Journal of Data Science **12**(2014), 107-136

A New Class of Survival Regression Models with Cure Fraction

Edwin M. M. Ortega[1]*, Gladys D. C. Barriga[2], Elizabeth M. Hashimoto[1],
Vicente G. Cancho[1] and Gauss M. Cordeiro[3]
[1] *Universidade de São Paulo,* [2] *Universidade Estadual Paulista "Júlio de Mesquita Filho" and* [3] *Universidade Federal de Pernambuco*

Abstract: In this paper, we propose a flexible cure rate survival model by assuming that the number of competing causes of the event of interest follows the negative binomial distribution and the time to event follows a generalized gamma distribution. We define the negative binomial-generalized gamma distribution, which can be used to model survival data. The new model includes as special cases some of the well-known cure rate models discussed in the literature. We consider a frequentist analysis and nonparametric bootstrap for parameter estimation of a negative binomial-generalized gamma regression model with cure rate. Then, we derive the appropriate matrices for assessing local influence on the parameter estimates under different perturbation schemes and present some ways to perform global influence analysis. Finally, we analyze a real data set from the medical area.

Key words: Cure fraction models, generalized gamma distribution, lifetime data, negative binomial distribution, sensitivity analysis.

1. Introduction

Models for survival data with a cure fraction (also known as cure rate models or long-term survival models) play an important role in reliability and survival analysis. Cure rate models cover situations where there are sampling units not susceptible to the occurrence of the event of interest. The proportion of such units is termed the cured fraction. These models have become very popular due to significant progress in treatment therapies leading to enhanced cure rates. The proportion of these "cured" units is termed the cure fraction. In clinical studies, the event of interest can be the death of a patient (which can happen due to different competing causes) or a tumor recurrence (which can be attributed to metastasis-component tumor cells left active after an initial treatment). Models

*Corresponding author.

to accommodate a cure fraction have been widely developed. Perhaps the most popular type of cure rate model is the mixture distribution introduced by Boag (1949) and Berkson and Gage (1952). Further, mixture models are based on the assumption that only a cause is responsible for the occurrence of the event of interest. However, in clinical studies, the patient's death, which is the event of interest, may happen due to different latent competing causes, in the sense that there is no information about which cause was responsible for the individual death. A tumor recurrence can be attributed to metastasis-component tumor cells left active after initial treatment. A metastasis-component tumor cell is a tumor cell with potential to metastasize (Yakovlev and Tsodikov, 1996). The literature on distributions which accommodates different latent competing causes is rich and growing rapidly. The book by Ibrahim *et al.* (2001), the review paper by Tsodikov *et al.* (2003) and the works by Cooner *et al.* (2007) and Rodrigues *et al.* (2009) can be mentioned as key references. In this paper, we propose a new model called the negative binomial-generalized gamma ("NBGG" for short) cure rate model, conceived inside a latent competing causes scenario, where the causes are modeled by the negative binomial (NB) distribution and the time for the corresponding cause to produce the event of interest (death or tumor recurrence) is modeled by the generalized gamma (GG) distribution. There is no information about which cause was responsible for the individual death or tumor recurrence, but only the minimum lifetime value among all causes is observed and a part of the population is not susceptible to the event of interest. The proposed model includes the traditional cure models as special cases (Boag, 1949; Berkson and Gage, 1952; Yakovlev and Tsodikov, 1996; Ortega *et al.*, 2009b; Cancho *et al.*, 2011). Also, we examine statistical inference aspects and formulate the NBGG model with covariates.

After fitting the model, it is important to check its assumptions and conduct robustness studies to detect possible influential or extreme observations that can cause distortions in the results of the analysis. In this paper, we discuss the influence diagnostic based on case-deletion, in which the influence of the i-th observation on the parameter estimates is studied by removing this observation from the analysis. We propose diagnostic measures based on case-deletion for the NBGG regression model with cure rate in order to determine which subjects might be influential in the analysis.

Nevertheless, when case-deletion is used, all information from a single subject is deleted at once and therefore it is hard to say whether that subject has some influence on a specific aspect of the model. A solution for this problem can be found in the local influence approach, where we again investigate how the results of the analysis change under small perturbations in the model or data. Cook (1986) proposed a general framework to detect the influence of the observations to indicate how sensitive the analysis is when small perturbations in the data or model occur. Several authors have applied the local influence methodology in regression analysis with censoring. Ortega *et al.* (2003) considered the problem of assessing local influence in generalized log-gamma regression models with censored observations; Silva *et al.* (2008) investigated local influence in log-Burr XII regression models with censored data; Fachini *et al.* (2008) adapted local influence methods to polyhazard models under the presence of covariates; Cancho *et al.* (2009) derived curvature calculations under various perturbation schemes in log-exponentiated Weibull regression models with cure rate; and Hashimoto *et al.* (2010) calculated the appropriate matrices for assessing local influences on the parameter estimates under different perturbation schemes in the log-exponentiated Weibull regression model for interval-censored data. We propose a similar method to detect influential subjects in the NBGG regression model with cure rate.

The plan of the next sections of the paper is as follows. Section 2 is dedicated to model formulation. Parameter inference is discussed in Section 3. In Section 4, we use nonparametric bootstrap for parameter estimation of the NBGG regression model with cure rate. In Section 5, we obtain the normal curvatures of local influence and derive the global influence under some usual perturbations. The results of an application to a real data set are reported in Section 6. Section 7 provides concluding remarks.

2. The Model

For an individual in the population, let N denote the unobservable number of causes of the event of interest for this individual. The time for the j-th cause to produce the event of interest is denoted by Z_j, $j = 1, \cdots, N$. We assume that, conditional on N, the Z_j are i.i.d. random variables with cumulative distribution function (cdf) $F(z)$ and survival function $S(z) = 1 - F(z)$. We also assume that N is independent of $Z_1, Z_2, \cdots$. The observable time to event is defined by $T = \min\{Z_1, \cdots, Z_N\}$, if $N \geq 1$ and $T = \infty$ if $N = 0$, with $P(T = \infty | N = 0) = 1$. Under this setup, the survival function for the population is given by

$$S_{\text{pop}}(t) = P(N = 0) + P(Z_1 > t, \cdots, Z_N > t | N \geq 1)\, P(N \geq 1). \qquad (1)$$

Tsodikov *et al.* (2003), among others, demonstrated that $S_{\text{pop}}(t) = A_p[S(t)]$, where $A_p(\cdot)$ is the probability generating function (pgf) of the number of competing causes (N). de Castro *et al.* (2010) considered that the number of competing causes follows the NB distribution with parameters $\theta > 0$ and $\alpha > -1/\theta$ (Piegorsch, 1990; Saha and Paul, 2005), with probability mass function

$$p_m = P(N = m) = \frac{\Gamma(\alpha^{-1} + m)}{\Gamma(\alpha^{-1})\, m!} \left(\frac{\alpha\,\theta}{1 + \alpha\,\theta}\right)^m (1 + \alpha\theta)^{-1/\alpha},$$

where $m = 0, 1, \cdots$ and $\Gamma(k) = \int_0^\infty w^{k-1} e^{-w} dw$ is the gamma function. Then,

$$E(N) = \theta \quad \text{and} \quad \mathrm{Var}(N) = \theta + \alpha\, \theta^2. \tag{2}$$

The pgf of N is given by $A_p(s) = \sum_{m=0}^{\infty} p_m\, s^m = [1 + \alpha\, \theta\, (1 - s)]^{-1/\alpha}$, $0 \le s \le 1$, so that the survival function for the population becomes

$$S_{\text{pop}}(t) = A_p\big[S(t)\big] = [1 + \alpha\theta F(t)]^{-1/\alpha}. \tag{3}$$

The cured fraction is given by $p_0 = (1+\alpha\theta)^{-1/\alpha}$. The probability density function (pdf) associated to (3) reduces to

$$f_{\text{pop}}(t) = [1 + \alpha\theta F(t)]^{-(1/\alpha)-1}\, \theta f(t),$$

where $f(t) = -S'(t)$ denotes the (proper) density function of the time to event Z in (1) with hazard rate function (hrf)

$$h_{\text{pop}}(t) = \theta\, f(t)[1 + \alpha\theta F(t)]^{-1}.$$

We note that $f_{\text{pop}}(t)$ and $h_{\text{pop}}(t)$ are improper functions, since $S_{\text{pop}}(t)$ is not a proper survival function. The probability distribution in (2) is a flexible model in the sense it provides nice links between the binomial and Poisson distributions, that is, for $\alpha = -1/r$ (r integer) we have that $N \sim \text{binomial}(r, \theta/r)$ provided that $0 \le \theta/r \le 1$ (Piegorsch, 1990). When $\alpha \to 0$, we have $N \sim \text{Poisson}(\theta)$. These results imply that for $\alpha = -1\, (r = 1)$, we have the classical mixture cure model (Boag, 1949; Berkson and Gage, 1952) with a high level of under-dispersion and for $\alpha \to 0$, we obtain the promotion time cure model (Yakovlev and Tsodikov, 1996). From (2) it follows that the variance of the number of competing causes under the NB model is flexible. If $-1/\theta \le \alpha < 0$, there is under-dispersion from the Poisson model. We illustrate this point with the mixture cure model. On the other hand, if $\alpha > 0$, the counts are over-dispersed.

As pointed out by Tournoud and Ecochard (2008), the parameters of the NB model have biological interpretations. In (2), θ is the mean number of competing causes, whereas α accounts for the inter-individual variance of the number of causes. Additionally, the NB distribution enables de Castro *et al.* (2010) to provide a probabilistic justification for the transformation introduced by Yin and Ibrahim (2005).

To make the model in (3) more flexible, we first assume that the times Z's to the event of interest have a GG distribution with pdf given by

$$f(z; \tau, \beta, k) = \frac{\beta}{\tau\Gamma(k)} \Big(\frac{z}{\tau}\Big)^{k\beta-1} \exp\Big[-\Big(\frac{z}{\tau}\Big)^{\beta}\Big], \tag{4}$$

where $\beta > 0$ and $k > 0$ are shape parameters and $\tau > 0$ is a scale parameter. This distribution is known as the three-parameter GG distribution (Stacy, 1962).

The times Z's to the event of interest, are unobservable which do not allow an appropriate choice of a parametric distribution for the Z's. Then, it is reasonable to consider the GG distribution, since this model includes all four most common types of the hrf: monotonically increasing, decreasing, bathtub and unimodal hazard rates (Cox *et al.*, 2007).

This version of the GG distribution (4), however, presents convergence problems that severely limit its usefulness (see for example, Lawless, 2003). To avoid this problem, consider the following re-parametrization: $\tau = \exp[\mu - \log(q^{-2})/\beta]$, $\beta = q/\sigma$ and $k = q^{-2}$. Then, the re-parameterized pdf is given by

$$f(z;\boldsymbol{\gamma}) = \begin{cases} \frac{|q|}{\sigma\Gamma(q^{-2})z}(q^{-2})^{q^{-2}}\exp\Big\{q^{-1}\left[\frac{\log(z)-\mu}{\sigma}\right] - q^{-2}\exp\Big\{q\left[\frac{\log(z)-\mu}{\sigma}\right]\Big\}\Big\}, & \\ & \text{if} \quad q \neq 0, \\ \frac{1}{\sqrt{2\pi}\sigma z}\exp\Big\{-\frac{1}{2}\left[\frac{\log(z)-\mu}{\sigma}\right]^2\Big\}, & \text{if} \quad q = 0, \end{cases} \tag{5}$$

where $\boldsymbol{\gamma} = (\mu, \sigma)^\top$ for $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < q < \infty$. The GG distribution with pdf (5) is known as the family of extended generalized gamma (EGG) models, described in details by Kalbfleisch and Prentice (2002) and Lawless (2003), and includes as special cases the exponential, Weibull, reciprocal Weibull, log-normal and gamma distributions.

Substituting the reparametrized GG distribution in (3), we obtain the NBGG cure rate model

$$S_{\text{pop}}(t;\boldsymbol{\gamma}) = \begin{cases} \left\{1 + \alpha\theta\left(1 - \Gamma\left\{q^{-2}\exp\left[q\left(\frac{\log(t)-\mu}{\sigma}\right)\right]; q^{-2}\right\}\right)\right\}^{-1/\alpha}, & \\ & \text{if} \quad q < 0; \\ \left\{1 + \alpha\theta\left(\Phi\left[\frac{\log(t)-\mu}{\sigma}\right]\right)\right\}^{-1/\alpha}, & \text{if} \quad q = 0, \\ \left\{1 + \alpha\theta\left(\Gamma\left\{q^{-2}\exp\left[q\left(\frac{\log(t)-\mu}{\sigma}\right)\right]; q^{-2}\right\}\right)\right\}^{-1/\alpha}, & \\ & \text{if} \quad q > 0, \end{cases} \tag{6}$$

where $\Phi(z)$ denotes the standard normal cumulative distribution and $\Gamma(k;z)$ is the incomplete gamma function ratio, which is given by

$$\Gamma(k;z) = \frac{1}{\Gamma(k)}\int_0^z x^{k-1}e^{-x}dx.$$

In Table 1, we list some characteristics of the NBGG model with cure rate.

Note that for $q = 1$ and $\beta = 1/\sigma$, the special case corresponds to the long-term survival model with cured fraction (de Castro *et al.*, 2009).

Table 1: Characteristics of the NBGG model with cure rate in a competitive-risk structure

Distribution Z and N	Mixture Cure Model $\alpha = -1$	Promotion Time Model $\alpha \to 0$	Alternative Cure Model $\alpha > 0$
$q = 1$, $\beta = 1/\sigma$ $\mu = \log(\tau)$	Mixture Weibull	Promotion Weibull	Alternative Weibull
$q = 1$, $\sigma = 1$ $\mu = \log(\tau)$	Mixture Exponential	Promotion Exponential	Alternative Exponential
$q = 0$	Mixture Log-normal	Promotion Log-normal	Alternative Log-normal
$q = \sigma$ $\mu = \log(\tau)$	Mixture Gamma	Promotion Gamma	Alternative Gamma
$q = -1$, $\beta = 1/\sigma$ $\mu = \log(\tau)$	Mixture Reciprocal Weibull	Promotion Reciprocal Weibull	Alternative Reciprocal Weibull

The survival function for the non-cured population, say $S_{\mathrm{NBGG}}(y) = P(Y > y|N \geq 1)$, is given by

$$S_{\mathrm{NBGG}}(t) = P(T > t|N \geq 1) = \frac{[1 + \alpha\theta F(t)]^{-1/\alpha} - (1 + \alpha\theta)^{-1/\alpha}}{1 - (1 + \alpha\theta)^{-1/\alpha}}, \; t > 0, \quad (7)$$

where $F(\cdot)$ is the GG cdf. We note that $S_{\mathrm{NBGG}}(0) = 1$ and $S_{\mathrm{NBGG}}(\infty) = 0$, so that it is a proper survival function. The pdf for the non-cured population (called the NBGG density function) is given by

$$f_{\mathrm{NBGG}}(t) = \frac{\theta\, f(t)\, [1 + \alpha\theta F(t)]^{-(1/\alpha+1)}}{1 - (1 + \alpha\theta)^{-1/\alpha}}, \; t > 0, \quad (8)$$

where $f(\cdot)$ is the GG pdf. From (8), we note that the parameter σ controls the scale of the distribution while the parameters α, θ and q control its shape. As $\alpha = -1$, the NBGG distribution reduces to the GG distribution. Figure 1 displays some plots of the NBGG density function for some fixed values of α and θ. These plots indicate that this distribution is very flexible and that the values of α and θ have a substantial effect on its skewness and kurtosis.

From (7) and (8), it is easy to verify that the hrf for the non-cured population is given by

$$h_{\mathrm{NBGG}}(t) = \frac{\theta h(t) S(t)\, [1 + \alpha\theta F(t)]^{-(1/\alpha+1)}}{[1 + \alpha\theta F(t)]^{-1/\alpha} - (1 + \alpha\theta)^{-1/\alpha}}, \;\; t > 0, \quad (9)$$

where $h(t)$ and $S(t)$ are the hazard rate and survival functions of the GG distribution, respectively. Based on (9), $h_{\mathrm{NBGG}}(y)/h(y)$ is increasing in $y > 0$ for $\theta > 0$ and $\alpha > -1/\theta$. Further, $h(t) \leq h_{\mathrm{NBGG}}(t)$ and then the limit behavior of

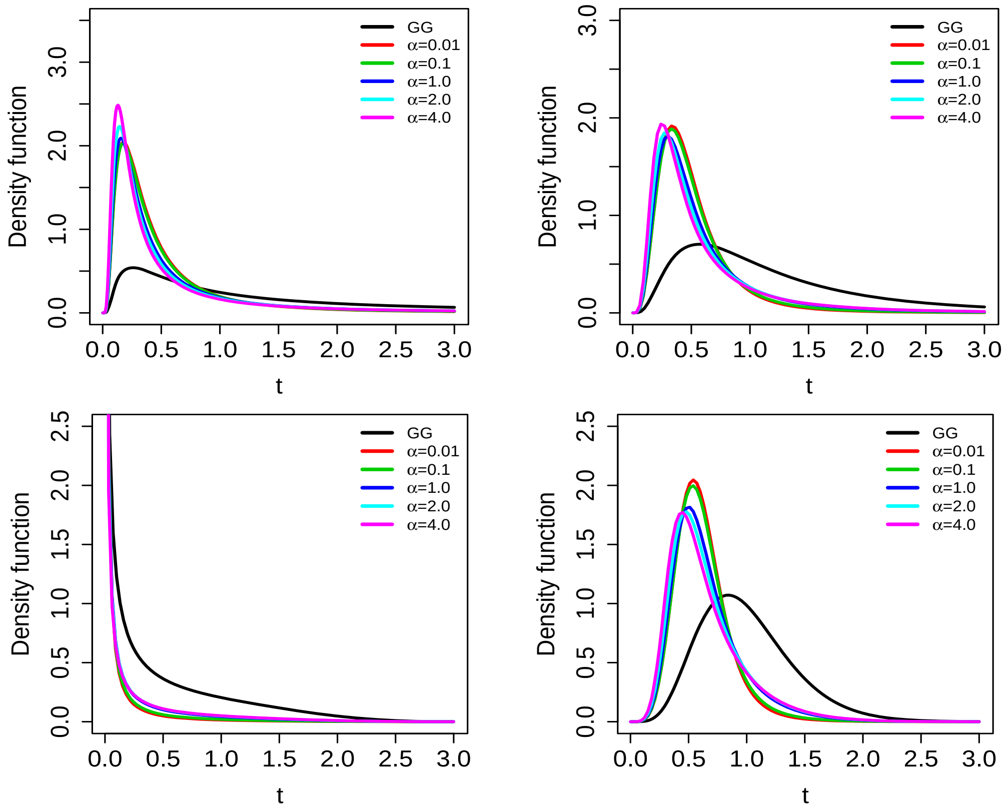

Figure 1: The NBGG pdf for some parameter values. The parameters are fixed at $\alpha = -1, 0.01, 0.1, 1, 2, 4$, $\theta = 5$, $\mu = 0$ and $\sigma = 1.5$, $q = 1$ (top left panel); $\sigma = 0.4$, $q = 0.4$ (top right panel); $\sigma = 1$, $q = 4$ (bottom left panel) and $\sigma = 0.75$, $q = 0$ (bottom right panel)

the hrf of the NBGG distribution is the same as the behavior of the GG hrf. Figure 2 displays some shapes of the NBGG hrf for some parameter values. When $\alpha \to 0$, the NBGG hrf approaches the Poisson generalized gamma hrf and, for $\alpha = 1$, it reduces to the geometric generalized gamma hrf (Ortega *et al.*, 2011). In Figure 2, we plot this hrf for some parameter values. There is a mathematical relationship between the model (6) and the mixture cure rate model (Boag, 1949; Berkson and Gage, 1952). We can write

$$S_{\text{pop}}(t) = (1 + \alpha\theta)^{-1/\alpha} + \left[1 - (1 + \alpha\theta)^{-1/\alpha}\right] S_{\text{NBGG}}(t),$$

where $S_{\text{NBGG}}(t)$ is given by (7). Thus, $S_{\text{pop}}(t)$ is a mixture cure rate model with cure rate equal to $p_0 = (1 + \alpha\theta)^{-1/\alpha}$ and survival function $S_{\text{NBGG}}(t)$ for the non-cured population. This results imply that every mixture cure rate model corresponds to some model of the form (6) for any α, θ and $F(\cdot)$ (this result holds for any distribution function).

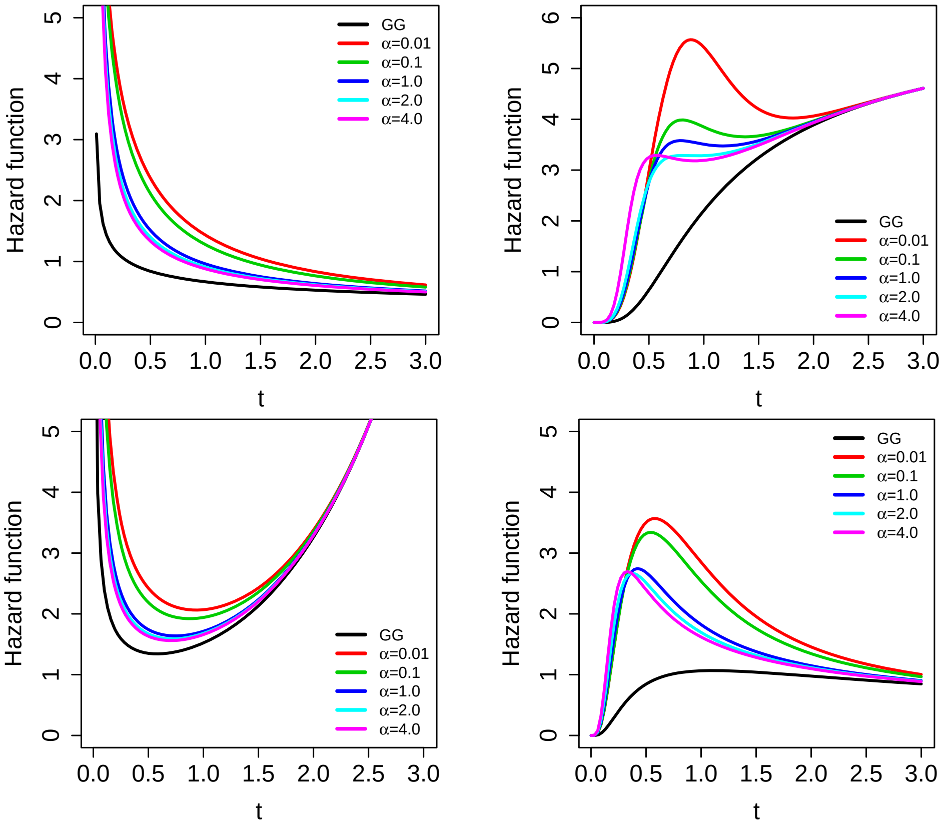

Figure 2: The hrf of the NBGG distribution. The parameters are fixed at $\alpha = -1, 0.01, 0.1, 1, 2, 4$, $\theta = 5$, $\mu = 0$ and $\sigma = 1.5$, $q = 1$ (top left panel); $\sigma = 0.4$, $q = 0.4$ (top right panel); $\sigma = 1$, $q = 4$ (bottom left panel) and $\sigma = 0.75$, $q = 0$ (bottom right panel)

3. Inference

Hereafter, we suppose that the time to the event is not completely observed and may be subjected to right censoring. Let C_i denote the censoring time. We observe that $Y_i = \min\{T_i, C_i\}$ and $\delta_i = I(T_i \leq C_i)$ are such that $\delta_i = 1$ if T_i is a time to event and $\delta_i = 0$ if it is right censored, $i = 1, \cdots, n$.

Following de Castro *et al.* (2009), we consider the Fisher parametrization of the NB distribution (Ross and Preece, 1985) for $\alpha \geq -1$. We define $\theta = (p_0^{-\alpha} - 1)/\alpha$, if $\alpha \neq 0$, and $\theta = -\log(p_0)$, if $\alpha = 0$. We incorporate covariates for the parametric cure rate model (6) through the cure parameter, p_0. When covariates are included, we have a different cure rate parameter, p_{0i}, for each subject, $i = 1, \cdots, n$. The cured fraction is linked to covariates $\boldsymbol{x}_i = (x_{i1}, \cdots, x_{ip})^\top$ by the logistic link, i.e.,

$$\log\left(\frac{p_{0i}}{1-p_{0i}}\right) = \boldsymbol{x}_i^\top\boldsymbol{\beta} \quad \text{or} \quad p_{0i} = \frac{\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})}, \tag{10}$$

where $\boldsymbol{\beta}$ stands for the vector of regression coefficients. Notice that regardless of the specific model (which depends on α), the covariates are associated with the cured fraction through a unique expression. We recall that covariates are traditionally used to model the expectation of the number of competing causes. For instance, in the promotion time cure model, we have $\theta_i = E(N_i) = \exp(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\beta})$ and $p_{0i} = e^{-\theta_i}$, so that $\log\{-\log(p_{0i})\} = \boldsymbol{x}_i^\top\boldsymbol{\beta}$. The connection between the cured fraction and the covariates is much more cumbersome in this expression than in the logistic link. Recently, de Castro *et al.* (2009) showed that this parametrization is identifiable.

From (2), $\mathrm{Var}(N_i) = E(N_i)\,p_{0i}^{-\alpha}$. Thus, extra variability in the number of competing causes due to omitted covariates is governed by the dispersion parameter α. Under this relation, the improper functions (6) can be rewritten as

$$S_{\text{pop}}(t_i;\boldsymbol{\gamma}) = \begin{cases} \{1+(p_{0i}^{-\alpha}-1)F(t_i;\boldsymbol{\gamma})\}^{-1/\alpha}, & \text{if} \quad \alpha \neq 0; \\ p_{0i}^{F(t_i;\boldsymbol{\gamma})}, & \text{if} \quad \alpha = 0; \end{cases} \tag{11}$$

where $F(t_i;\boldsymbol{\gamma})$ is the GG distribution. The density function corresponding to this model is given by

$$f_{\text{pop}}(t_i;\boldsymbol{\beta},\boldsymbol{\gamma},\alpha) = \begin{cases} \{1+(p_{0i}^{-\alpha}-1)F(t_i;\boldsymbol{\gamma})\}^{-1/\alpha-1}\,(\frac{p_{0i}^{-\alpha}-1}{\alpha})\,f(t_i;\boldsymbol{\gamma}), & \text{if} \quad \alpha \neq 0; \\ -\log(p_{0i})\,p_{0i}^{F(t_i;\boldsymbol{\gamma})}f(t_i;\boldsymbol{\gamma}), & \text{if} \quad \alpha = 0. \end{cases} \tag{12}$$

We refer to model (11) as the NBGG regression model with cure rate in a competitive-risk structure. Following the steps in the proof of Theorem 6.2 in Tournoud and Ecochard (2008), we conclude that if the covariates are linked to the parameter α too, identifiability is preserved.

Based on the NB distribution with (10), (11) and (12), we can write the likelihood of $\boldsymbol{\vartheta} = (\alpha,\boldsymbol{\beta}^\top,\boldsymbol{\gamma})^{\mathrm{T}}$ under non-informative censoring as

$$L(\boldsymbol{\vartheta};\boldsymbol{\mathcal{D}}) \propto \begin{cases} \prod_{i=1}^n \left\{\frac{p_{0i}^{-\alpha}-1}{\alpha}f(y_i;\boldsymbol{\gamma})\right\}^{\delta_i}\{1+(p_{0i}^{-\alpha}-1)F(y_i;\boldsymbol{\gamma})\}^{-\delta_i-1/\alpha}, & \\ & \text{if} \quad \alpha \neq 0; \\ \prod_{i=1}^n \{-\log(p_{0i})f(y_i;\boldsymbol{\gamma})\}^{\delta_i}\,p_{0i}^{F(y_i;\boldsymbol{\gamma})}, & \text{if} \quad \alpha = 0; \end{cases} \tag{13}$$

where $\boldsymbol{\vartheta} = (\alpha,\boldsymbol{\beta}^\top,\boldsymbol{\gamma}^\top)^\top$, $\boldsymbol{\mathcal{D}} = (n,\boldsymbol{y},\boldsymbol{\delta},\boldsymbol{x})$, and $\boldsymbol{x} = (\boldsymbol{x}_1^\top,\cdots,\boldsymbol{x}_n^\top)$. The maximization of (13) follows the same two steps for obtaining the maximum likelihood estimates (MLEs) of $\boldsymbol{\vartheta}$ under the uncensored case. Since in general it is reasonable to expect that the shape parameter q belongs to the interval [-3, 3] (Lawless, 2003), in the first step of the iterative process we set different q values in this interval. Then, we obtain the MLEs $\widetilde{\alpha}(q)$, $\widetilde{\boldsymbol{\beta}}(q)$ and $\widetilde{\gamma}(q)$, and determine the maximized log-likelihood function $L_{\max}(q)$. In this step, we use the MaxBFGS routine in the matrix programming language Ox (see, for instance, Doornik, 2002). In the second step, the log-likelihood $L_{\max}(q)$ is maximized, and then $\widehat{q}$ is obtained. The MLEs of α, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are $\widehat{\alpha} = \widetilde{\alpha}(\widehat{q})$, $\widehat{\boldsymbol{\beta}} = \widetilde{\boldsymbol{\beta}}(\widehat{q})$ and $\widehat{\gamma} = \widetilde{\gamma}(\widehat{q})$, respectively. The procedures discussed in this work are developed by assuming q fixed. An important point to take into account in GG models is related to the estimation of the parameter q. Several authors have dealt with this topic (Lawless, 2003; Ortega *et al.*, 2003; Ortega *et al.*, 2009a; among others) and pointed out difficulties to estimate q due to problems of unbounded and local maxima in the likelihood function.

The inference procedures for $\boldsymbol{\vartheta} = (\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ can be based on the asymptotic normal approximation

$$(\widehat{\alpha}, \widehat{\boldsymbol{\beta}}^\top, \widehat{\gamma}^\top)^\top \sim \mathcal{N}_{(p+3)}\Big\{(\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top, -\ddot{\boldsymbol{L}}^{-1}(\boldsymbol{\vartheta})\Big\},$$

where $-\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta}) = \{\partial^2 l(\boldsymbol{\vartheta})/\partial\boldsymbol{\vartheta}\boldsymbol{\vartheta}^T\}$ is the $(p+3) \times (p+3)$ observed information matrix

$$\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta}) = \begin{pmatrix} \boldsymbol{L}_{\alpha\alpha} & \boldsymbol{L}_{\alpha\beta_j} & \boldsymbol{L}_{\alpha\gamma_k} \\ . & \boldsymbol{L}_{\beta_j\beta_{j'}} & \boldsymbol{L}_{\beta_j\gamma_k} \\ . & . & \boldsymbol{L}_{\gamma_k\gamma_{k'}} \end{pmatrix},$$

whose sub-matrices are given in Appendix A.

Besides estimation, hypothesis testing is another key issue. Let $\boldsymbol{\vartheta}_1$ and $\boldsymbol{\vartheta}_2$ be proper disjoint subsets of $\boldsymbol{\vartheta}$. We aim to test $H_0 : \boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_{01}$ against $H_1 : \boldsymbol{\vartheta}_1 \neq \boldsymbol{\vartheta}_{01}$, $\boldsymbol{\vartheta}_2$ unspecified. Let $\hat{\boldsymbol{\vartheta}}_0$ maximize $L(\boldsymbol{\vartheta}; \boldsymbol{\mathcal{D}})$ constrained to H_0 and define the likelihood ratio statistic

$$LR = 2\log\left[\frac{L(\hat{\boldsymbol{\vartheta}}; \boldsymbol{\mathcal{D}})}{L(\hat{\boldsymbol{\vartheta}}_0; \boldsymbol{\mathcal{D}})}\right].$$

Under H_0 and some regularity conditions, the LR statistic converges to the chi-square distribution with $\dim(\boldsymbol{\vartheta}_1)$ degrees of freedom.

4. Bootstrap Re-Sampling Method

The bootstrap re-sampling method was proposed by Efron (1979). This method treats the observed sample as if it represents the population. From the information obtained from such a sample, B bootstrap samples of similar size to that of the observed sample are generated, from which it is possible to estimate various characteristics of the population, such as the mean, variance, percentiles and other quantities.

According to the literature, the re-sampling method can be nonparametric or parametric. In this study, the nonparametric bootstrap method is addressed, according to which the distribution function F can be estimated by an empirical distribution $\widehat{F}$.

Let $\boldsymbol{T} = (T_1, \cdots, T_n)$ be an observed random sample and $\widehat{F}$ be the empirical distribution of $\boldsymbol{T}$. Thus, a bootstrap sample $\boldsymbol{T}^*$ is constructed by re-sampling with replacement of n elements of the sample $\boldsymbol{T}$. For the B bootstrap samples generated, $T_1^*, \cdots, T_B^*$, the bootstrap replication of the parameter of interest for the b-th sample is given by

$$\hat{\boldsymbol{\theta}}_b^* = s(T_b^*),$$

that is, the value of $\hat{\boldsymbol{\theta}}$ for sample T_b^*, $b = 1, \cdots, B$.

The bootstrap estimator of the standard error (Efron and Tibshirani, 1993) is the standard deviation of these bootstrap samples. It is denoted by $\widehat{EP}_B$ and obtained by the following expression:

$$\widehat{EP}_B = \left[\frac{1}{(B-1)} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \bar{\theta}_B \right)^2 \right]^{1/2},$$

where $\bar{\theta}_B = \sum_{b=1}^{B} \hat{\theta}_b^* / B$. Note that B is the number of bootstrap samples generated. According to Efron and Tibshirani (1993), assuming $B \geq 200$, it is generally sufficient to present good results to determine the bootstrap estimates. However, to achieve greater accuracy, a reasonably high B value must be considered. In this study, we consider $B = 3000$ bootstrap samples. We describe the bias corrected and accelerated (BCa) method for constructing approximated confidence intervals based on the bootstrap re-sampling method. For further details on bootstrap intervals, see for example, Efron and Tibshirani (1993), DiCiccio and Efron (1996) and Davison and Hinkley (1997).

5. Diagnostic Analysis

In order to assess the sensitivity of the MLEs, global influence and local influence analysis are now carried out under three perturbation schemes.

5.1 Global Influence

A first tool to perform sensitivity analysis, as stated before, is by means of global influence analysis starting from case-deletion. Case-deletion is a common approach to study the effect of dropping the i-th case from the data set. The case-deletion model for (11) and (12) is given by

$$S_{\text{pop}}(y_l;\boldsymbol{\beta},\boldsymbol{\gamma},\alpha)=\begin{cases}\{1+(p_{0l}^{-\alpha}-1)F(y_l;\boldsymbol{\gamma})\}^{-1/\alpha}, & \text{if} \quad \alpha\neq 0;\\ p_{0l}^{F(y_l;\boldsymbol{\gamma})}, & \text{if} \quad \alpha=0;\end{cases}$$

and

$$f_{\text{pop}}(y_l;\boldsymbol{\beta},\boldsymbol{\gamma},\alpha)=\begin{cases}\{1+(p_{0l}^{-\alpha}-1)F(y_l;\boldsymbol{\gamma})\}^{-1/\alpha-1}\,\frac{p_{0l}^{-\alpha}-1}{\alpha}\,f(y_l;\boldsymbol{\gamma}), & \text{if} \quad \alpha\neq 0;\\ -\log(p_{0l})\,p_{0l}^{F(y_l;\boldsymbol{\gamma})}f(y_l;\boldsymbol{\gamma}), & \text{if} \quad \alpha=0;\end{cases}$$

where $l=1,\cdots,n$, $l\neq i$. In the following discussion, a quantity with subscript "(i)" means the original quantity with the i-th case deleted. For model (13), the log-likelihood function of $\boldsymbol{\vartheta}$ is denoted by $l_{(i)}(\boldsymbol{\vartheta})$. Let $\hat{\boldsymbol{\vartheta}}_{(i)}=(\hat{\alpha}_{(i)},\hat{\boldsymbol{\beta}}_{(i)}^{\top},\hat{\boldsymbol{\gamma}}_{(i)}^{\top})^{\top}$ be the MLE of $\boldsymbol{\vartheta}$ from $l_{(i)}(\boldsymbol{\vartheta})$. To assess the influence of the i-th case on the MLE $\hat{\boldsymbol{\vartheta}}=(\hat{\alpha},\hat{\boldsymbol{\beta}}^{\top},\hat{\boldsymbol{\gamma}}^{\top})^{\top}$, the basic idea is to compare the difference between $\hat{\boldsymbol{\vartheta}}_{(i)}$ and $\hat{\boldsymbol{\vartheta}}$. If deletion of a case seriously influences the estimates, more attention should be paid to that case. Hence, if $\hat{\boldsymbol{\vartheta}}_{(i)}$ is far from $\hat{\boldsymbol{\vartheta}}$, then the i-th case is regarded as an influential observation. A first measure of global influence is defined as the standardized norm of $\hat{\boldsymbol{\vartheta}}_{(i)}-\hat{\boldsymbol{\vartheta}}$ (generalized Cook distance)

$$GD_i(\boldsymbol{\vartheta})=(\hat{\boldsymbol{\vartheta}}_{(i)}-\hat{\boldsymbol{\vartheta}})^{\top}\big[-\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})\big](\hat{\boldsymbol{\vartheta}}_{(i)}-\hat{\boldsymbol{\vartheta}}).$$

Another alternative is to assess $GD_i(\alpha)$, $GD_i(\boldsymbol{\beta})$ or $GD_i(\boldsymbol{\gamma})$, whose values reveal the impact of the i-th case on the estimates of α, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, respectively. Another popular measure of the difference between $\hat{\boldsymbol{\vartheta}}_{(i)}$ and $\hat{\boldsymbol{\vartheta}}$ is the likelihood distance

$$LD_i(\boldsymbol{\vartheta})=2\Big\{l(\hat{\boldsymbol{\vartheta}})-l(\hat{\boldsymbol{\vartheta}}_{(i)})\Big\}.$$

Besides this, we can also compute $\hat{\beta}_j-\hat{\beta}_{j(i)}$ $(j=1,\cdots,p)$ to assess the difference between $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_{(i)}$. Alternative global influence measures are possible. One could think of the behavior of a test statistic, such as the Wald test for explanatory variables or censoring effect under a case-deletion scheme.

Since $\hat{\boldsymbol{\vartheta}}_{(i)}$ is required for every case, a heavy computational burden may be involved. In this case, the following one-step approximation for $\hat{\boldsymbol{\vartheta}}_{(i)}$ can be used to reduce the burden:

$$\hat{\boldsymbol{\vartheta}}_{(i)}\cong\hat{\boldsymbol{\vartheta}}+\ddot{\boldsymbol{L}}(\hat{\boldsymbol{\vartheta}})^{-1}\dot{l}_i(\hat{\boldsymbol{\vartheta}}),$$

where $\dot{l}_{(i)}(\hat{\boldsymbol{\vartheta}})=\partial l_{(i)}(\boldsymbol{\vartheta})/\partial\boldsymbol{\vartheta}$ is evaluated at $\boldsymbol{\vartheta}=\hat{\boldsymbol{\vartheta}}$ (see, for instance, Cook and Weisberg, 1982).

We can also apply the techniques developed by Wang *et al.* (1996) to evaluate how the i-th observation affects a set of parameter estimates. We define the following quantity as the influential estimate (IE) for individual i and for the parameters vectors $\boldsymbol{\vartheta}$, $\boldsymbol{\beta}$ and $(\mu,\sigma,\alpha)^\top$

$$IE(\boldsymbol{\vartheta})_i = \frac{1}{(p+3)}\sum_{v=1}^{p+3}\frac{|\hat{\vartheta}_v - \hat{\vartheta}_{(i)v}|}{SE(\hat{\vartheta}_v)}, \quad IE(\boldsymbol{\beta})_i = \frac{1}{p}\sum_{j=1}^{p}\frac{|\hat{\beta}_j - \hat{\beta}_{(i)j}|}{SE(\hat{\beta}_j)}, \quad \text{and}$$

$$IE(\mu,\sigma,\alpha)_i = \frac{1}{3}\left[\frac{|\hat{\mu}-\hat{\mu}_{(i)}|}{SE(\hat{\mu})} + \frac{|\hat{\sigma}-\hat{\sigma}_{(i)}|}{SE(\hat{\sigma})} + \frac{|\hat{\alpha}-\hat{\alpha}_{(i)}|}{SE(\hat{\alpha})}\right],$$

where $\hat{\vartheta}_v$, $\hat{\vartheta}_{(i)v}$, $\hat{\beta}_j$, $\hat{\beta}_{(i)j}$, $\hat{\mu}$, $\hat{\mu}_{(i)}$, $\hat{\sigma}$, $\hat{\sigma}_{(i)}$, $\hat{\alpha}$ and $\hat{\alpha}_{(i)}$ are the MLEs for the NBGG regression model with cure rate. The $IE(\cdot)$ value for individual i can be interpreted as the average relative coefficient changes for a set of estimates and it is useful for assessing the effect of the parameter estimates by exclusion of the i-th observation. Therefore, a relatively large value of $IE(\cdot)_i$ indicates a potential influential observation that might cause instability in the model fitting.

5.2 Local Influence

Another approach was suggested by Cook (1986), where instead of removing observations, weights are given to them. Local influence calculation can be carried out for model (11). If likelihood displacement $LD(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\vartheta}}) - l(\hat{\boldsymbol{\vartheta}}_{\boldsymbol{\omega}})\}$ is used, where $\hat{\boldsymbol{\vartheta}}_{\boldsymbol{\omega}}$ denotes the MLE under the perturbed model, the normal curvature for $\boldsymbol{\vartheta}$ in the direction of $\boldsymbol{d}$, $\| \boldsymbol{d} \| = 1$ is given by $C_{\boldsymbol{d}}(\boldsymbol{\vartheta}) = 2|\boldsymbol{d}^\top\boldsymbol{\Delta}^\top[\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})]^{-1}\boldsymbol{\Delta}\boldsymbol{d}|$, where $\boldsymbol{\Delta}$ is a $(p+3)\times n$ matrix that depends on the perturbation scheme, whose elements are given by $\Delta_{vi} = \partial^2 l(\boldsymbol{\vartheta}|\boldsymbol{\omega})/\partial\alpha_v\partial\omega_i$, $i = 1,\cdots,n$ and $v = 1,\cdots,p+3$, evaluated at $\hat{\boldsymbol{\vartheta}}$ and $\boldsymbol{\omega}_0$, where $\boldsymbol{\omega}_0$ is the no perturbation vector (see, Cook, 1986). For the NBGG regression model with a cure fraction, the elements of $\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})$ are given in Appendix A. We can also calculate normal curvatures $C_{\boldsymbol{d}}(\alpha)$, $C_{\boldsymbol{d}}(\boldsymbol{\beta})$ and $C_{\boldsymbol{d}}(\boldsymbol{\gamma})$ to perform various index plots, for instance, the index plot of $\boldsymbol{d}_{\max}$, the eigenvector corresponding to $C_{\boldsymbol{d}_{\max}}$, the largest eigenvalue of the matrix $\boldsymbol{B} = -\boldsymbol{\Delta}^\top[\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})]^{-1}\boldsymbol{\Delta}$ and the index plots of $C_{\boldsymbol{d}_i}(\alpha)$, $C_{\boldsymbol{d}_i}(\boldsymbol{\beta})$ and $C_{\boldsymbol{d}_i}(\boldsymbol{\gamma})$, called the total local influence, where $\boldsymbol{d}_i$ are the standard basis vectors of $\boldsymbol{R}^n$. Thus, the curvature in the direction of $\boldsymbol{d}_i$ takes the form $C_i = 2|\boldsymbol{\Delta}_i^\top[\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})]^{-1}\boldsymbol{\Delta}_i|$, where $\boldsymbol{\Delta}_i^\top$ denotes the i-th row of $\boldsymbol{\Delta}$. It is usual to point out those cases such that $C_i \geq 2\bar{C}$, where $\bar{C} = \sum_{i=1}^n C_i/n$. Another influence measure for the i-th observation is $U_i = \sum_{k=1}^{n_1}\lambda_k e_{ki}^2$, where $\{(\lambda_k, \boldsymbol{e}_k)|k = 1,\cdots,n\}$ are the eigenvalue-eigenvector pairs of $\boldsymbol{B}$ with $\lambda_1 \geq \cdots \geq \lambda_{n_1} \geq \lambda_{n_1+1} = \cdots = \lambda_n = 0$ and $\{\boldsymbol{e}_k = (e_{k1},\cdots,e_{kn})^\top\}$ is the associated orthonormal basis. Zhu and Zhang (2004) studied the influence measure u_i systematically under a case weight perturbation. Thus, this influence measure expresses local sensitivity to the log-likelihood of the perturbations. Recently, Ibacahche-Pulgar *et al.* (2012) studied influence diagnostics for elliptical semi parametric mixed models, Vanegas *et al.* (2012) determined the appropriate matrices for diagnostic procedures in Birnbaum-Saunders nonlinear regression models and Zeller *et al.* (2012) proposed the diagnostics in multivariate measurement error models under asymmetric heavy-tailed distributions.

5.3 Curvature Calculations

Next, we calculate for three perturbation schemes the matrix

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{vi})_{\left[(p+3)\times n\right]} = \left(\frac{\partial^2 l(\boldsymbol{\vartheta}|\boldsymbol{\omega})}{\partial\vartheta_v \boldsymbol{\omega}_i}\right)_{\left[(p+3)\times n\right]},$$

where $v = 1, \cdots, p+3$ and $i = 1, \cdots, n$. We consider the model defined in (11), (12) and its log-likelihood function given by (13).

5.3.1 Case-Weight Perturbation

First, we consider a case weight perturbation which modifies the weight given to each subject in the log-likelihood. Consider the vector of weights $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_n)^\top$.

In this case, the log-likelihood function takes the form

$$l(\boldsymbol{\vartheta}; \boldsymbol{\mathcal{D}}|\boldsymbol{\omega}) = \begin{cases} \sum\limits_{i=1}^{n} \omega_i \delta_i \log\left\{\left(\frac{p_{0i}^{-\alpha}-1}{\alpha}\right) f(y_i;\boldsymbol{\gamma})\right\} & \\ + \sum\limits_{i=1}^{n} \omega_i(-\delta_i - 1/\alpha)\log\left\{1 + (p_{0i}^{-\alpha} - 1)F(y_i;\boldsymbol{\gamma})\right\}, & \text{if } \alpha \neq 0; \\ \sum\limits_{i=1}^{n} \omega_i \delta_i \log\left\{-\log(p_{0i}) f(y_i;\boldsymbol{\gamma})\right\} + \sum\limits_{i=1}^{n} \omega_i F(y_i;\boldsymbol{\gamma})\log(p_{0i}), & \text{if } \alpha = 0, \end{cases}$$

where $0 \leq \omega_i \leq 1$, $\boldsymbol{\omega}_0 = (1, \cdots, 1)^\top$. The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top, \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top)^\top$ is given in Appendix B.

5.3.2 Response Perturbation

Since the y_i values have different variances, the is a need to scall the perturbation vector $\boldsymbol{\omega}$ by an estimator of the standard deviation of y_i. We consider here that each y_i is perturbed as $y_{iw} = y_i + \omega_i S_y$, where S_y is a scale factor that can be the estimated standard deviation of y and $\omega_i \in \boldsymbol{R}$.

Here, the perturbed log-likelihood function becomes

$$l(\boldsymbol{\vartheta};\boldsymbol{\mathcal{D}}|\boldsymbol{\omega}) = \begin{cases} \sum\limits_{i=1}^{n} \delta_i \log\left\{\left(\frac{p_{0i}^{-\alpha}-1}{\alpha}\right) f(y_i^*;\boldsymbol{\gamma})\right\} & \\ +\sum\limits_{i=1}^{n}(-\delta_i - 1/\alpha)\log\left\{1+(p_{0i}^{-\alpha}-1)F(y_i^*;\boldsymbol{\gamma})\right\}, & \text{if } \alpha \neq 0; \\ \sum\limits_{i=1}^{n} \delta_i \log\{-\log(p_{0i}) f(y_i^*;\boldsymbol{\gamma})\} + \sum\limits_{i=1}^{n} F(y_i^*;\boldsymbol{\gamma})\log(p_{0i}), & \text{if } \alpha = 0, \end{cases}$$

where $y_i^* = y_i + \omega_i S_y$ and $\boldsymbol{\omega}_0 = (0,\cdots,0)^\top$. The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top, \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top)^\top$ is given in Appendix C.

5.3.3 Explanatory Variable Perturbation

Cook (1986) described a general scheme for perturbing the whole design matrix X in linear regression models. Some authors have studied the perturbation of covariates. This perturbation has a more complicated impact on the estimates. The errors-in-variable model treats the error of covariates and so the local influence under the perturbation of covariates may be related to the errors-in-variable model. Consider now an additive perturbation on a particular continuous explanatory variable, say x_t, by setting $x_{it\omega} = x_{it} + \omega_i S_x$, where S_x is a scaled factor, $\omega_i \in \boldsymbol{R}$. This perturbation scheme leads to the following expressions for the perturbed log-likelihood function

$$l(\boldsymbol{\vartheta};\boldsymbol{\mathcal{D}}|\boldsymbol{\omega}) = \begin{cases} \sum\limits_{i=1}^{n} \delta_i \log\left\{\left(\frac{(p_{0i}^*)^{-\alpha}-1}{\alpha}\right) f(y_i;\boldsymbol{\gamma})\right\} & \\ +\sum\limits_{i=1}^{n}(-\delta_i - 1/\alpha)\log\left\{1+[(p_{0i}^*)^{-\alpha}-1]F(y_i;\boldsymbol{\gamma})\right\}, & \text{if } \alpha \neq 0; \\ \sum\limits_{i=1}^{n} \delta_i \log\{-\log(p_{0i}^*) f(y_i;\boldsymbol{\gamma})\} + \sum\limits_{i=1}^{n} F(y_i;\boldsymbol{\gamma})\log(p_{0i}^*), & \text{if } \alpha = 0, \end{cases}$$

where $p_{0i}^* = \exp(\boldsymbol{x}_i^{*\top}\boldsymbol{\beta})/(1+\exp(\boldsymbol{x}_i^{*\top}\boldsymbol{\beta}))$, $(\boldsymbol{x}_i^{*\top}\boldsymbol{\beta}) = \beta_1 x_{i1} + \cdots + \beta_t\big(x_{it} + \omega_i S_x\big) + \cdots + \beta_p x_{ip}$ and $\boldsymbol{\omega}_0 = (0,\cdots,0)^\top$. The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top, \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top)^\top$ is given in Appendix D.

6. Application

In this section, we discuss an application of the local influence methodology to a set of real data on cancer recurrence. The data come from a study on cutaneous melanoma (a type of malignant cancer) for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alfa-2b) in order to prevent recurrence. Patients were included in the study from 1991 to 1995, and follow-up was conducted until 1998. The data were collected by Ibrahim *et al.* (2001) and represent the survival times (T), until the patient's death. The original sample size was $n = 427$ patients, 10 of whom did not present a value for explanatory variable tumor thickness. When such cases were removed, a sample of size $n = 417$ patients was retained. The percentage of censored observations was 56%. The following variables were associated with each participant, $i = 1, \cdots, 417$: y_i: observed time (in years); x_{i1}: treatment (0:observation, 1:interferon); x_{i2}: age (in years); x_{i3}: nodule (nodule category: 1 to 4); x_{i4}: sex (0:male, 1:female); x_{i5}: p.s. (performance status-patient's functional capacity scale as regards his daily activities: 0:fully active, 1:other) and x_{i6}: tumor (tumor thickness in mm).

First, we consider the NBGG regression model given in (11) with all regressor variables,

$$\log\left(\frac{p_{0i}}{1-p_{0i}}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6}.$$

To obtain the MLEs of the model parameters, we use the MaxBFGS subroutine in Ox. To estimate $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ of the NBGG regression model with cure rate, we fix different values for q. We choose the value of q that maximizes the likelihood function over several values of $q \in (-3, 3)$, thus obtaining $q = 0.7$. Iterative maximization of the logarithm of the likelihood function in (13) begins with an initial guess $\alpha = 1$, $\boldsymbol{\beta} = \mathbf{1}$, $\sigma = 1$ and $\mu = 1/\bar{y}$ (a moment estimator for μ from i.i.d. exponential observations), where $\bar{y} = \sum y_i/n$. Of course, this choice is not foolproof; it is advisable to run the BFGS method several times from different starting values. Table 2 gives the MLEs for the proposed model. At a 5% significance level, all regression coefficients but nodule category (β_3) are non-significant.

Table 2: MLEs of the parameters of the NBGG regression model with cure rate fraction fitted to the cutaneous melanoma data

Parameter	Estimate	Standard Error	p-value
μ	1.371	0.354	–
σ	0.550	0.073	–
α	2.995	2.132	–
β_0	1.329	0.673	0.048
β_1	0.127	0.183	0.487
β_2	-0.010	0.007	0.134
β_3	-0.411	0.132	0.002
β_4	0.057	0.203	0.779
β_5	-0.142	0.221	0.520
β_6	-0.009	0.027	0.746

Now, using the nonparametric bootstrap method with $B = 3000$, we obtain the bootstrap estimates and the BCa confidence intervals, as described in Section 4. The estimates are presented in Table 3. Note that the estimates from the two methods taken for illustration are very similar, as expected. However, since these methods are based on the likelihood, and asymptotic normality is expected for this sample size ($n = 417$), we can continue the analysis using the MLEs.

Table 3: Nonparametric bootstrap from the NBGG regression model with cure rate fraction fitted to the cutaneous melanoma data set

Parameter	Estimate	Standard Error	95% C.I. Bca
μ	1.326	0.291	(0.875, 1.924)
σ	0.515	0.103	(0.273, 0.668)
α	3.846	4.357	(0.267, 14.139)
β_0	1.557	0.674	(0.408, 2.939)
β_1	0.087	0.224	(-0.404, 0.485)
β_2	-0.011	0.007	(-0.028, 0.002)
β_3	-0.450	0.134	(-0.717, -0.215)
β_4	0.095	0.233	(-0.328, 0.580)
β_5	-0.152	0.257	(-0.681, 0.357)
β_6	-0.014	0.032	(-0.089, 0.038)

6.1 Global Influence Analysis

In this section, we use Ox to compute the case-deletion measures $GD_i(\boldsymbol{\vartheta})$ and $LD_i(\boldsymbol{\vartheta})$ presented in Section 4.1. These influence measures are plotted in Figure 3.

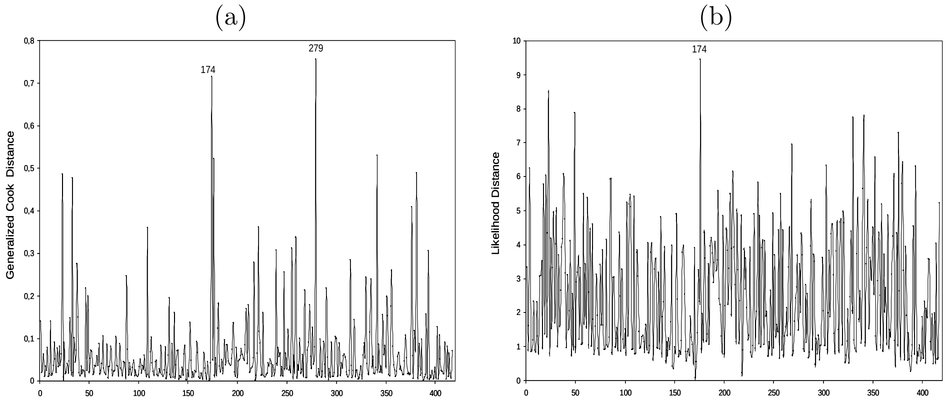

Figure 3: Index plot of $GD_i(\boldsymbol{\vartheta})$, generalized Cook's distance (Figure 3a). Index plot of $LD_i(\boldsymbol{\vartheta})$, likelihood distance (Figure 3b)

Figure 3 indicates that the cases 174 and 279 are possible influential observations. Similarly to the $GD_i(\boldsymbol{\vartheta})$ and $LD_i(\boldsymbol{\vartheta})$ statistics, we calculate the new measures $IE(\boldsymbol{\theta})_i$, $IE(\boldsymbol{\gamma})_i$ and $IE(\boldsymbol{\beta})_i$. The index plots for these influence measures are displayed in Figure 4. Clearly, the most influential observations are 174 and 259.

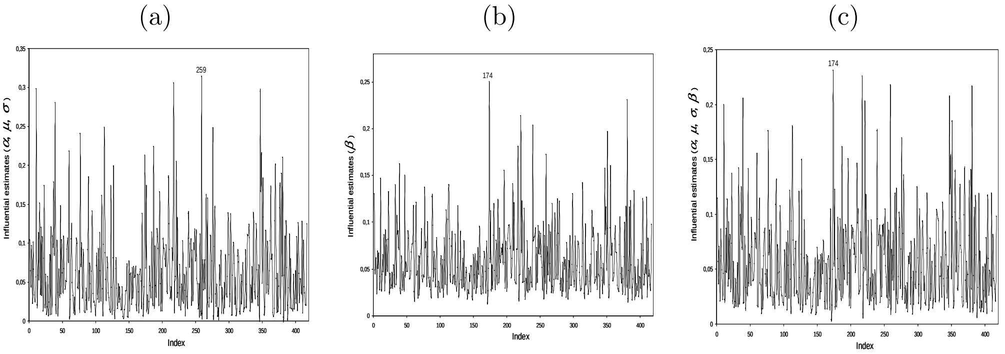
(a) (b) (c)

Figure 4: Index plot of $IE(\boldsymbol{\theta})_i$ (Figure 4a). Index plot of $IE(\boldsymbol{\gamma})_i$ (Figure 4b). Index plot of $IE(\boldsymbol{\beta})_i$ (Figure 4c)

6.2 Local Influence Analysis

In this section, we analyze the local influence for the cancer data.

6.2.1 Explanatory Variable Perturbation

By applying the local influence methodology developed in Section 4, where case-weight perturbation is used, we obtain the values $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\vartheta}) = 1.57$, $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\gamma})$ $= 1.24$ and $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\beta}) = 1.47$ as maximum curvatures. In Figure 5, the index plots of $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$, C_i and U_i for all points are presented. Clearly, the most influential observations on $\hat{\boldsymbol{\vartheta}}$ are the cases 174, 279 and 381 (see Figure 4).

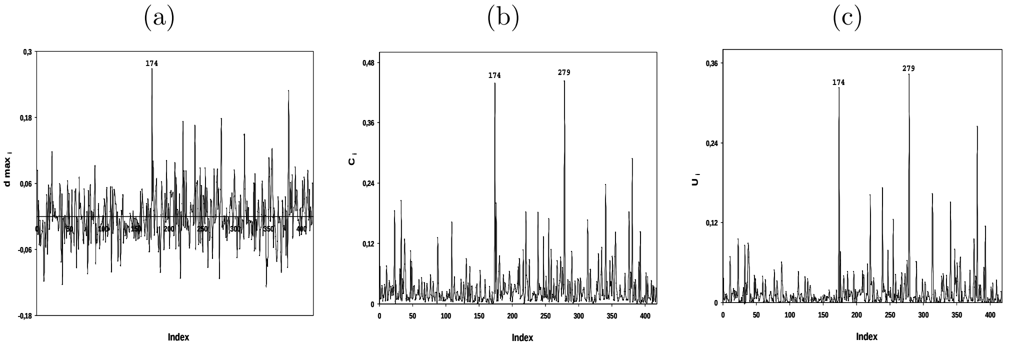
(a) (b) (c)

Figure 5: Index plots of $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$ (Figure 5a), C_i (Figure 5b) and U_i (Figure 5c) under the case-weight perturbation scheme

6.2.2 Influence Using Response Variable Perturbation

Next, we examine the influence of perturbations on the observed survival times. The values for the maximum curvature were $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\vartheta}) = 19.60$, $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\gamma}) = 1.58$ and $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\beta}) = 7.34$. Figure 6 displays the plots for $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$, C_i and U_i for all points. The plots in Figures 6a, 6b and 6c indicate that the observations 279 and 341 as the most influential on $\hat{\boldsymbol{\vartheta}}$.

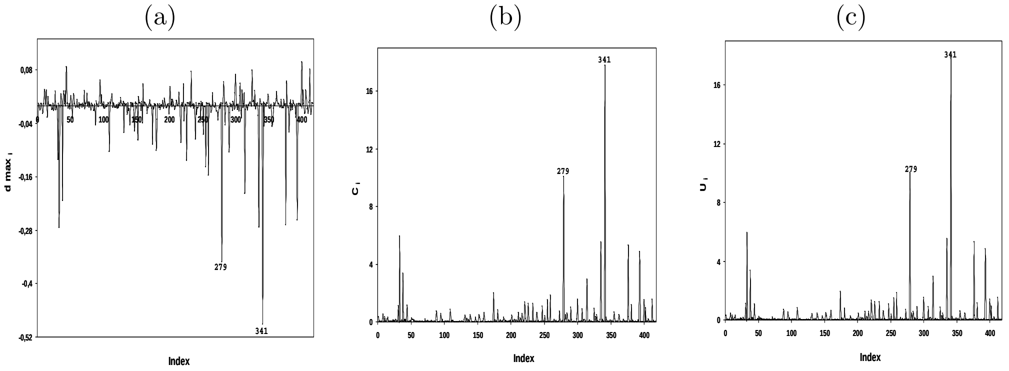

Figure 6: Index plots of $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$ (Figure 6a), C_i (Figure 6b) and U_i (Figure 6c) under the response perturbation scheme

6.2.3 Influence Using Explanatory Variable Perturbation

The perturbation of the explanatory variable age (x_2) is investigated here. After the perturbation of this explanatory variable, the values $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\vartheta}) = 1.12$, $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\gamma}) = 1.10$ and $C_{\boldsymbol{d}_{\max}}(\boldsymbol{\beta}) = 0.95$ were obtained as maximum curvatures. The respective index plots of $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$, C_i and U_i are displayed in Figure 7. The plots in Figures 7a, 7b and 7c indicate that the observation 279 is the most influential on $\hat{\boldsymbol{\vartheta}}$.

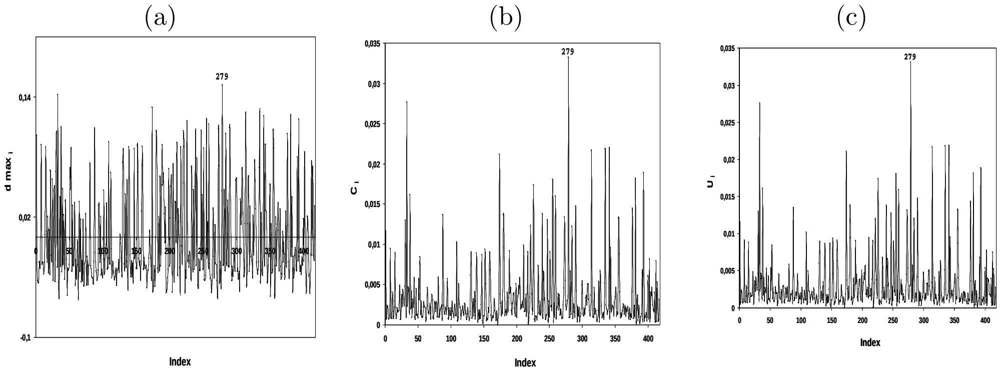

Figure 7: Index plots of $\boldsymbol{d}_{\max}(\boldsymbol{\vartheta})$ (Figure 7a), C_i (Figure 7b) and U_i (Figure 7c) under perturbation of the explanatory variable age

6.3 Impact of the Detected Influential Observations

The diagnostic analysis (global influence and local influence) detected the following three cases as potentially influential: 174, 279 and 341. In order to reveal the impact of these three observations on the parameter estimates, we refitted the model under some situations. First, we individually eliminated each one of these three cases. Next, we removed all potentially influential observations from set "A" (original data set) .

Table 4 gives the relative changes (in percentage) of each parameter estimate, defined by $\boldsymbol{RC}_{\vartheta_j} = [(\hat{\vartheta}_j - \hat{\vartheta}_{j(I)})/\hat{\vartheta}_j] \times 100$, parameter estimates and the corresponding p-values, where $\hat{\vartheta}_{j(I)}$ denotes the MLE of ϑ_j after the set "I" of observations being removed. Note that $I_1 = \{174\}$, $I_2 = \{279\}$, $I_3 = \{341\}$, $I_4 = \{174, 279\}$, $I_5 = \{174, 341\}$, $I_6 = \{279, 341\}$, and $I_7 = \{174, 279, 341\}$.

Table 4 indicates that the MLEs from the NBGG regression model with a cure fraction are not highly sensitive under deletion of the outstanding observations. In general, the significance of the parameter estimates does not change (at 5%) after removing set I. Therefore, we do not have inferential changes after removing the observations handed out in the diagnostic plots The largest variations in the parameter estimates occur with the estimates that are not significant, which should be removed from the model.

We fit the NBGG, Poisson-generalized-gamma (PGG) cure rate and mixture-generalized-gamma (MGG) regression models to these data. For details, see for example, Ortega *et al.* (2009b) and Ortega *et al.* (2009a). The fitted models can be compared employing the Akaike information criterion (AIC). Table 5 gives the estimates (and their standard errors) of the parameters for both regression models and the AIC values in increasing order. The NBGG model yields the best fitting according to these criteria.

The QQ plot of the normalized randomized quantile residuals (Dunn and Smyth, 1996; Rigby and Stasinopoulos, 2005) in Figure 8 suggests that the NBGG regression model yields an acceptable fit. Each point in Figure 8 corresponds to the median of five sets of ordered residuals. Hence, in the rest of this section we adopt this model.

Finally, we end up our application dealing with the estimation of the surviving fraction (p_0). To estimate the proportion of cured individuals, we use (10) and the invariance property of the MLEs, namely

$$\hat{p}_{0i} = \frac{\exp(0.900 - 0.409x_3)}{1 + \exp(0.900 - 0.409x_3)}, \quad \text{and} \quad \hat{p}_0 = \frac{\sum_{i=1}^{n} \hat{p}_{0i}}{n} = 0.487.$$

From these results, we note that the estimate of the parameter α is equal to 2.830 > 0, providing favorable indications for the alternative cure model and

Table 4: Relative changes [-***RC***-in %], estimates and the corresponding *p*-values in parentheses for the regression coefficients to explain the expected log-survival time

Dropped	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
None	-	-	-	-	-	-	-	-	-	-
	1.37	0.55	3.00	1.33	0.13	-0.01	-0.41	0.06	-0.14	-0.01
	(-)	(-)	(-)	(0.05)	(0.49)	(0.13)	(0.00)	(0.78)	(0.52)	(0.75)
I_1	[6]	[2]	[18]	[13]	[38]	[-18]	[-5]	[115]	[-40]	[-45]
	1.28	0.56	2.46	1.50	0.08	-0.01	-0.43	0.12	-0.08	-0.01
	(-)	(-)	(-)	(0.01)	(0.67)	(0.08)	(0.00)	(0.53)	(0.72)	(0.64)
I_2	[6]	[0]	[13]	[-4]	[25]	[5]	[-3]	[-16]	[49]	[-39]
	1.29	0.55	2.59	1.38	0.10	-0.01	-0.43	0.07	-0.07	-0.01
	(-)	(-)	(-)	(0.03)	(0.61)	(0.17)	(0.00)	(0.75)	(0.76)	(0.66)
I_3	[0]	[4]	[-7]	[-2]	[14]	[-4]	[1]	[-43]	[-12]	[-5]
	1.37	0.53	3.20	1.35	0.11	-0.01	-0.41	0.08	-0.16	-0.01
	(-)	(-)	(-)	(0.04)	(0.54)	(0.11)	(0.00)	(0.68)	(0.46)	(0.72)
I_4	[10]	[-1]	[24]	[-13]	[55]	[-10]	[-6]	[-108]	[91]	[-68]
	1.24	0.55	2.29	1.50	0.06	-0.01	-0.44	0.12	-0.01	-0.01
	(-)	(-)	(-)	(0.01)	(0.76)	(0.12)	(0.00)	(0.55)	(0.96)	(0.60)
I_5	[6]	[2]	[11]	[-14]	[51]	[-21]	[-3]	[-157]	[31]	[-45]
	1.29	0.54	2.67	1.51	0.06	-0.01	-0.43	0.15	-0.10	-0.01
	(-)	(-)	(-)	(0.01)	(0.73)	(0.07)	(0.00)	(0.44)	(0.68)	(0.63)
I_6	[6]	[3]	[7]	[-6]	[39]	[1]	[-2]	[-58]	[40]	[-44]
	1.29	0.53	2.78	1.40	0.08	-0.01	-0.42	0.09	-0.09	-0.01
	(-)	(-)	(-)	(0.03)	(0.67)	(0.15)	(0.00)	(0.65)	(0.71)	(0.64)
I_7	[9]	[3]	[17]	[-13]	[68]	[-13]	[-4]	[-148]	[85]	[-68]
	1.24	0.53	2.50	1.51	0.04	-0.01	-0.43	0.14	-0.02	-0.01
	(-)	(-)	(-)	(0.01)	(0.82)	(0.10)	(0.00)	(0.47)	(0.93)	(0.59)

Table 5: MLEs for the regression model with a cure fraction fitted and in parenthesis the standard error to the cutaneous melanoma data set and information criteria considering other models

Parameter	NBGG	PGG	MGG
μ	1.335 (0.248)	0.883 (0.082)	0.719 (0.067)
σ	0.556 (0.054)	0.642 (0.045)	0.666 (0.046)
α	2.830 (1.257)	—	—
β_0	0.900 (0.307)	-1.258 (0.200)	1.153 (0.269)
β_3	-0.409 (0.097)	0.368 (0.069)	-0.473 (0.105)
AIC	1024.009	1029.934	1038.993

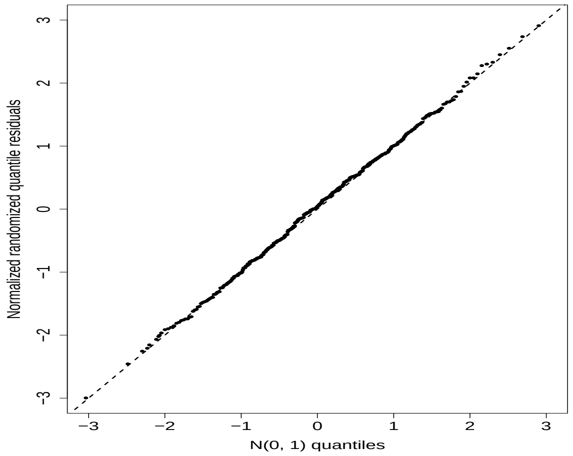

Figure 8: QQ plot of the normalized randomized quantile residuals with identity line for the GG regression model

with super-dispersion. We also observe that the parameter β_3 is significant (at the 5% level), which indicates the nodule size influences the survival time of the patients. Besides this, we note that the estimate of β_3 is negative, which implies that the larger the nodule, the smaller the estimated probability of the patient's survival is. In this study, we could also verify that approximately 49% of the patients are cured of skin cancer.

The MLEs of the survival function and Kaplan-Meier estimate are presented in Figure 9. The model provides a good fit for the cured fraction.

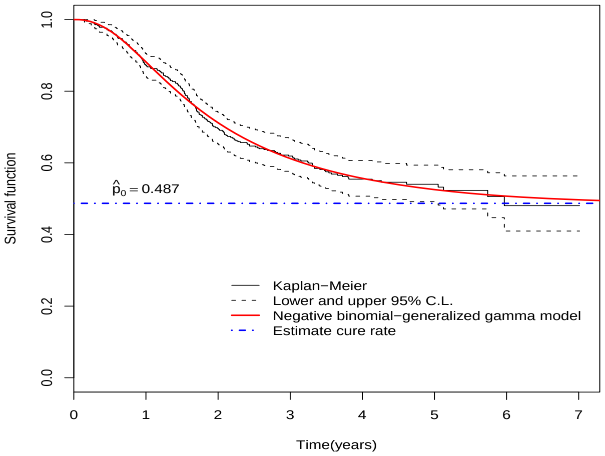

Figure 9: Kaplan-Meier curves (solid lines), estimated survival, estimates of the cure fraction and and upper and lower 95% confidence limits for the cutaneous melanoma data

Next, we turn to a simplified model retaining the nodule category as the only covariate. The estimates of the surviving fraction of patients stratified by nodule category from 1 to 4 (and standard error) are 0.620(0.0406), 0.521(0.0421), 0.420(0.0382) and 0.344(0.0328), respectively, where standard error were obtained after an application of the delta method. Figure 10 displays the surviving function stratified by nodule category from 1 to 4 jointly with the Kaplan-Meier estimate (left panel). Also, Figure 10 (right panel) displays the surviving function stratified by nodule category for non-cured patients.

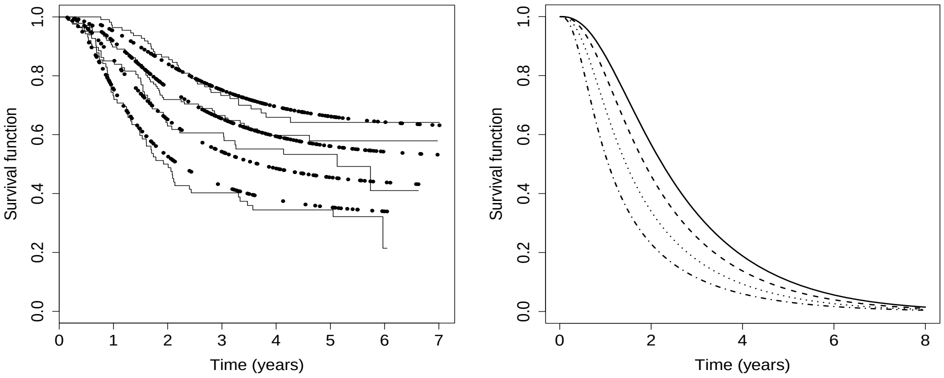

Figure 10: Kaplan-Meier curves (solid lines) and the estimated of the survival function for the NBGG model (left panel) and the estimated of the survival function for the non-cured pacients (right panel) stratified by nodule category (1-4, from top to bottom).

7. Concluding Remarks

In this paper, we propose a model for lifetime data, called the negative binomial generalized gamma (NBGG for short) cure rate model, which was conceived inside a latent competing causes scenario with cure fraction. Maximum likelihood inference is implemented straightforwardly and asymptotic theory may be considered for generating confidence intervals for the parameters and hypothesis tests. Also, we provide applications of influence diagnostics (global, local and total influence) in the NBGG model with covariates. The necessary matrices for application of the techniques were obtained by taking into account some usual perturbations in the model/data. Therefore, the NBGG regression model with a cure fraction can be an interesting option to explain/predict the log-survival time and long-term individuals.

Appendix A: Hessian Matrix $\ddot{\boldsymbol{L}}(\boldsymbol{\vartheta})$

Here, we derive the necessary formulas to obtain the second-order partial derivatives of the log-likelihood function. After some algebraic manipulations, we obtain

$$
\begin{aligned}
\boldsymbol{L}_{\alpha\alpha} &= \sum_{i=1}^{n} \delta_i \frac{p_{0i}^{-\alpha}\log(p_{0i})[-\alpha\log(p_{0i}) - p_{0i}^{-\alpha}]}{h_i^2} - \frac{1}{\alpha^2}\sum_{i=1}^{n} \frac{F(y_i;\gamma)p_{0i}^{-\alpha}\log(p_{0i})}{g_i} \\
&\quad -\frac{2}{\alpha^3}\sum_{i=1}^{n}\log\left[1 + h_iF(y_i;\boldsymbol{\gamma})\right] + \sum_{i=1}^{n}\frac{[F(y_i;\boldsymbol{\gamma})]^2[\log(p_{0i})]^2(\delta_i+\alpha^{-1})p_{0i}^{-2\alpha}}{g_i^2} \\
&\quad +\sum_{i=1}^{n}\frac{F(y_i;\boldsymbol{\gamma})\log(p_{0i})p_{0i}^{-\alpha}\big[-\alpha^{-2} - (\delta_i+\alpha^{-1})\log(p_{0i})\big]}{g_i};
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{L}_{\alpha\beta_j} &= \sum_{i=1}^{n}\frac{\delta_i x_{ij}p_{0i}^{-\alpha}}{h_i^2[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]}\Big\{-h_i[1-\alpha\log(p_{0i})] - \alpha p_{0i}^{-\alpha}\log(p_{0i})\Big\} \\
&\quad -\sum_{i=1}^{n}\frac{x_{ij}p_{0i}^{-\alpha}F(y_i;\boldsymbol{\gamma})}{g_i[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]}\Big\{\alpha^{-1} - (\delta_i+\alpha^{-1})[1-\alpha\log(p_{0i})]\Big\} \\
&\quad -\sum_{i=1}^{n}\frac{x_{ij}\alpha p_{0i}^{-2\alpha}\log(p_{0i})(\delta_i+\alpha^{-1})[F(y_i;\boldsymbol{\gamma})]^2}{g_i^2[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]};
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{L}_{\alpha\gamma_k} &= \sum_{i=1}^{n}\frac{\delta_i[\dot{f}(y_i;\boldsymbol{\gamma})]_{\gamma_k}}{h_i^2 f(y_i;\gamma)}\left[h_i + \alpha p_{0i}^{-\alpha}\log(p_{0i})\right] \\
&\quad +\sum_{i=1}^{n}[\dot{F}(y_i\boldsymbol{\gamma})]_{\gamma_k}\Big[\frac{-h_i}{\alpha^2 g_i} + \frac{p_{0i}^{-\alpha}\log(p_{0i})}{g_i^2(\delta_i-\alpha^{-1})^{-1}}\Big];
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{L}_{\beta_j\beta_{j'}} &= \sum_{i=1}^{n}\frac{\delta_i x_{ij}x_{ij'}p_{0i}}{h_i^2[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]^2}\Big\{h_i[1-\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})] - p_{0i}\Big\} \\
&\quad -\sum_{i=1}^{n}\frac{(\delta_i+\alpha^{-1})F(y_i;\boldsymbol{\gamma})x_{ij}x_{ij'}p_{0i}}{g_i^2[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]^2}\Big\{g_i[1-\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})] - p_{0i}F(y_i;\boldsymbol{\gamma})\Big\};
\end{aligned}
$$

$$
\boldsymbol{L}_{\beta_j\gamma_k} = \alpha\sum_{i=1}^{n}\frac{(\delta_i+\alpha^{-1})x_{ij}p_{0i}^{-\alpha}[\dot{F}(y_i;\boldsymbol{\gamma})]_{\gamma_k}}{g_i^2[1+\exp(\boldsymbol{x}_i^\top\boldsymbol{\beta})]};
$$

$$
\begin{aligned}
\boldsymbol{L}_{\gamma_k\gamma_{k'}} &= \alpha\sum_{i=1}^{n}\frac{\delta_i}{h_i[f(y_i;\boldsymbol{\gamma})]^2}\Big\{[\ddot{f}(y_i;\boldsymbol{\gamma})]_{\gamma_k\gamma_{k'}}f(y_i;\boldsymbol{\gamma}) - [\dot{f}(y_i;\boldsymbol{\gamma})]_{\gamma_k}[\dot{f}(y_i;\boldsymbol{\gamma})]_{\gamma_{k'}}\Big\} \\
&\quad -\sum_{i=1}^{n}\frac{h_i(\delta_i+\alpha^{-1})}{g_i^2}\Big\{[\ddot{F}(y_i;\boldsymbol{\gamma})]_{\gamma_k\gamma_{k'}}g_i - h_i\dot{F}(y_i;\boldsymbol{\gamma})]_{\gamma_k}[\dot{F}(y_i;\boldsymbol{\gamma})]_{\gamma_{k'}}\Big\};
\end{aligned}
$$

where

$$
\begin{aligned}
&h_i = p_{0i}^{-\alpha} - 1, \quad g_i = 1 + h_i F(y_i; \boldsymbol{\gamma}),\\
&[\dot{f}(y_i; \boldsymbol{\gamma})]_{\gamma_k} = \partial f(y_i; \boldsymbol{\gamma})/\partial \gamma_k, \quad [\ddot{f}(y_i; \boldsymbol{\gamma})]_{\gamma_k \gamma_{k'}} = \partial^2 f(y_i; \boldsymbol{\gamma})/\partial \gamma_k \partial \gamma_{k'},\\
&[\dot{F}(y_i; \boldsymbol{\gamma})]_{\gamma_j} = \partial F(y_i; \boldsymbol{\gamma})/\partial \gamma_k, \quad [\ddot{F}(y_i; \boldsymbol{\gamma})]_{\gamma_k \gamma_{k'}} = \partial^2 F(y_i; \boldsymbol{\gamma})/\partial \gamma_k \partial \gamma_{k'},
\end{aligned}
$$

with p_{0i} is defined in Section 3, $i = 1, \cdots, n$, $j, j' = 1, \cdots, p$ and $k, k' = 1, 2$.

Appendix B: Case-Weight Perturbation Scheme

Here, we provide the elements considering the case-weight perturbation scheme. The elements of the matrix $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top(p \times n), \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top(2 \times n)]^\top$ are expressed as

$$
\begin{aligned}
(\Delta_\alpha)_i &= \frac{\delta_i \big\{ -\hat{p}_{0i}^{-\hat{\alpha}}[\hat{\alpha} \log(\hat{p}_{0i}) + 1] + 1 \big\}}{\hat{\alpha} \hat{h}_i} + \frac{\log(\hat{g}_i)}{\hat{\alpha}^2} + \frac{(\delta_i + \hat{\alpha}^{-1}) \hat{p}_{0i}^{-\hat{\alpha}} \log(\hat{p}_{0i}) F(y_i; \hat{\boldsymbol{\gamma}})}{\hat{g}_i};\\
(\Delta_{\beta_j})_i &= \frac{-\hat{\alpha} x_{ij} \hat{p}_{0i}^{-\hat{\alpha}}}{[1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]} \left[\frac{\delta_i}{\hat{h}_i} - \frac{(\delta_i + \hat{\alpha}^{-1}) F(y_i; \hat{\boldsymbol{\gamma}})}{\hat{g}_i} \right];\\
(\Delta_{\gamma_k})_i &= \frac{\hat{\alpha} \delta_i [\dot{f}(y_i; \boldsymbol{\gamma})]_{\gamma_k}}{\hat{h}_i f(y_i; \hat{\boldsymbol{\gamma}})} - \frac{(\delta_i + \hat{\alpha}^{-1}) \hat{h}_i [\dot{F}(y_i; \boldsymbol{\gamma})]_{\gamma_k}}{\hat{g}_i}.
\end{aligned}
$$

Appendix C: Response Perturbation Scheme

Here, we provide the elements $\boldsymbol{\Delta}_{ji}$ considering the response variable perturbation scheme. The elements of the matrix $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top(p \times n), \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top(2 \times n)]^\top$ are expressed as

$$
\begin{aligned}
(\Delta_\alpha)_i &= \frac{[\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i}}{\hat{g}_i} \left[\frac{\hat{h}_i}{\hat{\alpha}^2} + \frac{(\delta_i + \hat{\alpha}^{-1}) \hat{p}_{0i}^{-\hat{\alpha}} \log(\hat{p}_{0i})}{\hat{g}_i} \right];\\
(\Delta_{\beta_j})_i &= \frac{\hat{\alpha} x_{ij} \hat{p}_{0i}^{-\hat{\alpha}} (\delta_i + \hat{\alpha}^{-1}) [\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i}}{\hat{g}_i^2 [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]};\\
(\Delta_{\gamma_k})_i &= \frac{\hat{\alpha}}{\hat{h}_i [f(y_i; \hat{\boldsymbol{\gamma}})]^2} \Big\{ [\ddot{f}(y_i^*; \boldsymbol{\gamma})]_{\omega_i \gamma_k} f(y_i; \hat{\boldsymbol{\gamma}}) - [\dot{f}(y_i^*; \boldsymbol{\gamma})]_{\gamma_k} [\dot{f}(y_i^*; \boldsymbol{\gamma})]_{\omega_i} \Big\}\\
&\quad - \frac{(\delta_i + \hat{\alpha}^{-1}) \hat{h}_i}{\hat{g}_i^2} \Big\{ \hat{g}_i [\ddot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i \gamma_k} - \hat{h}_i [\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i} [\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\gamma_k} \Big\}.
\end{aligned}
$$

Appendix D: Explanatory Variable Perturbation

Here, we provide the elements $\boldsymbol{\Delta}_{ji}$ considering the explanatory variable perturbation scheme. The elements of $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_{\boldsymbol{\beta}}^\top(p \times n), \boldsymbol{\Delta}_{\boldsymbol{\gamma}}^\top(2 \times n)]^\top$ are given by

$$
\begin{aligned}
(\Delta_\alpha)_i = {} & \frac{\delta_i \hat{\alpha} \log(\hat{p}_{0i}) \hat{p}_{0i}^{-\hat{\alpha}} \hat{\beta}_t S_x (1 + \hat{p}_{0i}^{-\hat{\alpha}})}{\hat{h}_i [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]} + \frac{\hat{\alpha} [F(y_i; \hat{\boldsymbol{\gamma}})]^2 (\delta_i + \hat{\alpha}^{-1}) \log(\hat{p}_{0i}) \hat{\beta}_t S_x}{\hat{g}_i^2 [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]} \\
& - \frac{F(y_i; \hat{\boldsymbol{\gamma}}) \hat{p}_{0i}^{\hat{\alpha}} \hat{\beta}_t S_x}{\hat{g}_i [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]} \Big\{ \hat{\alpha}^{-1} - (\delta_i + \hat{\alpha}^{-1}) \big[- \hat{\alpha} \log(\hat{p}_{0i}) + 1 \big] \Big\};
\end{aligned}
$$

$$
(\Delta_{\gamma_k})_i = \frac{\hat{\alpha} \hat{\beta}_t S_x \hat{p}_{0i}^{-\hat{\alpha}}}{[1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]} \left\{ \frac{\hat{\alpha} \delta_i [\dot{f}(y_i; \boldsymbol{\gamma})]_{\gamma_k}}{h_i^2 f(y_i; \hat{\boldsymbol{\gamma}})} + \frac{(\delta_i + \hat{\alpha}^{-1}) [\dot{F}(y_i; \boldsymbol{\gamma})]_{\gamma_k}}{\hat{g}_i^2} \right\}.
$$

For $j \neq t$,

$$
\begin{aligned}
(\Delta_{\beta_j})_i = {} & \frac{\hat{\alpha} \delta_i x_{ij} \hat{\beta}_t S_x \hat{p}_{0i}^{-\hat{\alpha}}}{\hat{h}_i [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]^2} \Big\{ - \hat{h}_i [\hat{\alpha} + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})] + \hat{\alpha} \hat{p}_{0i}^{-\hat{\alpha}} \Big\} \\
& + \frac{\hat{\alpha} x_{ij} \hat{\beta}_t S_x \hat{p}_{0i}^{-\hat{\alpha}} (\delta_i + \hat{\alpha}^{-1}) F(y_i; \hat{\boldsymbol{\gamma}})}{\hat{g}_i^2 [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]^2} \left\{ \frac{\hat{\alpha} + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})}{[1 + \hat{h}_i F(y_i; \hat{\boldsymbol{\gamma}})]^{-1}} - \hat{\alpha} \hat{p}_{0i}^{-\hat{\alpha}} F(y_i; \hat{\boldsymbol{\gamma}}) \right\}.
\end{aligned}
$$

For $j = t$,

$$
\begin{aligned}
(\Delta_{\beta_t})_i = {} & \frac{\hat{\alpha} \delta_i S_x \hat{p}_{0i}^{-\hat{\alpha}}}{\hat{h}_i^2 [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]^2} \Big\{ \hat{h}_i \big[- x_{it} \hat{\beta}_t (\hat{\alpha} + \exp(\boldsymbol{x}_i^\top) \hat{\boldsymbol{\beta}}) \big] + \hat{\alpha} x_{it} \beta_t \hat{p}_{0i}^{-\hat{\alpha}} \Big\} \\
& + \frac{\hat{\alpha} S_x (\delta_i + \hat{\alpha}^{-1}) [F(y_i; \hat{\boldsymbol{\gamma}})]^2}{\hat{g}_i{}^2 \hat{p}_{0i}^{\hat{\alpha}} [1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})]^2} \left\{ \frac{x_{it} \hat{\beta}_t \hat{\alpha} - 1 + \exp(\boldsymbol{x}_i^\top \hat{\boldsymbol{\beta}})(1 + x_{it} \hat{\beta}_t)}{F(y_i; \hat{\boldsymbol{\gamma}}) \hat{g}_i^{-1}} - \frac{\hat{\alpha} x_{it} \hat{\beta}_t}{\hat{p}_{0i}^{\hat{\alpha}}} \right\},
\end{aligned}
$$

where

$$
\hat{h}_i = \hat{p}_{0i}^{-\hat{\alpha}} - 1, \quad \hat{g}_i = 1 + \hat{h}_i F(y_i; \hat{\boldsymbol{\gamma}}), \quad [\dot{f}(y_i; \boldsymbol{\gamma})]_{\gamma_k} = \partial f(y_i; \boldsymbol{\gamma}) / \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}},
$$

$$
[\dot{F}(y_i; \boldsymbol{\gamma})]_{\gamma_k} = \partial F(y_i; \boldsymbol{\gamma}) / \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}}, \quad [\dot{f}(y_i^*; \boldsymbol{\gamma})]_{\omega_i} = \partial f(y_i^*; \boldsymbol{\gamma}) / \partial \omega_i \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0},
$$

$$
\begin{aligned}
[\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i} &= \partial F(y_i^*; \boldsymbol{\gamma}) / \partial \omega_i \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}, \\
[\dot{f}(y_i^*; \boldsymbol{\gamma})]_{\gamma_k} &= \partial f(y_i^*; \boldsymbol{\gamma}) / \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}, \\
[\dot{F}(y_i^*; \boldsymbol{\gamma})]_{\gamma_k} &= \partial F(y_i^*; \boldsymbol{\gamma}) / \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}, \\
[\ddot{f}(y_i^*; \boldsymbol{\gamma})]_{\omega_i \gamma_k} &= \partial^2 f(y_i^*; \boldsymbol{\gamma}) / \partial \omega_i \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}, \\
[\ddot{F}(y_i^*; \boldsymbol{\gamma})]_{\omega_i \gamma_k} &= \partial^2 F(y_i^*; \boldsymbol{\gamma}) / \partial \omega_i \partial \gamma_k \big|_{\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0},
\end{aligned}
$$

where p_{0i} is defined in Section 3, $i = 1, \cdots, n$, $j = 1, \cdots, p$, and $k = 1, 2$.

Acknowledgements

The financial support from FAPESP, CAPES and CNPq is gratefully acknowledged. We thank the associate editor and the reviewer for constructive comments.

References

Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501-515.

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical, Society B* **11**, 15-53.

Cancho, V. G., Ortega, E. M. M. and Bolfarine, H. (2009). The log-exponentiated-weibull regression models with cure rate: local influence and residual analysis. *Journal of Data Science* **7**, 433-458.

Cancho, V. G., Rodrigues, J. and de Castro, M. (2011). A flexible model for survival data with a cure rate: a Bayesian approach. *Journal of Applied Statistics* **38**, 57-70.

Cook, R. D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society, Series B* **48**, 133-169.

Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* Chapman and Hall, Boca Raton, Florida.

Cooner, F., Banerjee, S., Carlin, B. P. and Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association* **102**, 560-572.

Cox, C., Chu, H., Schneider, M. F. and Muñoz, A. (2007). Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in medicine* **26**, 4352-4374.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press, Cambridge.

de Castro, M., Cancho, V. G. and Rodrigues, J. (2009). A Bayesian long-term survival model parametrized in the cured fraction. *Biometrical Journal* **51**, 443-455.

de Castro, M., Cancho, V. G. and Rodrigues, J. (2010). A note on a unified approach for cure rate models. *Brazilian Journal of Probability and Statistics* **24**, 100-103.

DiCiccio, T. J. and Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science* **11**, 189-212.

Doornik, J. A. (2002). *Object-Oriented Matrix Programming Using Ox*, 3rd edition. Timberlake Consultants Press and Oxford, London.

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236-244.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* **7**, 1-26.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, page 436. Chapman and Hall, New York.

Fachini, J. B., Ortega, E. M. M. and Louzada-Neto, F. (2008). Influence diagnostics for polyhazard models in the presence of covariates. *Statistical Methods and Applications* **17**, 413-433.

Hashimoto, E. M., Ortega, E. M. M., Cancho, V. G. and Cordeiro, G. M. (2010). The log-exponentiated Weibull regression model for interval-censored data. *Computational Statistics & Data Analysis* **54**, 1017-1035.

Ibacache-Pulgar, G., Paula, G. A. and Galea, M. (2012). Influence diagnostics for elliptical semiparametric mixed models. *Statistical Modelling* **12**, 165-193.

Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis.* Springer, New York.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* Wiley, New York.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data.* Wiley, New York.

Ortega, E. M. M., Cancho, V. G. and Lachos, V. H. (2009a). A generalized log-gamma mixture model for cure rate: estimation and sensitivity analysis. *Shankhya, Series B* **71**, 1-29.

Ortega, E. M. M., Cordeiro, G. M. and Pascoa, M. A. R. (2011). The generalized gamma geometric distribution. *Journal of Statistical Theory and Applications* **10**, 433-454.

Ortega, E. M. M., Bolfarine, H. and Paula, G. A. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistical* & *Data Analysis* **42**, 165-186.

Ortega, E. M. M., Cancho, V. G. and Paula, G. A. (2009b). Generalized log-gamma regression models with cure fraction. *Lifetime Data Analysis* **15**, 79-106.

Piegorsch, W. W. (1990). Maximum likelihood estimation for the negative binomial dispersion parameter. *Biometrics* **46**, 863-867.

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). *Applied Statistics* **54**, 507-554.

Rodrigues, J., Cancho, V. G., de Castro, M. and Louzada-Neto, F. (2009). On the unification of the long-term survival models. *Statistics* & *Probability Letters* **79**, 753-759.

Ross, G. J. S. and Preece, D. A. (1985). The negative binomial distribution. *Journal of the Royal Statistical Society, Series D* **34**, 323-336.

Saha, K. and Paul, S. (2005). Bias-corrected maximum likelihood estimator of the negative binomial dispersion parameter. *Biometrics* **61**, 179-185.

Silva, G. O., Ortega, E. M. M., Cancho, V. G. and Barreto, M. L. (2008). Log-Burr XII regression models with censored data. *Computational Statistical and Data Analysis* **52**, 3820-3842.

Stacy, E. W. (1962). A generalization of the gamma distribution. *Annals of Mathematical Statistics* **33**, 1187-1192.

Tournoud, M. and Ecochard, R. (2008). Promotion time models with time-changing exposure and heterogeneity: application to infectious diseases. *Biometrical Journal* **50**, 395-407.

Tsodikov, A. D., Ibrahim, J. G. and Yakovlev, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association* **98**, 1063-1078.

Vanegas, L. H., Rondón, L. M. and Cysneiros, F. J. A. (2012). Diagnostic procedures in Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis* **56**, 1662-1680.

Wang, P., Puterman, M. L., Cockburn, I. and Le, N. (1996). Mixed poisson regression models with covariate dependent rates. *Biometrics* **52**, 381-400.

Yakovlev, A. Y., Tsodikov, A. D. and Asselain, B. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore.

Yin, G. and Ibrahim, J. G. (2005). Cure rate models: a unified approach. *Canadian Journal of Statistics* **33**, 559-570.

Zeller, C. B., Carvalho, R. R. and Lachos, V. H. (2012). On diagnostics in multivariate measurement error models under asymmetric heavy-tailed distributions. *Statistical Papers* **53**, 665-683.

Zhu, H. and Zhang, H. (2004). A diagnostic procedure based on local influence. *Biometrika* **91**, 579-589.

Received August 31, 2012; accepted July 4, 2013.

Edwin M. M. Ortega
Departamento de Ciências Exatas
Universidade de São Paulo
13418-900, Piracicaba, SP, Brazil
edwin@usp.br

Gladys D. C. Barriga
Departamento de Engenharia de Produção
Universidade Estadual Paulista "Júlio de Mesquita Filho"
17033-360, Bauru, SP, Brazil
gladyscacsire@yahoo.com.br

Elizabeth M. Hashimoto
Departamento de Ciências Exatas
Universidade de São Paulo
13418-900, Piracicaba, SP, Brazil
emhashim@usp.br

Vicente G. Cancho
Departamento de Matemática Aplicada e Estatítica
Universidade de São Paulo
13560-970, São Carlos, SP, Brazil
garibay@icmc.usp.br

Gauss M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco
50740-540, Recife, PE, Brazil
gauss@de.ufpe.br