
γ_1	0.307	0.004	0.028	0.029	94.5	0.307	0.003	0.036	0.038	94.2
γ_2	-0.75	-0.002	0.116	0.121	94.5	-0.75	-0.011	0.135	0.14	94.4
n=800, LC										
β_0	0.405	0.005	0.147	0.152	94	-0.405	0.002	0.146	0.149	94.9
β_1	0.327	0.002	0.081	0.081	95.9	0.501	0.001	0.078	0.08	94.6
γ_0	0.571	0	0.024	0.024	95	0.571	-0.002	0.027	0.027	95
γ_1	0.307	0.001	0.016	0.017	93.4	0.307	0.001	0.02	0.02	94.6
γ_2	-0.75	0	0.074	0.073	95.7	-0.75	-0.004	0.084	0.086	94.2
n=800, HC										
β_0	0.405	-0.001	0.181	0.187	93.8	-0.405	-0.009	0.189	0.187	95.9
β_1	0.327	0.004	0.095	0.094	95.1	0.501	0.006	0.094	0.095	96.1
γ_0	0.571	-0.002	0.026	0.025	95.1	0.571	-0.002	0.029	0.028	94.6
γ_1	0.307	0.001	0.02	0.02	95.1	0.307	0.003	0.026	0.027	95.1
γ_2	-0.75	-0.003	0.081	0.083	94.2	-0.75	-0.002	0.095	0.095	94.5

Table 7. Bias, SE, RMSE, and CP for the cure rates of the Geometric cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0 (0.002)	0.049 (0.049)	0.051 (0.05)	74 (76.7)
		p_{02}	0.325	-0.001 (0)	0.03 (0.03)	0.03 (0.029)	93.5 (94.4)
		p_{03}	0.257	-0.001 (0)	0.028 (0.029)	0.027 (0.027)	94.6 (95.3)
		p_{04}	0.2	0.001 (0.001)	0.036 (0.037)	0.036 (0.036)	88 (88.8)
	heavy	p_{01}	0.4	-0.001 (0)	0.059 (0.064)	0.058 (0.066)	71.4 (70.5)
		p_{02}	0.325	-0.002 (0)	0.035 (0.038)	0.035 (0.04)	91.7 (91.2)
		p_{03}	0.257	-0.002 (0)	0.03 (0.032)	0.03 (0.032)	94.8 (94.5)
		p_{04}	0.2	0 (0.001)	0.038 (0.041)	0.039 (0.04)	87.9 (89.4)
800	light	p_{01}	0.4	0 (0.001)	0.035 (0.035)	0.035 (0.035)	73.9 (75.3)
		p_{02}	0.325	-0.001 (0.001)	0.021 (0.022)	0.022 (0.022)	92.1 (92.8)
		p_{03}	0.257	0 (0.001)	0.019 (0.02)	0.019 (0.02)	94.9 (94.3)
		p_{04}	0.2	0 (0.001)	0.025 (0.026)	0.024 (0.025)	89.5 (87.4)
	heavy	p_{01}	0.4	-0.001 (0)	0.042 (0.045)	0.042 (0.045)	71 (70.7)
		p_{02}	0.325	-0.001 (0)	0.025 (0.027)	0.025 (0.027)	92.1 (90.2)
		p_{03}	0.257	0 (-0.001)	0.021 (0.022)	0.02 (0.022)	95.9 (96.1)
		p_{04}	0.2	0.001 (0)	0.026 (0.028)	0.026 (0.027)	89.9 (90.8)
400	light	p_{01}	0.6	0 (0.002)	0.049 (0.047)	0.049 (0.047)	78.9 (79.7)
		p_{02}	0.476	0 (0.003)	0.033 (0.032)	0.034 (0.032)	92 (93.5)
		p_{03}	0.355	-0.001 (0.002)	0.03 (0.031)	0.031 (0.029)	94.4 (95.6)
		p_{04}	0.25	-0.001 (0.003)	0.038 (0.039)	0.039 (0.037)	87.6 (89.2)
	heavy	p_{01}	0.6	-0.001 (0.005)	0.061 (0.061)	0.063 (0.066)	73.4 (75.3)
		p_{02}	0.476	0 (0.006)	0.04 (0.042)	0.041 (0.044)	88.6 (88.8)
		p_{03}	0.355	-0.001 (0.005)	0.033 (0.035)	0.033 (0.033)	95.2 (95.1)
		p_{04}	0.25	0 (0.005)	0.041 (0.042)	0.041 (0.04)	89.9 (92.2)
800	light	p_{01}	0.6	0 (0.003)	0.035 (0.033)	0.035 (0.034)	77.5 (77.4)
		p_{02}	0.476	0 (0.003)	0.024 (0.023)	0.024 (0.024)	92 (91.6)
		p_{03}	0.355	0 (0.002)	0.021 (0.022)	0.021 (0.022)	94.4 (93.8)
		p_{04}	0.25	0.001 (0.001)	0.027 (0.027)	0.026 (0.026)	89.4 (88.1)
	heavy	p_{01}	0.6	0 (0.002)	0.044 (0.044)	0.044 (0.046)	74 (74.7)
		p_{02}	0.476	0 (0.003)	0.029 (0.03)	0.029 (0.032)	90.1 (87.4)
		p_{03}	0.355	0 (0.003)	0.023 (0.024)	0.024 (0.024)	95.3 (94.7)
		p_{04}	0.25	0 (0.003)	0.028 (0.029)	0.03 (0.028)	89.5 (91.9)

cure rate (0.6, 0.25) if the lifetime follows proportional odds model under log-logistic baseline with parameter $\gamma_0 = 01.75$, $\gamma_1 = 3.25$, $\gamma_2 = -0.75$. Here, we carry out the model discrimination by two methods, namely, Likelihood-based method and information-based method.

6.1. Likelihood-based method

We consider a likelihood ratio test for the null hypothesis H_0 that the competing risk follows one of the three special cases of COM-Poisson distribution, namely, Geometric $\phi = 0$, Poisson $\phi = 1$, and Bernoulli $\phi \longrightarrow \infty$ versus the alternative hypothesis H_a that the competing risk follows the COM-Poisson distribution. The test statistic is taken as $\Lambda = -2(\hat{l}_0 - \hat{l})$, where $\hat{l}_0$ and $\hat{l}$ are the values of the maximized log-likelihood function under the null and alternative hypotheses, respectively. The asymptotic distribution of the test statistic Λ, under $H_0 : \phi = 1$ follows a χ^2 distribution with one degree of freedom. However, the boundary distribution of the test statistic Λ when $\phi = 0$ (Geometric) and $\phi \longrightarrow \infty$ (Poisson) has a mixture distribution of χ^2_0 and χ^2_1 distributions such that $P(\Lambda \leq \lambda) = \frac{1}{2} + \frac{1}{2}\chi^2_1$, where χ^2_0 is chi-square distribution with 0 degrees of freedom and χ^2_1 is the chi-square distribution with one degree of freedom.

The values of ϕ used in the profile likelihood approach for the COM-Poisson distribution are {0,0.25,0.5,2/3,1,1.5,2,4,∞}. Figure 1 provides the histograms of the test statistics Λ on the Poisson cure rate model with proportional odds assumption under log-logistic baseline when sample size equals to 400 and 4000 over 1000 generated datasets. These plots also display the probability density of chi-square distribution with one

Table 8. Bias, SE, RMSE, and CP for the cure rates of the Poisson cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0 (0)	0.049 (0.05)	0.048 (0.049)	75.9 (75)
		p_{02}	0.325	-0.001 (-0.001)	0.03 (0.03)	0.029 (0.03)	93.9 (93.5)
		p_{03}	0.257	-0.001 (-0.001)	0.028 (0.028)	0.027 (0.028)	95.7 (95.3)
		p_{04}	0.2	0 (0.001)	0.035 (0.036)	0.035 (0.036)	88.2 (88.9)
	heavy	p_{01}	0.4	-0.001 (0)	0.058 (0.062)	0.059 (0.06)	73.2 (73.6)
		p_{02}	0.325	-0.001 (-0.001)	0.034 (0.037)	0.034 (0.036)	91.8 (91.8)
		p_{03}	0.257	-0.001 (-0.001)	0.029 (0.031)	0.029 (0.032)	95 (94.3)
		p_{04}	0.2	0.001 (0.001)	0.038 (0.039)	0.038 (0.041)	88.8 (87.2)
800	light	p_{01}	0.4	0 (0)	0.035 (0.035)	0.034 (0.035)	77 (75.3)
		p_{02}	0.325	0 (-0.001)	0.021 (0.022)	0.021 (0.021)	93.3 (94.7)
		p_{03}	0.257	0 (0)	0.019 (0.02)	0.02 (0.02)	94.4 (95.3)
		p_{04}	0.2	0.001 (0)	0.025 (0.025)	0.025 (0.025)	89.3 (87.1)
	heavy	p_{01}	0.4	-0.001 (-0.001)	0.042 (0.044)	0.042 (0.044)	71 (72.2)
		p_{02}	0.325	-0.001 (0)	0.025 (0.026)	0.025 (0.026)	91.2 (91.4)
		p_{03}	0.257	0 (0.001)	0.021 (0.022)	0.021 (0.022)	96.1 (94.8)
		p_{04}	0.2	0.001 (0.002)	0.026 (0.027)	0.027 (0.028)	88 (88.2)
400	light	p_{01}	0.6	0.001 (-0.001)	0.049 (0.049)	0.048 (0.051)	79 (77.8)
		p_{02}	0.476	0.001 (-0.002)	0.033 (0.033)	0.034 (0.033)	91 (92.2)
		p_{03}	0.355	0.001 (-0.002)	0.03 (0.031)	0.031 (0.031)	94.1 (95.1)
		p_{04}	0.25	0.001 (-0.001)	0.038 (0.039)	0.038 (0.04)	87.5 (86.8)
	heavy	p_{01}	0.6	0 (-0.004)	0.06 (0.065)	0.061 (0.063)	73.5 (76.3)
		p_{02}	0.476	0 (-0.002)	0.04 (0.043)	0.041 (0.042)	90.5 (90.3)
		p_{03}	0.355	-0.002 (-0.001)	0.033 (0.034)	0.033 (0.034)	94.7 (95.2)
		p_{04}	0.25	-0.001 (0.001)	0.041 (0.042)	0.04 (0.042)	90.3 (91.7)
800	light	p_{01}	0.6	0.001 (0)	0.035 (0.035)	0.036 (0.035)	77.6 (76.4)
		p_{02}	0.476	0.001 (0)	0.024 (0.024)	0.024 (0.024)	91.9 (90.5)
		p_{03}	0.355	0.001 (0)	0.021 (0.022)	0.021 (0.021)	95.2 (94.3)
		p_{04}	0.25	0.001 (0)	0.026 (0.027)	0.026 (0.025)	88.5 (90.7)
	heavy	p_{01}	0.6	0.001 (0)	0.043 (0.046)	0.043 (0.047)	75 (74.2)
		p_{02}	0.476	0.001 (0.001)	0.029 (0.031)	0.029 (0.031)	90 (89.5)
		p_{03}	0.355	0.001 (0.001)	0.023 (0.024)	0.023 (0.024)	94.6 (95.3)
		p_{04}	0.25	0.001 (0.002)	0.028 (0.029)	0.029 (0.029)	90.1 (91)

degree of freedom, and 90% quantiles. The histogram of Λ is not close to the asymptotic distribution when sample size is small while they become close as the sample size increases. This suggests that a parametric bootstrap method would be better when the sample size is small, and we will describe this method in detail in the illustrative example section. Incidentally, the same was observed under the proportional odds assumption with Weibull baseline.

6.2. Information-based method

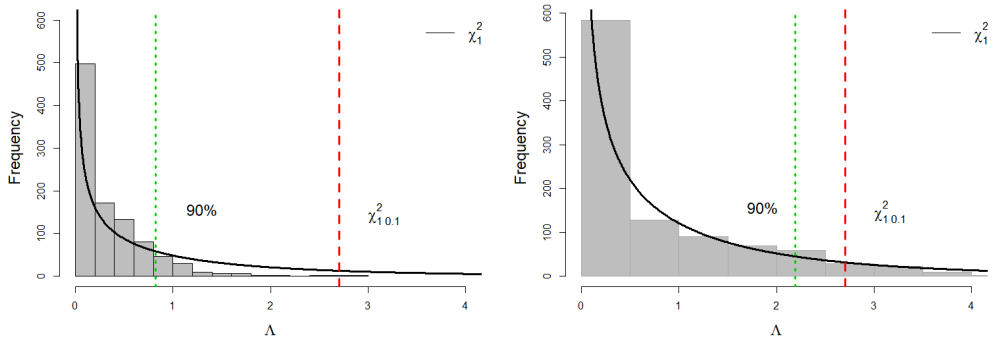
We use Akaike information criterion (AIC) and Bayesian information criterion (BIC) for model selection among Geometric, Poisson, and Bernoulli distribution cure rate models. The AIC and BIC are given by

$$\text{AIC} = 2k - 2\hat{l} \text{ and } \text{BIC} = k \log n - 2\hat{l}, \tag{38}$$

where k is the number of model parameters to be estimated, $\hat{l}$ is the maximized likelihood value, and n is the sample size. We select the model with the smallest value of AIC or BIC. AIC and BIC would give us the same selection rate since $k = 5$, $n = 400$ or 800 are the same in each of the scenarios among Geometric, Poisson, and Bernoulli distributions, i.e., the model with the largest $\hat{l}$ is the model that fits the data best. We examine the total relative bias (TRB) and total root mean square error (TRMSE) due to misspecification of the cure rate model. TRB is the sum of the absolute bias of the estimated cure rates to that of the true cure rates for each of the four groups. Similarly, TRMSE is the sum of the absolute MSE of the estimated cure rates. TRB and TRMSE due to misspecification is

Table 9. Bias, SE, RMSE, and CP for the cure rates of the Bernoulli cure rate model under proportional odds with log-logistic (Weibull) baseline.

n	C	p_0	True	Bias	SE	RMSE	CP(95%)
400	light	p_{01}	0.4	0.001 (-0.002)	0.049 (0.05)	0.051 (0.049)	74.4 (75.1)
		p_{02}	0.325	0 (-0.002)	0.03 (0.03)	0.031 (0.03)	92.2 (92.1)
		p_{03}	0.257	0 (-0.002)	0.028 (0.028)	0.028 (0.029)	94.8 (93)
		p_{04}	0.2	0.001 (-0.001)	0.036 (0.036)	0.035 (0.037)	88 (86.4)
	heavy	p_{01}	0.4	0.004 (0.002)	0.059 (0.06)	0.061 (0.063)	71.3 (71.9)
		p_{02}	0.325	0.001 (0)	0.035 (0.036)	0.035 (0.037)	91.6 (90.9)
		p_{03}	0.257	-0.001 (-0.001)	0.03 (0.031)	0.029 (0.031)	95.6 (95.5)
		p_{04}	0.2	0 (0)	0.038 (0.039)	0.038 (0.04)	89.3 (88.3)
800	light	p_{01}	0.4	0 (-0.001)	0.035 (0.035)	0.035 (0.036)	76 (72.3)
		p_{02}	0.325	0 (-0.001)	0.021 (0.021)	0.021 (0.022)	92 (92.3)
		p_{03}	0.257	0 (-0.001)	0.019 (0.02)	0.019 (0.02)	95.5 (93.7)
		p_{04}	0.2	0 (-0.001)	0.025 (0.025)	0.024 (0.025)	88.5 (87.2)
	heavy	p_{01}	0.4	0.002 (0.001)	0.042 (0.043)	0.043 (0.045)	72.6 (71.2)
		p_{02}	0.325	0.001 (0)	0.025 (0.026)	0.025 (0.027)	90.5 (90)
		p_{03}	0.257	0 (0)	0.021 (0.021)	0.02 (0.022)	96.5 (94.5)
		p_{04}	0.2	0 (0)	0.026 (0.027)	0.025 (0.027)	89.4 (89.1)
400	light	p_{01}	0.6	0.004 (0.001)	0.049 (0.049)	0.049 (0.049)	78.4 (76.9)
		p_{02}	0.476	0.002 (0)	0.033 (0.033)	0.033 (0.034)	92.8 (92.1)
		p_{03}	0.355	0 (-0.002)	0.03 (0.031)	0.031 (0.032)	94.6 (94.4)
		p_{04}	0.25	0 (-0.001)	0.038 (0.039)	0.039 (0.039)	88.3 (87)
	heavy	p_{01}	0.6	-0.002 (-0.001)	0.061 (0.063)	0.058 (0.063)	76 (75.3)
		p_{02}	0.476	0 (0)	0.04 (0.042)	0.038 (0.042)	92.5 (90.8)
		p_{03}	0.355	0.001 (0)	0.033 (0.034)	0.032 (0.034)	95.7 (95.5)
		p_{04}	0.25	0.003 (0.001)	0.041 (0.043)	0.041 (0.042)	90 (88.6)
800	light	p_{01}	0.6	0.001 (-0.001)	0.035 (0.035)	0.035 (0.035)	77.1 (76.3)
		p_{02}	0.476	0 (-0.001)	0.024 (0.024)	0.024 (0.024)	92 (90.9)
		p_{03}	0.355	-0.001 (-0.001)	0.021 (0.022)	0.022 (0.023)	93.7 (93.1)
		p_{04}	0.25	-0.001 (0)	0.026 (0.027)	0.027 (0.028)	88.8 (86.3)
	heavy	p_{01}	0.6	0 (0.001)	0.043 (0.045)	0.041 (0.045)	75.8 (73.9)
		p_{02}	0.476	0 (0.001)	0.029 (0.03)	0.028 (0.029)	90.8 (90.7)
		p_{03}	0.355	0.001 (0)	0.023 (0.024)	0.023 (0.024)	94.7 (95.4)
		p_{04}	0.25	0.001 (0)	0.028 (0.029)	0.028 (0.03)	90 (89.7)



(a) (b)

Figure 1. Histogram of Λ for the Poisson cure rate model under Proportional odds assumption with log-logistic baseline, *n*=400 (left), *n*=4000 (right).

Table 10. TRB (in %) in the estimation of cure proportions when fitting different models for a given true model under proportional odds due to misspecification.

			Weibull Baseline			Loglogistic Baseline		
n	censoring	Fitted Model	Geometric	Poisson	Bernoulli	Geometric	Poisson	Bernoulli
400	light	Geometric	-	1.4	-1.5	-	0.3	-0.5
		Poisson	48	-	-0.4	0.2	-	0.8
		Bernoulli	80.8	1.4	-	0.4	0.3	-
	heavy	Geometric	-	0.9	1.5	-	2.6	1.1
		Poisson	45.9	-	1.1	0.4	-	-1.4
		Bernoulli	99.8	0	-	1.1	1.3	-
800	light	Geometric	-	1.4	-0.3	-	1.2	1.4
		Poisson	47.1	-	1.7	1	-	1.5
		Bernoulli	82.7	2	-	-0.1	-1.1	-
	heavy	Geometric	-	-0.1	1.5	-	2.5	1.5
		Poisson	48.3	-	2.5	0.9	-	1.4
		Bernoulli	100.6	-0.5	-	0.8	0.3	-

Table 11. TRMSE in the estimation of cure proportions when fitting different models for a given true model under proportional odds due to misspecification.

			Weibull Baseline			Loglogistic Baseline		
n	censoring	Fitted Model	Geometric	Poisson	Bernoulli	Geometric	Poisson	Bernoulli
400	light	Geometric	-	0.999	0.961	-	0.3	-0.5
		Poisson	2.081	-	1.407	0.2	-	0.8
		Bernoulli	3.354	1.434	-	0.4	0.3	-
	heavy	Geometric	-	1.013	1.023	-	2.6	1.1
		Poisson	1.637	-	1.266	0.4	-	-1.4
		Bernoulli	3.73	0.991	-	1.1	1.3	-
800	light	Geometric	-	1.026	1	-	1.2	1.4
		Poisson	3.692	-	2.071	1	-	1.5
		Bernoulli	6.693	1.539	-	-0.1	-1.1	-
	heavy	Geometric	-	0.967	0.97	-	2.5	1.5
		Poisson	3.26	-	1.38	0.9	-	1.4
		Bernoulli	7.868	0.996	-	0.8	0.3	-

defined as the difference of TRB and TRMSE between the true model and the wrongly specified model (see Tables 10 and 11).

Table 12 shows that the selection rates under AIC or BIC for the correct models increase as the sample size increases, while it decrease as censoring rate increases, and the selection rates for the correct models are always the highest among all the cases. If the true model is under Weibull baseline, the rate to select log-logistic as the true baseline is low, and the rate becomes even lower for large sample size, or light censoring.

7. Illustration with melanoma data

In this section, we consider a cutaneous melanoma (a type of manlignant cancer) data to illustrate the performance of the proposed methodology. The data was first introduced by [13], and subsequently studied by many authors including [2], [4], [5], [6],[1], [21]. These data were taken from [11], and were originally used to detect the prospective treatment performance on the high-dose interferon alfa-2b therapy in order to prevent the recurrence

Table 12. Selection rates based on Akaike information criterion under different settings. The data were simulated from proportional odds with Weibull baseline (γ_0= 0.571 γ_1= 0.307 γ_2= -0.75).

				True COM-Poisson Model				
n	censoring	Fitted Model		$\phi=0$	$\phi=0.5$	$\phi=1$	$\phi=2$	$\phi\longrightarrow\infty$
400	light	Weibull	Geometric	**0.47**	0.0551	0.0828	0.0262	0.004
			Poisson	0.235	0.2906	**0.4919**	0.2553	0.2913
			Bernoulli	0.19	0.3567	0.3758	0.5146	**0.6996**
		log-logistic	Geometric	0.039	0.0802	0.0141	0.0585	0
			Poisson	0.037	0.1283	0.0222	0.0898	0.003
			Bernoulli	0.029	0.0892	0.0131	0.0555	0.002
	heavy	Weibull	Geometric	**0.344**	0.0452	0.1047	0.0241	0.011
			Poisson	0.197	0.2199	0.3797	0.204	0.279
			Bernoulli	0.229	0.3624	0.3867	0.5065	**0.677**
		log-logistic	Geometric	0.065	0.1014	0.0393	0.0894	0.01
			Poisson	0.103	0.1627	0.0584	0.0985	0.012
			Bernoulli	0.062	0.1084	0.0312	0.0774	0.011
800	light	Weibull	Geometric	**0.587**	0.0512	0.0849	0.0271	0.002
			Poisson	0.266	0.4137	**0.5905**	0.3601	0.2402
			Bernoulli	0.105	0.2801	0.3104	0.4935	**0.7578**
		log-logistic	Geometric	0.015	0.0753	0.0051	0.0411	0
			Poisson	0.012	0.1265	0.0081	0.0431	0
			Bernoulli	0.015	0.0532	0.001	0.0351	0
	heavy	Weibull	Geometric	**0.433**	0.051	0.1085	0.0301	0.002
			Poisson	0.245	0.301	**0.5136**	0.2417	0.283
			Bernoulli	0.154	0.311	0.3126	0.5035	**0.709**
		log-logistic	Geometric	0.055	0.088	0.0111	0.0662	0.003
			Poisson	0.074	0.159	0.0352	0.0832	0.001
			Bernoulli	0.039	0.09	0.0191	0.0752	0.002

of the disease. The study included 427 patients in total from years 1991 to 1995 and follow up until year 1998. Among them, 10 patients were removed in our analysis due to the missingness of the tumor thickness data. The overall percentage of censored observations is 55.6%. The mean and standard deviation of the observed lifetimes are 3.18 and 1.69 in years, respectively. We choose the nodule categories based on the tumor thickness as the only covariate. The subjects were therefore divided into four different categories ($x=0,1,2,3$), with corresponding sample sizes $n_1=111$, $n_2=137$, $n_3=87$, $n_4=82$. The percentage of censored observations for each group are 67.57%, 61.31%, 52.87%, 32.93%. See Figure 2 for the lifetimes of susceptibles.

The initial values for the EM algorithm are chosen in the following way. We consider the censored rate as the over-estimated cured rate of groups one and four, and then calculated β_0 and β_1. The initial guess for $\boldsymbol{\gamma}$ are estimated from the linear relationship between Nelson-Aalen estimates of log odds and log t, i.e.,

$$\log O(t;\boldsymbol{\gamma}) = -\gamma_1 \log t + \gamma_1 \log\gamma_0 + x\gamma_2 \tag{39}$$

$$\log O(t;\boldsymbol{\gamma}) = x\gamma_2 - \log(e^{(\gamma_1 t)^{1/\gamma_0}} - 1) \approx -\frac{1}{\gamma_0}\log t - \frac{1}{\gamma_0}\log\gamma_1 + x\gamma_2, \tag{40}$$

for log-logistic and Weibull baseline distributions, respectively. We then fitted these data by Geometric ($\phi=0$), COM-Poisson ($\phi=0.5$), Poisson ($\phi=1$), COM-Poisson ($\phi=2$) and Bernoulli ($\phi\approx\infty$) cure rate model. These models, along with cure rate models proposed by [6] and [1], are compared on the basis of AIC and BIC. From Table 13, we observe that $\hat{\phi}\approx\infty$ and $\hat{\phi}=0$ provide the maximized log-likelihood values, which means that the Bernoulli and Geometric cure rate models provide a good fit for the data under log-logistic odds and Weibull odds models, respectively. Moreover, the maximized log-likelihood value increases and decreases as ϕ increases

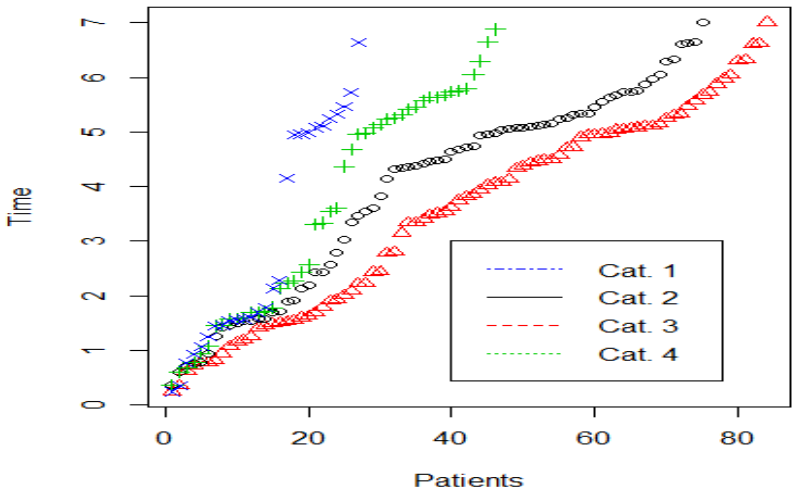


Figure 2. Cutaneous Melanoma data

under the log-logistic odds and Weibull odds, respectively. The proposed model based on Weibull odds provides smaller values of $\hat{l}$, AIC, and most of BIC values than the model for Weibull lifetimes, and the proportional hazards lifetimes with a Weibull baseline. Our proposed model based on log-logistic odds provides the smallest $\hat{l}$, AIC and most of BIC values among all the models. In conclusion, $\hat{\phi} = 0$ under proportional odds model with a log-logistic distribution as the baseline gives the best fit for the data.

From Table 13, we can see that the maximum difference between $\hat{l}$ is only 0.09, which means that the difference between the maximized likelihood value is $e^{0.09} = 1.09$ among different ϕ's. Table 14 presents the estimates, standard errors and 95% CI of the cure rates stratified by nodule category. The confidence interval for all the models do not overlap for the first and forth nodule categories. Table 15 presents the MLES and their standard errors for different cure rate models under proportional odds assumption. It is to be noted that $\hat{l}$ is very close to each other under proportional odds model with log-logistic baseline even-though the estimates are quite different. The same behaviour was also seen in the model discrimination section.

In order to further investigate the effect of ϕ under the COM-Poisson distribution, we fix ϕ from 0 to 5 with an increment of 0.1 and evaluate the maximum log-likelihood value for each ϕ through likelihood approach. And we test the null hypothesis $H_0 : \phi = \infty$ vs. $H_1 : 0 \leq \phi < \infty$ using the likelihood ratio test for the log-logistic odds baseline. And $H_0 : \phi = 0$ vs. $H_1 : \phi > 0$ under the Weibull odds baseline. The test statistic is given by $\Lambda = -2(\hat{l}_0 - \hat{l})$. Figure 3 shows that the likelihood ratio test statistic decreases and increases as ϕ increases for log-logistic and Weibull odds, respectively, which suggest that the maximized likelihoods increase and decrease as ϕ increases. As we mentioned during the model discrimination, the asymptotic distribution is not suitable when the sample size is small. So, we use a bootstrap method to obtain the distribution of the likelihood ratio test statistic Λ. We generated 1000 samples from Geometric, Poisson, Bernoulli cure rate models under proportional odds model with log-logistic and Weibull distributions, respectively. For each of the dataset, we fit the true cure rate model as well as the COM-Poisson cure rate model, then we calculate the values of Λ. The histograms of Λ are given in Figure 4. The p-value is the proportion of times Λ greater than the corresponding value determined from the data. We obtained p-values of 0.142, 0.132 and 0.681 if we test for Geometric, Poisson and Bernoulli cure rate models with log-logistic odds. Also, we obtained p-values of 0.599, 0.001 and 0.000 if we test for Geometric, Poisson and Bernoulli cure rate models under Weibull odds. Moreover, it would be of interest to get an acceptable range of ϕ if we are using the Weibull baseline for the proportional odds model. Figure 3 present the values of Λ against ϕ with

Table 13. AIC, BIC and $\hat{l}$ under different cure rate models

	COM-Poisson	Geometric $\phi=0$	$\phi=0.5$	Poisson $\phi=1$	$\phi=2$	Bernoulli $\phi=\infty$
Proportional	AIC	1022.863	1022.845	1022.821	1022.768	**1022.683**
Odds	BIC	1043.029	1043.011	1042.986	1042.933	**1042.849**
log-logistic	$\hat{l}$	-506.432	-506.423	-506.41	-506.384	**-506.342**
Proportional	AIC	**1025.644**	1026.014	1026.374	1026.862	1027.414
Odds	BIC	**1045.809**	1046.179	1046.539	1047.027	1047.579
Weibull	$\hat{l}$	**-507.822**	-508.007	-508.187	-508.431	-508.707
Proportional	AIC	**1028.6766**	1032.4676	1034.1606	1036.0434	1038.948
Hazard	BIC	**1048.842**	1052.633	1054.326	1056.2088	1059.114
Weibull*	$\hat{l}$	**-509.3383**	-511.2338	-512.0803	-513.0217	-514.474
	AIC	**1026.838**	1032.388	1034.788	1037.792	1043.182
Weibull**	BIC	**1042.97**	1048.52	1050.92	1053.924	1059.314
	$\hat{l}$	**-509.419**	-512.194	-513.394	-514.896	-517.591

Section $*$ is taken from [6]. Section $**$ is taken from [1].

Table 14. Estimates, standard errors and 95% CI for the cure rates

Cure rate model, Baseline	Bernoulli, log-logistic			Geometric, Weibull		
X	$\hat{p_0}$	SE	95% C.I.	$\hat{p_0}$	SE	95% C.I.
1	0.602	0.053	(0.497 , 0.706)	0.63	0.047	(0.538 , 0.721)
2	0.508	0.038	(0.434 , 0.583)	0.537	0.034	(0.471 , 0.603)
3	0.415	0.037	(0.343 , 0.486)	0.442	0.033	(0.377 , 0.507)
4	0.327	0.049	(0.231 , 0.422)	0.351	0.046	(0.261 , 0.44)

Table 15. MLEs and SEs of the model parameters

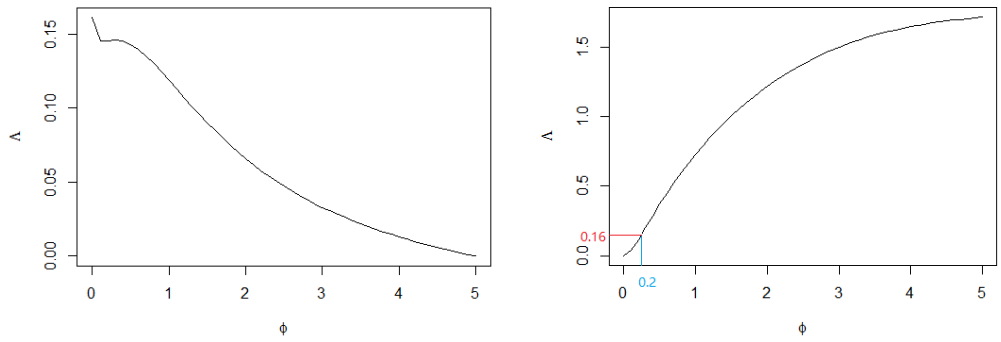
Cure rate model,PO Baseline	Bernoulli, log-logistic		Geometric, Weibull	
Param	MLEs	SE	MLEs	SE
β_0	-0.413	0.2226	-0.53	0.2003
β_1	0.379	0.1151	0.382	0.1049
γ_0	2.461	0.2995	0.488	0.0361
γ_1	2.266	0.1877	0.298	0.0338
γ_2	-0.473	0.1254	-0.293	0.1385

$\phi \in [0,5]$. We may reject the null hypothesis $H_0 : \phi = 0$ with 10% level of significance if Λ is greater than 0.16. This implies that $\phi \in [0, 0.2)$, and the Geometric model under Weibull odds adequately fits the data.

We also set up a test on the effect of the proportional odds parameter: $H_0 : \gamma_2 = 0$ as null hypothesis vs. $H_1 : \gamma_2 \neq 0$ as alternative hypothesis for Geometric, Poisson, Bernoulli cure rate models with $\phi = 0, 1, \infty$ under Weibull (log-logistic) baseline. Note that the covariate or the nodule categories would not affect the analysis if $\gamma_2 = 0$, and the lifetime would just follow a Weibull (log-logistic) distribution. The test statistic turned out to be 3.194 (14.95), 10.414 (19.564), 17.767 (25.99) with corresponding p-values 0.0739(0.00011), 0.00125 (9.73×10^{-6}), 0.000025 (3.4×10^{-7}). Most of the p-values were less than 0.05, which shows that the proportional odds model provides a better fit than a constant lifetime model over the four nodule categories.

Deviance residual is examined to check the error, which is defined as

$$D_i = sign(I_i + \log\hat{S}_p(t_i))\sqrt{-2(I_i + \log\hat{S}_p(t_i) + I_i \log[-(\log\hat{S}_p(t_i))])} \tag{41}$$



(a) (b)

Figure 3. The plot of $\Lambda = -2(\hat{l_0} - \hat{l_1})$ vs ϕ under log-logistic baseline (left) and Weibull baseline (right), for cutaneous melanoma data.

Figure 7, 8 and 5, 6 present the deviance residuals as well as qq plot for various fitted cure rate models. It can be seen that the deviance residuals are distributed around 0, and satisfied the normality assumption.

7.1. Model diagnosis

In this section, we check if the proportional odds model with log-logistic baseline assumption on the lifetime is met for the cutaneous melanoma data that we have discussed in the previous section. From (40), we can see that $log\hat{O}(t_{(i)})$ and $logt_{(i)}$ should have a linear relationship between them. For the cutaneous melanoma data, we calculated the cumulative hazard function $\hat{H}(t)$ as the observed hazard through the non-parametric Nelson-Aalen estimator, and then get the log-odds function accordingly as

$$\hat{H}(t) = \sum_{t_i \le t} \frac{d_i}{n_i};\ \hat{S}(t) = e^{-\hat{H}(t)};\ log(\hat{O}) = log(\hat{S}(t)) - log(1 - \hat{S}(t)), \tag{42}$$

where d_i and n_i are the number of events and total individuals at risk at time t_i, respectively. Figure 9 presents the scatter plot of $log\hat{O}(t_{(i)})$ vs. $logt_{(i)}$ for each category. The plot shows almost a linear relationship among the four groups. It is to be noticed that there is an intersection among $x = 2$ and $x = 3$ for short survival times which violates our proportional odds assumption. Figure 10 shows the difference of log-odds between nodule categories 1, 2, 3 and baseline 0. In this figure, the log odds for each of the nodule categories are calculated by using the linear interpolation within the range of discrete points from 0.4 to 2. From our proportional odds assumption, $logO - logO_0 = \boldsymbol{x\gamma}_2$, we know that the difference should be a linear horizontal line and does not depend on the time. However, the lines in Figure 10 look parallel but do show give a little curvature. Since our data include the cured individuals and are not independent, linear regression test may not be good for model diagnosis.

We use the parametric bootstrap and Monte Carlo methods to develop a goodness of fit test to check whether the Bernoulli cure rate model with proportional odds assumption under log-logistic baseline is sufficient. The critical region for this test will be to the left. We simulated 1000 data based on Bernoulli cure rate model with proportional odds survival assumption under log-logistic baseline. The parameters are $\beta_0 = -0.413$, $\beta_1 = 0.379$, $\gamma_0 = 2.461$, $\gamma_1 = 2.266$, $\gamma_2 = -0.473$. The censored proportion for the nodule categories are $(0.676, 0.613, 0.529, 0.329)$, respectively. We calculated the values of maximum likelihood $\hat{l}_1, ..., \hat{l}_{1000}$ for each of these generated datasets, and order them $(\hat{l})_1, ..., \hat{l}_{(1000)}$. Then, we determine the proportion of times the $\hat{l}$ is smaller than the maximum likelihood we obtained from the data as -506.342. Figure 11 presents the histogram of the log-likelihood over 1000

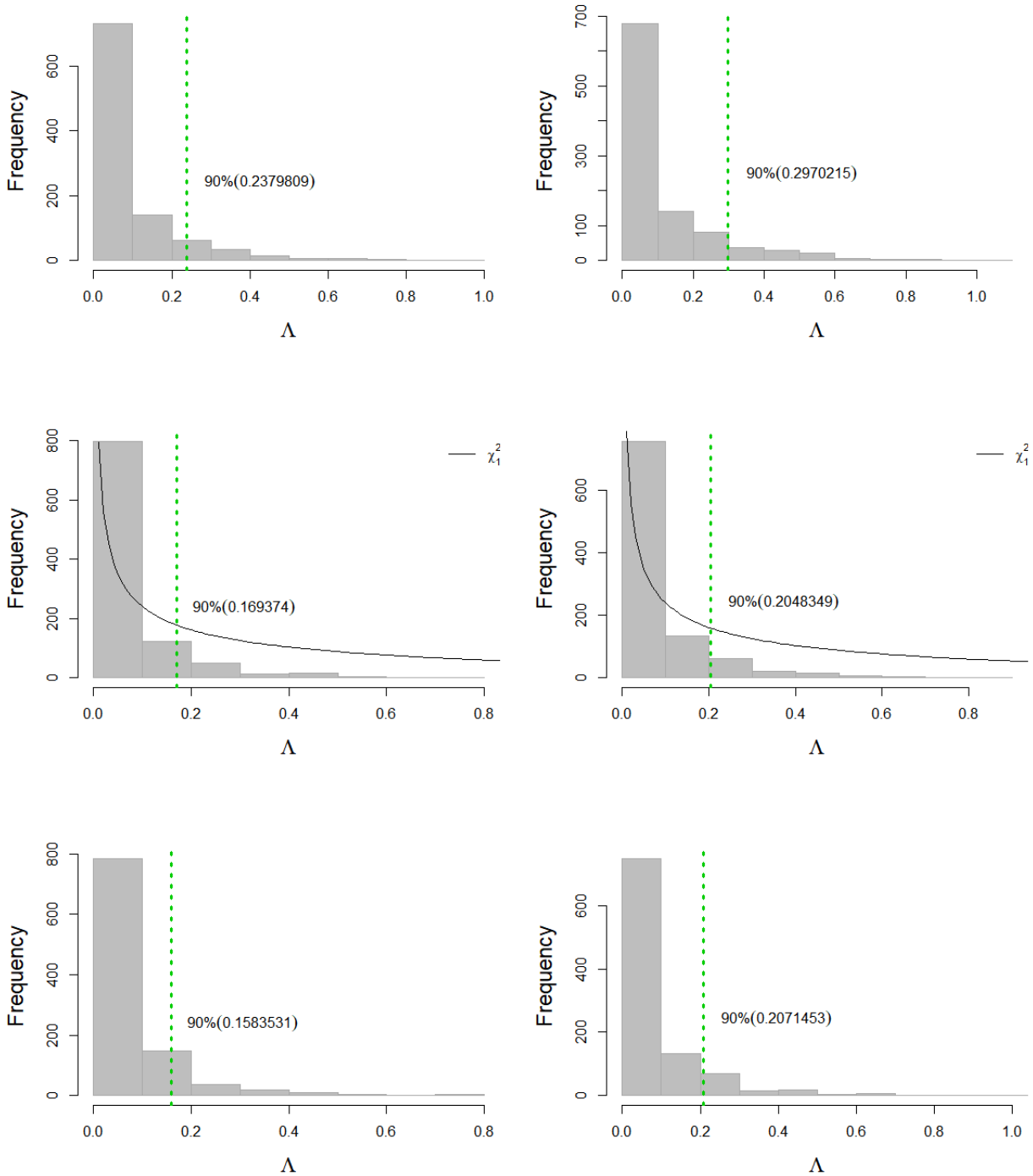


Figure 4. The histogram of $\Lambda = -2(\hat{l} - \hat{l_0})$ from 1000 generated datasets with respect to MLEs on Geometric (top), Poisson (middle), Bernoulli (bottom) cure rate model with the lifetime distribution as a proportional odds model with log-logistic(left) and Weibull (right) baseline .

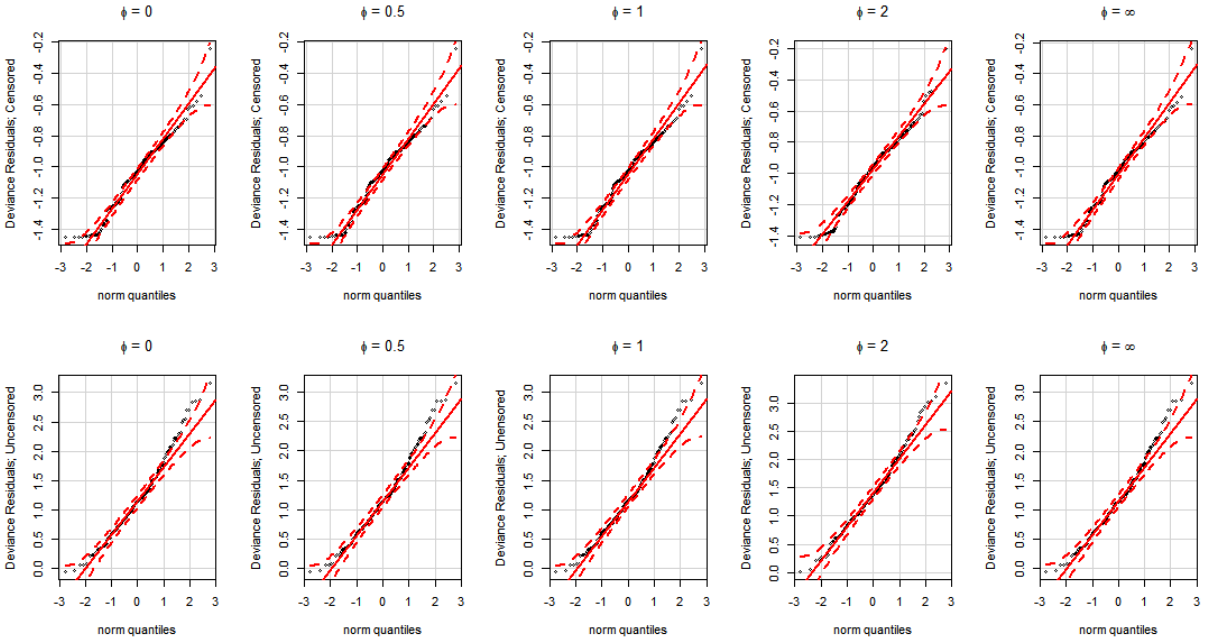


Figure 5. QQ plot for deviance residual on proportional odds model with log-logistic baseline for cutaneous melanoma data.

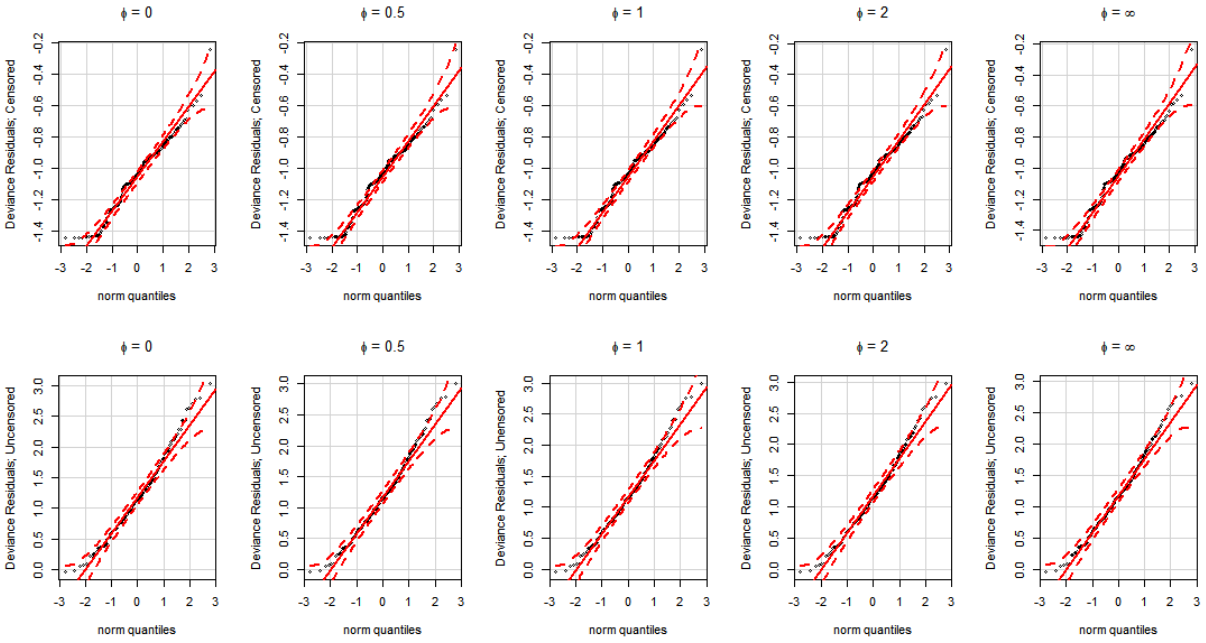


Figure 6. QQ plot for deviance residual on proportional odds model with Weibull baseline for cutaneous melanoma data.

simulated datasets. We obtained the p-value for this test as 0.895, which indicates that our model is quite good as suitable for these data.

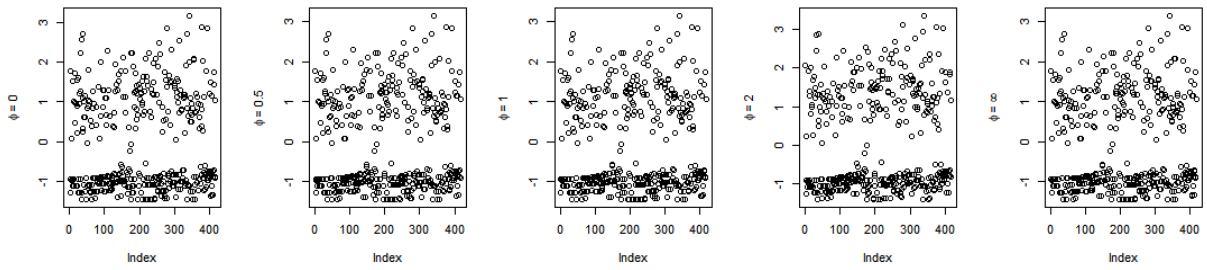


Figure 7. Deviance residual on proportional odds model with log-logistic baseline for cutaneous melanoma data.

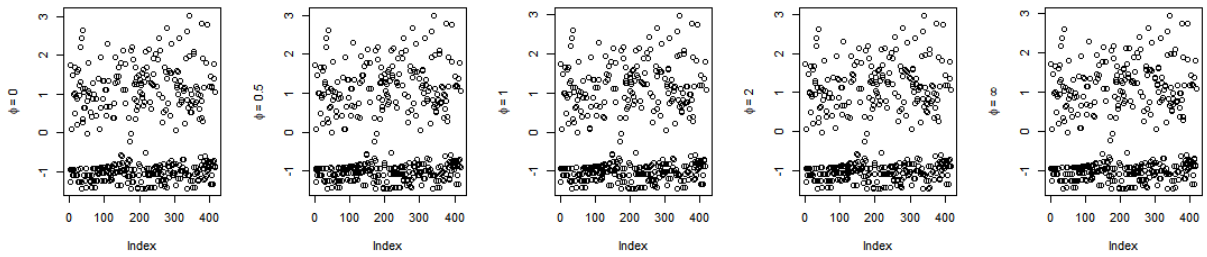


Figure 8. Deviance residual on proportional odds model with log-logistic baseline for cutaneous melanoma data.

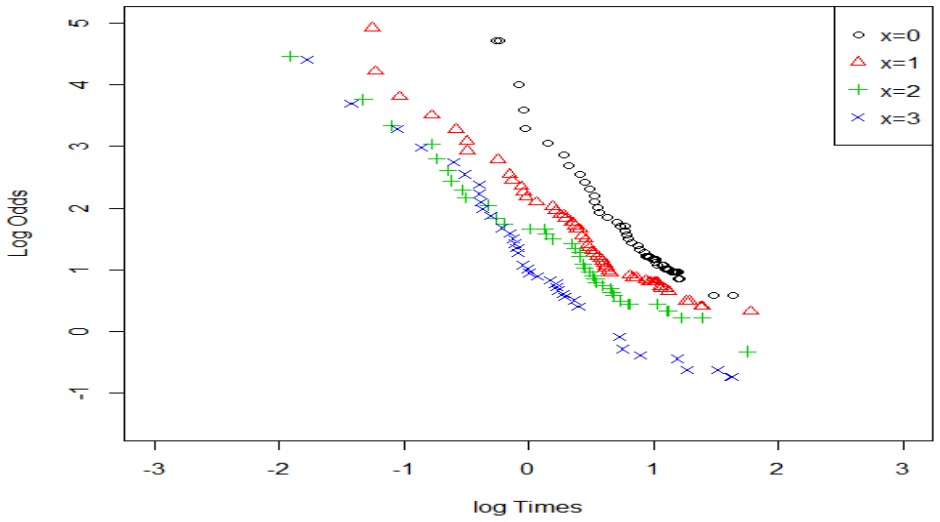


Figure 9. Log odds against log t for cutaneous melanoma data based on Nelson-Aalen estimator on patients who died.

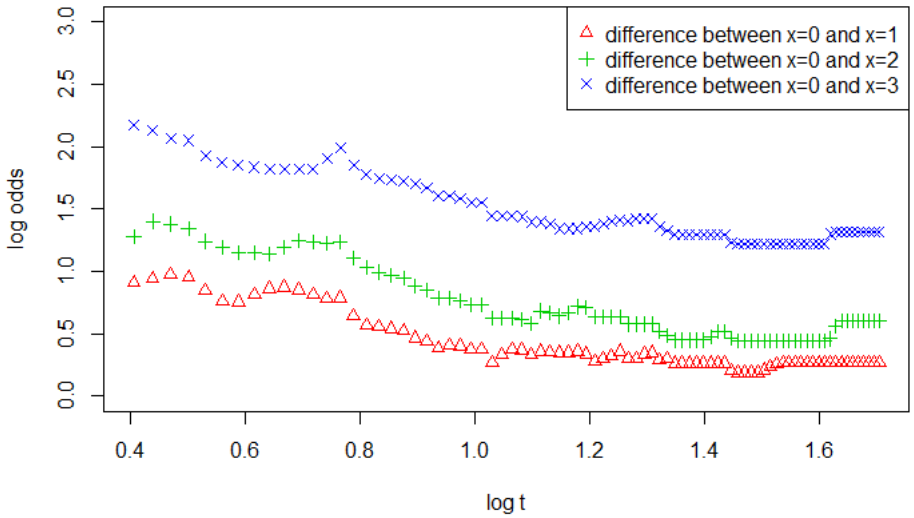


Figure 10. Difference of log odds between nodule categories against log t for cutaneous melanoma data based on Nelson-Aalen estimator.

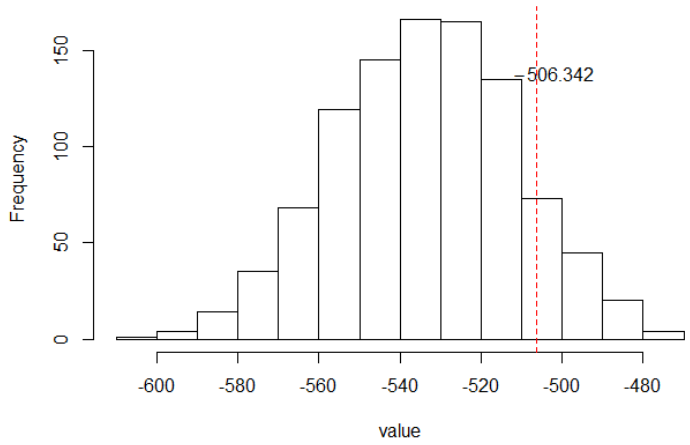


Figure 11. Histogram of $\hat{l}$

8. Concluding remarks

In this paper, we develop a flexible COM-Poisson cure rate model under a proportional odds assumption for the lifetime distribution of susceptibles with the baseline function being that of a Weibull distribution or log-logistic distribution. An EM algorithm is developed for the maximum likelihood estimation of the parameters from the proposed cure rate model. We perform an extensive Monte Carlo simulation study by varying sample sizes, censoring proportion, cure rates, parameters in different distributions to evaluate the performance of our proposed

methodology. Overall, our methodology provides accurate estimates of the model parameters as well as of the cure rates. Moreover, a real data on cutaneous melanoma is analyzed and model diagnosis is performed for illustrative purpose. There are many potential future works in this direction. One may consider the use of a non-parametric specification of the baseline distribution in the proportional odds model of the lifetimes of susceptible group. In addition, a destructive cure survival rate model, by including a damage or destruction term for the initial risk factors, can also be considered in the context of proportional odds model.

Acknowledgements

Our sincere thanks go to the reviewers and the guest editors for their useful comments and suggestions on an earlier version of this manuscript which led to this improved version.

Appendix

Results for the Special cases of COM-Poisson cure rate model

As mentioned earlier, the COM-Poisson distribution includes the Bernoulli, Poisson and Geometric distributions as special cases. Here, we detail the steps of the EM algorithm for these three special cure models.

Bernoulli cure rate model Let the competing cause random variable M follow a Bernoulli distribution with probability of success $\eta/(1+\eta)$. The probability density function for the whole population can then be expressed as

$$f_p(t_i,\boldsymbol{\theta}) = \frac{\eta}{1+\eta} f(t_i;\boldsymbol{\gamma}). \tag{43}$$

The survival function for the susceptible group is just the survival function for the time to event W, i.e., $S_s(t_i;\boldsymbol{\theta}) = S(t_i;\boldsymbol{\gamma})$. The inverse of the cure rate under this setting is $1/p_0 = 1+\eta$. We, therefore, have $H_\phi(\eta) = 1+\eta$ under the logistic link with a fixed value of ϕ, which implies $\eta = e^{\boldsymbol{x}_i'\boldsymbol{\beta}}$. The $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ function is then given by

$$\sum_{i\in\Delta_1} \boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{i\in\Delta_1} \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma}) - \sum_{i\in\Delta*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i'\boldsymbol{\beta} + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma}). \tag{44}$$

It is readily seen that some of the terms in the Q function are only corresponding to $\boldsymbol{\beta}$ while the others are only corresponding to $\boldsymbol{\gamma}$. So, it can be split into two parts as follows:

$$Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)}) = Q_1(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) + Q_2(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}), \tag{45}$$

$$Q_1(\boldsymbol{\gamma}, \boldsymbol{\pi}^{(k)}) = \sum_{I\in\Delta_1} \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma}) + \sum_{I_0} \boldsymbol{\pi}_i^{(k)} \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma}), \tag{46}$$

$$Q_2(\boldsymbol{\beta}, \boldsymbol{\pi}^{(k)}) = \sum_{i\in\Delta_1} \boldsymbol{x}_i'\boldsymbol{\beta} - \sum_{i\in\Delta*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}_i^{(k)} \boldsymbol{x}_i'\boldsymbol{\beta}, \tag{47}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i;\boldsymbol{\gamma}^{(k)})}{1 + e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}} S(t_i;\boldsymbol{\gamma}^{(k)})} \tag{48}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\frac{\partial Q}{\partial \gamma_j} = \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j},$$
$$\frac{\partial Q}{\partial \beta_l} = \sum_{i\in\Delta_1} x_{il} - \sum_{i\in\Delta_*} \frac{x_{il} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} x_{il},$$
$$\frac{\partial^2 Q}{\partial \beta_l \partial \beta_{l'}} = -\sum_{i\in\Delta_*} \frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2},$$
$$\frac{\partial^2 Q}{\partial \gamma_j \partial \gamma_{j'}} = \sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\partial \log S^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}},$$

for $l, l' = 0, \ldots, p$, $j, j' = 1, 2$, $h = 21, \ldots, 2p$, $i = 1, \ldots, n$.

Poisson cure rate model Let the competing cause random variable M follow a Poisson distribution. The probability density function for the whole population in this case can be expressed as

$$f_p(t_i;\boldsymbol{\theta}) = [\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] f(t_i;\boldsymbol{\gamma})(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{(S(t_i;\boldsymbol{\gamma})-1)}, \tag{49}$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = [(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{\gamma})} - 1] e^{-\boldsymbol{x}_i\boldsymbol{\beta}}. \tag{50}$$

The cure rate is $p_0 = e^{-\eta}$. We would then have $H_\phi(\eta) = e^\eta$ under the logistic link with a fixed value of ϕ, which implies that $\eta = \ln(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$. The $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ function is then given by

$$\begin{aligned} Q = &\sum_{I\in\Delta_1} \log[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})] + \sum_{I\in\Delta_1} \log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma}) + \sum_{I\in\Delta_1} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma}) \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) \\ &- \sum_{i\in\Delta_*} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) + \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \log((1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})} - 1) \end{aligned} \tag{51}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = 1 - (1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}})^{-S(t_i;\boldsymbol{\gamma}^{(k)})} \tag{52}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*, \boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\begin{aligned} \frac{\partial Q}{\partial \beta_l} = &\sum_{i\in\Delta_1} \frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} - \sum_{i\in\Delta_*} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} \\ &+ \sum_{i\in\Delta_0} \frac{\boldsymbol{\pi}^{(k)} S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}} \frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}, \\ \frac{\partial Q}{\partial \gamma_j} = &\sum_{i\in\Delta_1} \frac{\partial \log f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} + \sum_{i\in\Delta_1} \frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} \log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}) \\ &+ \sum_{i\in\Delta_0} \boldsymbol{\pi}^{(k)} \frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}} \frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j}, \end{aligned}$$

$$\frac{\partial^2 Q}{\partial\beta_l\partial\beta_{l'}} = \sum_{i\in\Delta_1}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}\{\frac{1}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}[1-\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}]+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\}-\sum_{i\in\Delta_*}\frac{x_{il}x_{il'}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}[1-\frac{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}-1}],$$
$$\frac{\partial^2 Q}{\partial\beta_l\partial\gamma_j} = \sum_{i\in\Delta_1}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}$$
$$+\sum_{i\in\Delta_0}\frac{\boldsymbol{\pi}^{(k)}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\frac{1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}(1+S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}))}{[1-(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},$$
$$\frac{\partial^2 Q}{\partial\gamma_j\partial\gamma_{j'}} = \sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_1}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})}{1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}$$
$$-\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}[\log(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})]^2}{[1-(1+e^{\boldsymbol{x}_i\boldsymbol{\beta}})^{-S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}]^2}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_{j'}},$$

for $l,l'=0,\ldots,p$, $j,j'=1,2$, $h=21,\ldots,2p$, $i=1,\ldots,n$.

Geometric cure rate model Let the competing cause random variable M follow a Geometric distribution. The probability density function for the whole population in this case can be expressed as

$$f_p(t_i;\boldsymbol{\theta}) = \frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}f(t_i,\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})^2}, \tag{53}$$

and the survival function for the susceptible group as

$$S_s(t_i;\boldsymbol{\theta}) = \frac{S(t_i;\boldsymbol{\gamma})}{R_G(t_i,\boldsymbol{\theta})}, \tag{54}$$

where $R_G(t_i,\boldsymbol{\theta}) = 1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}-e^{\boldsymbol{x}_i'\boldsymbol{\beta}}S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})$. The cure rate under this setting is $p_0=1-\eta$, and under the logistic link with a fixed value of ϕ, we would have $H_\phi(\eta)=(1-\eta)^{-1}$, which implies that $\eta = e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^{-1}$. The $Q(\boldsymbol{\theta}^*,\boldsymbol{\pi}^{(k)})$ function is then given by

$$Q = \sum_{I\in\Delta_1}\boldsymbol{x}_i'\boldsymbol{\beta}+\sum_{I\in\Delta_1}\log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})-\sum_{I\in\Delta_1}2\log R_G(t_i,\boldsymbol{\theta})-\sum_{i\in\Delta_0}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\boldsymbol{x}_i'\boldsymbol{\beta}$$
$$+\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\log S(t_i;\boldsymbol{\gamma})-\sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\log R_G(t_i,\boldsymbol{\theta}), \tag{55}$$

with the update step

$$\boldsymbol{\pi}_i^{(k)} = \frac{S(t_i;\boldsymbol{\gamma}^{(k)})e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}^{(k)}}} \tag{56}$$

for the ith censored observation. The required first- and second-order derivatives of $Q(\boldsymbol{\theta}^*,\boldsymbol{\pi}^{(k)})$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are as follows:

$$\frac{\partial Q}{\partial \beta_l} = \sum_{i\in\Delta_1} x_{il} - 2\sum_{i\in\Delta_1}\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \beta_l} - \sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}x_{il} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \beta_l},$$
$$\frac{\partial Q}{\partial \gamma_j} = \sum_{i\in\Delta_1}\Big(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} - 2\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \gamma_j}\Big) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\Big(\frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j} - \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \gamma_j}\Big),$$

$$\frac{\partial^2 Q}{\partial \beta_l \partial \beta_{l'}} = -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \beta_{l'}} - \sum_{i\in\Delta_0}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \beta_{l'}},$$
$$\frac{\partial^2 Q}{\partial \beta_l \partial \gamma_j} = -2\sum_{i\in\Delta_1}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j} - \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j},$$
$$\frac{\partial Q^2}{\partial \gamma_j \partial \gamma_{j'}} = \sum_{i\in\Delta_1}\Big(\frac{\partial \log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} - 2\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}}\Big) + \sum_{i\in\Delta_0}\boldsymbol{\pi}^{(k)}\Big(\frac{\partial \log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial \gamma_j \partial \gamma_{j'}} - \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}}\Big),$$

$$\frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \gamma_j} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial \gamma_j}, \quad \frac{\partial \log R_G(t_i,\boldsymbol{\theta})}{\partial \beta_l} = \frac{x_{ih}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})},$$
$$\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S^2(t_i,\boldsymbol{\theta})}{\partial \gamma_j \partial \gamma_{j'}} - \frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial \gamma_j}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial \gamma_{j'}},$$
$$\frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \beta_l} = \frac{x_{il}x_{il'}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}(1-S(t_i,\boldsymbol{\theta})), \quad \frac{\partial \log R_G^2(t_i,\boldsymbol{\theta})}{\partial \beta_l \partial \gamma_j} = -\frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial \gamma_j},$$

for $l, l' = 0, \ldots, p$, $j, j' = 1, 2$, $h = 21, \ldots, 2p$, $i = 1, \ldots, n$.

Observed information matrix

COM-Poisson cure rate model: The score functions, for a fixed value of ϕ, are

$$\frac{\partial l}{\partial \beta_l} = -\sum_{i\in\Delta_*} x_{il}\frac{e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}} + \sum_{i\in\Delta_1} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\frac{z_{21,i}}{z_{2,i}z_{01,i}}x_{il} + \sum_{i\in\Delta_0}\frac{z_{2,i}x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{z_{01,i}(1+z_{1i})},$$
$$\frac{\partial l}{\partial \gamma_h} = \sum_{i\in\Delta_1}\frac{\partial \log f(t_i,\boldsymbol{\gamma})}{\partial \gamma_h} + \sum_{i\in\Delta_1}\Big(\frac{z_{21,i}}{z_{2,i}} - 1\Big)\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial \gamma_h} + \sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial \log S(t_i,\boldsymbol{\gamma})}{\partial \gamma_h}.$$

Hence, the components of the observed information matrix, for a fixed value of ϕ, are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{-\sum_{i\in\Delta*}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}})^2}+\sum_{i\in\Delta_1}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{2,i}z_{01,i})^2}[z_{2,i}(z_{21,i}z_{01,i}+z_{31,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}})-z_{21,i}[z_{21,i}+\frac{z_{2,i}z_{02,i}}{z_{01,i}}]e^{\boldsymbol{x}'_i\boldsymbol{\beta}}]\\
&\quad+\sum_{i\in\Delta_0}\frac{x_{il}x'_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(z_{1,i}+1)^2z_{01,i}^2}[(\frac{z_{21,i}}{z_{01,i}}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}+z_{2,i})(z_{1,i}+1)z_{01,i}-z_{2,i}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}[\frac{z_{02,i}}{z_{01,i}}(z_{1,i}+1)+z_{2,i}]]\},\\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_h} &= -\{\sum_{i\in\Delta_1}x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2z_{01,i}}\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}+\sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{z_{01,i}(1+z_{1,i})}(z_{21,i}-\frac{z_{2,i}^2}{1+z_{1,i}})\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_h}\},\\
-\frac{\partial^2 l}{\partial\gamma_h\partial\gamma_{h'}} &= -\{\sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}+\sum_{i\in\Delta_1}(\frac{z_{21,i}}{z_{2,i}}-1)\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}\\
&\quad+\sum_{i\in\Delta_1}(\frac{z_{31,i}z_{2,i}-z_{21,i}^2}{z_{2,i}^2})\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\\
&\quad+\sum_{i\in\Delta_0}\frac{z_{2,i}}{1+z_{1,i}}\frac{\partial\log S^2(t_i,\boldsymbol{\gamma})}{\partial\gamma_h\partial\gamma_{h'}}+\sum_{i\in\Delta_0}\frac{z_{21,i}z_{1,i}+z_{21,i}-z_{2,i}^2}{(1+z_{1,i})^2}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_h}\frac{\partial\log S(t_i,\boldsymbol{\gamma})}{\partial\gamma_{h'}}\},
\end{aligned}
$$

for $l,l'=0,\ldots,p$, $x_{i0}\equiv 1$, $h,h'=0,1,j^*$, $j^*=21,22,\ldots,2p$, $i=1,\ldots,n$.

Bernoulli cure rate model: The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_l} &= \sum_{i\in\Delta_1}x_{il}-\sum_{i\in\Delta*}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}+\sum_{i\in\Delta_0}w_ix_{il},\\
\frac{\partial l}{\partial\gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial\log f(t_i;\boldsymbol{x}_i,\boldsymbol{\gamma})}{\partial\gamma_j}+\sum_{i\in\Delta_0}w_i\frac{\partial\log S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j},
\end{aligned}
$$

where $w_i=\frac{e^{\boldsymbol{x}'_i\boldsymbol{\beta}}S(t_i;\boldsymbol{\gamma})}{1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}S(t_i;\boldsymbol{\gamma})}$.

Hence, the components of the observed information matrix are

$$
\begin{aligned}
-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{-\sum_{i\in\Delta*}\frac{x_{il}x_{il'}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}})^2}+\sum_{i\in\Delta_0}x_{il}x_{il'}w_i(1-w_i)\},\\
-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} &= -\{\sum_{i\in\Delta_0}x_{il}w_i(1-w_i)\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\},\\
-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} &= -\{\sum_{i\in\Delta_1}\frac{\partial\log f(t_i;\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_0}w_i\frac{\partial\log S^2(t_i;\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_0}w_i(1-w_i)\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\frac{\partial\log S(t_i;\boldsymbol{\gamma})}{\partial\gamma_{j'}}\},
\end{aligned}
$$

for $l,l'=0,\ldots,p$, $x_{i0}\equiv 1$, $h,h'=0,1,j^*$, $j^*=21,22,\ldots,2p$, $i=1,\ldots,n$.

Poisson cure rate model: The score functions are

$$
\begin{aligned}
\frac{\partial l}{\partial\beta_l} &= \sum_{i\in\Delta_1}\frac{1}{\log(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}})}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}}+\sum_{i\in\Delta*}\frac{x_{il}e^{\boldsymbol{x}'_i\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma})-1)}{1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}},\\
\frac{\partial l}{\partial\gamma_j} &= \sum_{i\in\Delta_1}\frac{\partial\log f(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}+\sum_{i\in\Delta*}\frac{\partial S(t_i;\boldsymbol{\gamma})}{\partial\gamma_j}\log(1+e^{\boldsymbol{x}'_i\boldsymbol{\beta}}).
\end{aligned}
$$

Hence, the components of the observed information matrix are

$$-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} = -\{\sum_{i\in\Delta_1}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})-e^{\boldsymbol{x}_i'\boldsymbol{\beta}})}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2[\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})]^2}+\sum_{i\in\Delta_*}\frac{x_{il}x_{il}'e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{[1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}]^2}(S(t_i;\boldsymbol{\gamma})-1)\},$$

$$-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} = -\{\sum_{i\in\Delta_*}\frac{\partial S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\},$$

$$-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} = -\{\sum_{i\in\Delta_1}\frac{\partial\log f^2(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}+\sum_{i\in\Delta_*}\frac{\partial^2 S(t_i;\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_j'}\log(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})\},$$

for $l,l'=0,\ldots,p$, $x_{i0}\equiv 1$, $h,h'=0,1,j^*$, $j^*=21,22,\ldots,2p$, $i=1,\ldots,n$.

Geometric cure rate model: The score functions are

$$\frac{\partial l}{\partial\beta_l} = \sum_{i\in\Delta_1}x_{il}-2\sum_{i\in\Delta_1}\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}-\sum_{i\in\Delta_0}\frac{x_{il}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}+\sum_{i\in\Delta_0}\frac{S(t_i;\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}(x_{il}-\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}),$$

$$\frac{\partial l}{\partial\gamma_j} = \sum_{i\in\Delta_1}(\frac{\partial\log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j}-\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j})-\sum_{i\in\Delta_*}\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j}.$$

Hence, the components of the observed information matrix are

$$\begin{aligned}-\frac{\partial^2 l}{\partial\beta_l\partial\beta_{l'}} &= -\{-2\sum_{i\in\Delta_1}\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}}+\sum_{i\in\Delta_0}\frac{x_{il}x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}(S(t_i;\boldsymbol{\gamma})-1)\\ &\quad-\sum_{i\in\Delta_0}\frac{S(t_i;\boldsymbol{\gamma})e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}{1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}}}\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_{l'}}-\sum_{i\in\Delta_0}\frac{x_{il'}e^{\boldsymbol{x}_i'\boldsymbol{\beta}}S(t_i;\boldsymbol{\gamma})}{(1+e^{\boldsymbol{x}_i'\boldsymbol{\beta}})^2}\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l}\},\end{aligned}$$

$$-\frac{\partial^2 l}{\partial\beta_l\partial\gamma_j} = -\{-\sum_{i\in\Delta_1}\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j}-\sum_{i\in\Delta_*}\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j}\},$$

$$-\frac{\partial^2 l}{\partial\gamma_j\partial\gamma_{j'}} = -\{\sum_{i\in\Delta_1}(\frac{\partial\log f(t_i,\boldsymbol{x}_{ic},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_{j'}}-\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}})-\sum_{i\in\Delta_*}\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}\},$$

where $R_G(t_i,\boldsymbol{\theta})=1-e^{\boldsymbol{x}_i'\boldsymbol{\beta}}(S(t_i;\boldsymbol{\gamma})-1)$.

$$\frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\gamma_j} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j},\quad \frac{\partial\log R_G(t_i,\boldsymbol{\theta})}{\partial\beta_l} = \frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})},$$

$$\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}} = \frac{-e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})}\frac{\partial S^2(t_i,\boldsymbol{\theta})}{\partial\gamma_j\partial\gamma_{j'}}-\frac{e^{2\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_{j'}},$$

$$\frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\beta_l} = \frac{x_{il}x_{il'}e^{\boldsymbol{x}'\boldsymbol{\beta}}(1-S(t_i,\boldsymbol{\theta}))}{R_G(t_i,\boldsymbol{\theta})^2},\quad \frac{\partial\log R_G^2(t_i,\boldsymbol{\theta})}{\partial\beta_l\partial\gamma_j} = -\frac{x_{il}e^{\boldsymbol{x}'\boldsymbol{\beta}}}{R_G(t_i,\boldsymbol{\theta})^2}\frac{\partial S(t_i,\boldsymbol{\theta})}{\partial\gamma_j},$$

for $l,l'=0,\ldots,p$, $x_{i0}\equiv 1$, $h,h'=0,1,j^*$, $j^*=21,22,\ldots,2p$, $i=1,\ldots,n$.

References

1. N Balakrishnan, S Barui, and FS Milienos. Proportional hazards under Conway–Maxwell-Poisson cure rate model and associated inference. *Statistical Methods in Medical Research*, 26(5): 2055–2077, 2017.
2. N Balakrishnan and S Pal. EM algorithm-based likelihood estimation for some cure rate models. *Journal of Statistical Theory and Practice*, 6(4):698–724, 2012.
3. N Balakrishnan and S Pal. Lognormal lifetimes and likelihood-based inference for flexible cure rate models based on COM-Poisson family. *Computational Statistics & Data Analysis*, 67:41–67, 2013.

4. N Balakrishnan and S Pal. COM-Poisson cure rate models and associated likelihood-based inference with exponential and weibull lifetimes. In: *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference.* Frenkel IB, Karagrigoriou A, Lisnianski A et al. (eds.) John Wiley & Sons: Chichester, UK, pp. 308–348, 2014.
5. N Balakrishnan and S Pal. An EM algorithm for the estimation of parameters of a flexible cure rate model with generalized gamma lifetime and model discrimination using likelihood-and information-based methods. *Computational Statistics*, 30(1):151–189, 2015.
6. N Balakrishnan and S Pal. Expectation maximization-based likelihood inference for flexible cure rate models with Weibull lifetimes. *Statistical Methods in Medical Research*, 25(4):1535–1563, 2016.
7. J Berkson and R P Gage. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47(259):501–515, 1952.
8. J W Boag. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B*, 11(1):15–53, 1949.
9. R W Conway and W L Maxwell. A queuing model with state dependent service rates. *Journal of Industrial Engineering*, 12(2):132–136, 1962.
10. Y Gu, D Sinha, and S Banerjee. Analysis of cure rate survival data under proportional odds model. *Lifetime Data Analysis*, 17(1):123–134, 2011.
11. J G Ibrahim, M-H Chen, and D Sinha. *Bayesian Survival Analysis*. Springer: New York, 2001.
12. J B Kadane, G Shmueli, T P Minka, S Borle, P Boatwright, et al. Conjugate analysis of the Conway-Maxwell-Poisson distribution. *Bayesian Analysis*, 1(2):363–374, 2006.
13. J M Kirkwood, J G Ibrahim, V K Sondak, J Richards, L E Flaherty, M S Ernstoff, T J Smith, U Rao, M Steele, and R H Blum. High- and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial e1690/s9111/c9190. *Journal of Clinical Oncology*, 18(12):2444–2458, 2000.
14. C-S Li and J MG Taylor. A semi-parametric accelerated failure time cure model. *Statistics in Medicine*, 21(21):3235–3247, 2002.
15. X Liu, Y Peng, D Tu, and H Liang. Variable selection in semiparametric cure models based on penalized likelihood, with application to breast cancer clinical trials. *Statistics in Medicine*, 31(24):2882–2891, 2012.
16. T A Louis. Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society, Series B , (Methodological)*, pp. 226–233, 1982.
17. W Lu. Efficient estimation for an accelerated failure time model with a cure fraction. *Statistica Sinica*, 20:661-674, 2010.
18. M Mao and J-L Wang. Semiparametric efficient estimation for a class of generalized proportional odds cure models. *Journal of the American Statistical Association*, 105(489):302–311, 2010.
19. G McLachlan and T Krishnan. *The EM Algorithm and Extensions*. John Wiley & Sons, Hoboken, New Jersey, 2007.
20. Y Peng and K BG Dear. A nonparametric mixture model for cure rate estimation. *Biometrics*, 56(1):237–243, 2000.
21. J Rodrigues, M de Castro, V G Cancho, and N Balakrishnan. COM–Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference*, 139(10):3605–3611, 2009.
22. J Rodrigues, V G Cancho, M A Castro, F Louzada. On the unification of the long-term survival models. *Statistics & Probability Letters*, 79:753-759, 2009 .
23. G Shmueli, T P Minka, J B Kadane, S Borle, and P Boatwright. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 54(1):127–142, 2005.
24. J P Sy and J MG Taylor. Estimation in a cox proportional hazards cure model. *Biometrics*, 56(1):227–236, 2000.
25. J Zhang and Y Peng. A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine*, 26(16):3157–3171, 2007.
26. J Zhang and Y Peng. Accelerated hazards mixture cure model. *Lifetime Data Analysis*, 15(4):455-467, 2009.