

Article

An Alternative Promotion Time Cure Model with Overdispersed Number of Competing Causes: An Application to Melanoma Data

Diego I. Gallardo ¹, Mário de Castro ² and Héctor W. Gómez ^{3,*}

¹ Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó 1530000, Chile; diego.gallardo@uda.cl

² Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos 13560-095, Brazil; mcastro@icmc.usp.br

³ Departamento de Matemática, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile

* Correspondence: hector.gomez@uantof.cl; Tel.: +56-552-637-278

Abstract: A cure rate model under the competing risks setup is proposed. For the number of competing causes related to the occurrence of the event of interest, we posit the one-parameter Bell distribution, which accommodates overdispersed counts. The model is parameterized in the cure rate, which is linked to covariates. Parameter estimation is based on the maximum likelihood method. Estimates are computed via the EM algorithm. In order to compare different models, a selection criterion for non-nested models is implemented. Results from simulation studies indicate that the estimation method and the model selection criterion have a good performance. A dataset on melanoma is analyzed using the proposed model as well as some models from the literature.

Keywords: bell distribution; EM algorithm; long-term survival model; maximum likelihood; model comparison



Citation: Gallardo, D.I.; de Castro, M.; Gómez, H.W. An Alternative Promotion Time Cure Model with Overdispersed Number of Competing Causes: An Application to Melanoma Data. *Mathematics* **2021**, *9*, 1815. <https://doi.org/10.3390/math9151815>

Academic Editor: Sorana D. Bolboacă

Received: 3 July 2021
Accepted: 28 July 2021
Published: 31 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The usual approach in models for time-to-event data is built on the assumption that all subjects will experience the event of interest if the follow-up time is long enough. In many studies in different areas, there are subjects insusceptible to the event of interest. We call cure rate the proportion of insusceptible subjects. Cure rate models, also known as long-term survival models, are capable to deal with this situation.

Cure rate models have had an intense research activity in last years. From the seminal contributions in [1,2] and the approach based on competing risks for cure rate models in [3], we find many papers in recent years on this research area. The competing risks approach is summarized as follows. Let M be an unobserved variable denoting the initial number of competing causes related to the occurrence of the event of interest. For instance, in cancer studies M represents the number of carcinogenic cells at the end of treatment that can produce a detectable cancer. For M following the Bernoulli or the Poisson distribution, we obtain the models in [2,3], respectively. The time for the j -th competing cause to produce the event of interest (that is, the promotion time) is denoted by Z_j , $j = 1, \dots, M$. We also assume that, conditional on M , the latent times Z_1, \dots, Z_M are independent and identically distributed with cumulative distribution function (cdf) $F(z; \lambda)$ and survival function $S(z; \lambda) = 1 - F(z; \lambda)$, where λ is the parameter vector. Under a competing risks setup, the time elapsed until the event of interest is given by $T = \min(Z_0, Z_1, \dots, Z_M)$, where Z_0 is a random variable degenerate at $+\infty$. Our paper is based on the competing risks approach for cure rate models.

Many cure rate models proposed in last years are formulated from different distributions for the number of competing causes M . This research line includes the negative

binomial [4], COM-Poisson [5], power series [6], Yule-Simon [7], polylogarithm [8], modified power series [9], zero inflated power series [10] and zero-modified geometric [11] distributions, among many others. In this vein, we adopt the one-parameter Bell distribution recently proposed in [12], with probability mass function (pmf)

$$P(M = m; \theta) = \frac{B_m \theta^m e^{-e^\theta + 1}}{m!}, \quad m = 0, 1, 2, \dots, \quad (1)$$

for $\theta > 0$, where B_m is the Bell number defined as $B_m = e^{-1} \sum_{k=0}^{\infty} k^m / k!$. The first few Bell numbers are $B_0 = B_1 = 1, B_2 = 2, B_3 = 5, B_4 = 15, B_5 = 52, B_6 = 203, B_7 = 877, B_8 = 4140, B_9 = 21147$ and $B_{10} = 115975$. We denote $M \sim \text{Bell}(\theta)$ to say that M follows a distribution with pmf in (1). Denoting the mean and the variance of M by $\mathbb{E}(M)$ and $\text{Var}(M)$, respectively, we have that $\mathbb{E}(M) = \theta e^\theta$ and $\text{Var}(M) = \theta(1 + \theta)e^\theta$. Since $\text{Var}(M) > \mathbb{E}(M)$, the Bell distribution accommodates overdispersed counts. For cure rate models, this distribution is interesting due to some characteristics:

1. The distribution has only one parameter, being an alternative to traditional discrete distributions such as Poisson and Geometric.
2. The probability generating function (pgf) of the distribution has a simple expression. In fact, if $M \sim \text{Bell}(\theta)$, then its pgf is given by $G(s; \theta) = \exp(e^{\theta s} - e^\theta)$, for $|s| < 1$. See Proposition 1 in [12]. This fact is relevant from the point of view of cure rate models because the population survival function depends on this function.
3. $P(M = 0; \theta) = \exp(-e^\theta + 1)$ has a simple form. This fact is important because this probability is the cure rate of the model. The simplicity of this term allows, among other things, to reparameterize the model in terms of the cure rate.
4. The distribution belongs to the power series family of distributions [13] with pmf $P(M = m; \theta) = a_m \theta^m / A(\theta)$, where $a_m = B_m / m!$, $m = 0, 1, 2, \dots$, and $A(\theta) = \sum_{m=0}^{\infty} a_m \theta^m = \exp(e^\theta - 1)$. Recently, ref. [14] proposed an EM-type algorithm for a class of cure rate models based on the power series family. The maximization (M) step of the EM algorithm is decomposed in two steps involving the distribution of the number of competing causes M and the distribution of the latent times Z .
5. To the best of our knowledge, the Bell distribution has not yet been proposed to model the number of competing causes in a cure rate models context.

The remaining of the paper is organized as follows. In Section 2, we present the cure rate model. Inference methods are presented in Section 3. Section 4 is dedicated to a simulation study covering properties of the estimators and model selection. A dataset on melanoma is analyzed in Section 5. Closing remarks are given in Section 6.

2. Model

Let M be an unobserved variable denoting the initial number of competing causes related to the occurrence of an event of interest. We assume that $M \sim \text{Bell}(\theta)$. Under the competing risks setup in Section 1, from Theorem 2 in [4], see also [15], we have that the (improper) population survival function of the Bell cure rate (Bellcr) model is given by

$$S_{pop}(t; \theta, \lambda) = P(T > t; \theta, \lambda) = G(S(t; \lambda); \theta) = \exp[e^{\theta S(t; \lambda)} - e^\theta]. \quad (2)$$

It is immediate that the cure rate of the model is given by $p = \lim_{t \rightarrow \infty} S_{pop}(t; \theta) = \exp(1 - e^\theta) = P(M = 0; \theta)$. Henceforth, we adopt the parameterization $p = \exp(1 - e^\theta)$, i.e., $\theta = \log[1 - \log(p)]$. In this way, covariates can be directly linked to the cure rate p , allowing to compare regression coefficients among different models parameterized in terms of the cure rate.

An interesting family of cure rate models parameterized in the cure rate is studied in [16]. Such family includes the Bernoulli cure rate (Berncr) model, also known as mixture cure model [1,2], the Poisson cure rate (Pocr) model, also known as the promotion time cure model [3,17], the logarithmic cure rate (Locr) model and the negative binomial cure

rate (NBcr) model, noticing that the NBcr model has an additional parameter q ($q > 0$) and $q = 1$ corresponds to the geometric cure rate (Geocr) model.

Table 1 and Figure 1 show the variance of the number of competing causes M in terms of the cure rate for some models. Note that the curve for the Bellcr model lies between the Geocr and Pocr models. Also, the curves for the NBcr model with $q = 0.3, 5$ and 10 are close to the Locr, Bellcr and Pocr models, respectively. In fact, by applying the L'Hopital's rule, we have $\lim_{q \rightarrow \infty} q(1 - p^{1/q})p^{-2/q} = -\log(p)$ so that the proximity between the NBcr($q = 10$) and Pocr curves is justified.

Table 1. Variance of the number of competing causes as a function of the cure rate for some models parameterized in the cure rate (p).

Model	Var(M)	Model	Var(M)
Berncr	$p(1 - p)$	NBcr	$q(1 - p^{1/q})p^{-2/q}$
Pocr	$-\log(p)$	Locr *	$\theta[\theta + \log(1 - \theta)]/[(1 - \theta) \log(1 - \theta)]^2$
Geocr	$(1 - p)p^{-2}$	Bellcr	$[1 - \log(p)] \log[1 - \log(p)]\{1 + \log[1 - \log(p)]\}$

* $\theta = \theta(p) = 1 + pW[-\exp(-1/p)/p]$, where $W(\cdot)$ denotes the Lambert function [18].

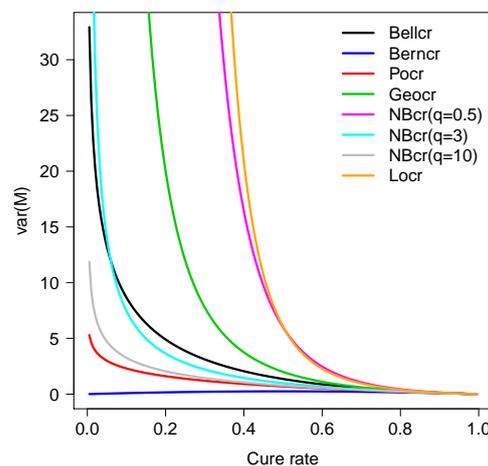


Figure 1. Variance of the number of competing causes as a function of the cure rate for some models.

With the parameterization in the cure rate p , the population survival function for the Bellcr model in (2) is recast, leading to

$$S_{pop}(t; p, \lambda) = p \exp\{[1 - \log(p)]^{S(t;\lambda)} - 1\}, \tag{3}$$

so that the population density and hazard functions are given by

$$f_{pop}(t; p, \lambda) = S_{pop}(t; p, \lambda)[1 - \log(p)]^{S(t;\lambda)} \log[1 - \log(p)]f(t; \lambda) \tag{4}$$

$$\text{and } h_{pop}(t; p, \lambda) = [1 - \log(p)]^{S(t;\lambda)} \log[1 - \log(p)]f(t; \lambda).$$

In a sample of size n , for the i -th subject the covariates are represented by $x_i = (1, x_{1i}, \dots, x_{ri})^\top$, $i = 1, \dots, n$, where the symbol “ \top ” denotes the transpose operator. The cure rate p_i is linked to x_i through the logistic function, that is,

$$\log\left(\frac{p_i}{1 - p_i}\right) = x_i^\top \beta, \tag{5}$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_r)^\top$ is the vector of regression coefficients of dimension $r + 1$. Therefore, interpretations about the regression coefficients can be obtained in terms of the odds ratio for the cure probability.

3. Inference

In this section, we present some details about the estimation procedure for the parameters of the model in Section 2. We consider a framework in which the lifetimes are subject to right censoring. Let Y_i and C_i be the failure and censoring time variables for the i -th subject, respectively. In a sample of size n , the observed variables are $T_i = \min(Y_i, C_i)$ and $\delta_i = I(Y_i \leq C_i)$, where $\delta_i = 1$ and $\delta_i = 0$ denote either if a failure or a censored time was observed for the i -th subject, respectively, $i = 1, \dots, n$. Under the usual assumption of non-informative censoring and with p_i as in (5), the log-likelihood function of $\boldsymbol{\psi} = (\boldsymbol{\beta}^\top, \boldsymbol{\lambda}^\top)^\top$ is given by

$$\ell(\boldsymbol{\psi}) = \sum_{i=1}^n \left(\log(p_i) + \log\{[1 - \log(p_i)t]^{S(t;\boldsymbol{\lambda})} - 1\} + \delta_i\{S(t_i; \boldsymbol{\lambda}) \log[1 - \log(p_i)] + \log\{(\log[1 - \log(p_i)])\} + \log[f(t_i; \boldsymbol{\lambda})]\} \right). \tag{6}$$

The maximum likelihood (ML) estimator of $\boldsymbol{\psi}$, $\hat{\boldsymbol{\psi}}$ say, is obtained by maximizing (6) with respect to $\boldsymbol{\psi}$. However, direct maximization is not computationally simple because some terms in (6) involve both $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$. For this reason and taking advantage of the EM algorithm developed for the power series family of cure rate models in [14], ML estimates in the Bellcr model are computed using this algorithm. In short, the k -th iteration of the EM algorithm goes as follows:

- **E step:** For $i = 1, \dots, n$, compute $\tilde{M}_i^{(k)} = \mu_i^{(k-1)} \exp(\mu_i^{(k-1)}) + \delta_i(1 + \mu_i^{(k-1)})$, where $\mu_i^{(k-1)} = \theta_i^{(k-1)} S(t_i; \boldsymbol{\lambda}^{(k-1)})$ with $\theta_i^{(k-1)} = \log[1 - \log(p_i^{(k-1)})]$ and $p_i^{(k-1)}$ comes from (5).
- **M step 1:** Given $\tilde{\boldsymbol{M}}^{(k)} = (\tilde{M}_1^{(k)}, \dots, \tilde{M}_n^{(k)})^\top$, find $\boldsymbol{\beta}^{(k)}$ that maximizes $Q_1(\boldsymbol{\beta} | \boldsymbol{\psi}^{(k)})$ with respect to $\boldsymbol{\beta}$, where

$$Q_1(\boldsymbol{\beta} | \boldsymbol{\psi}^{(k)}) = \sum_{i=1}^n \left(\tilde{M}_i^{(k)} \log\{\log[1 - \log(p_i)]\} - \log(p_i) \right). \tag{7}$$

- **M step 2:** Given $\tilde{\boldsymbol{M}}^{(k)}$, find $\boldsymbol{\lambda}^{(k)}$ that maximizes $Q_2(\boldsymbol{\lambda} | \boldsymbol{\psi}^{(k)})$ with respect to $\boldsymbol{\lambda}$, where

$$Q_2(\boldsymbol{\lambda} | \boldsymbol{\psi}^{(k)}) = \sum_{i=1}^n \{ \tilde{M}_i^{(k)} \log[S(t_i; \boldsymbol{\lambda})] + \delta_i \log[h(t_i; \boldsymbol{\lambda})] \}. \tag{8}$$

The E and M steps are cycled until a suitable convergence criterion is attained, for instance, $\|\boldsymbol{\psi}^{(k)} - \boldsymbol{\psi}^{(k-1)}\| < \epsilon$, where “ $\|\cdot\|$ ” denotes the Euclidean norm and ϵ is a specified tolerance. Maximization of the functions in (7) and (8) can be performed using the optim function in R language [19].

We stress that the proposed EM algorithm procedure is general in the sense that it can accommodate different distributions for the latent time Z in Section 1. In this work, we assume the Weibull distribution, denoted by $\text{Wei}(\alpha, \nu)$, because it is a suitable model in a cure rate model context. With $\boldsymbol{\lambda} = (\alpha, \nu)^\top$, the survival function for this distribution is given by $S(t; \boldsymbol{\lambda}) = \exp(-e^\alpha t^\nu)$, for $t > 0, \nu > 0$ and $\alpha \in \mathbb{R}$.

On the other hand, the normal theory asymptotic covariance matrix of the maximum likelihood estimator $\hat{\boldsymbol{\psi}}$ can be estimated from minus the Hessian matrix of the log-likelihood function in (6) evaluated at $\boldsymbol{\psi} = \hat{\boldsymbol{\psi}}$. The numDeriv R package [20] provides a numerical approximation to this matrix. Interval estimates of the parameters are computed from the asymptotic standard errors. Computational codes are available from the authors upon request.

Finally, inference usually is made on the cure rate. However, we might also be interested in the cure rate for subjects who survived up to a certain time t_0 , denoted by p_{t_0} . First, for $t > t_0$, from (3) we have $P(T > t | T > t_0) = P(T > t) / P(T > t_0) =$

$S_{pop}(t; p, \lambda) / S_{pop}(t_0; p, \lambda) = \exp\{[1 - \log(p)]^{S(t; \lambda)} - [1 - \log(p)]^{S(t_0; \lambda)}\}$. Then, omitting the dependence on the covariates and passing to the limit for $t \rightarrow \infty$ we obtain the conditional cure rate

$$p_{t_0} = \lim_{t \rightarrow \infty} P(T > t) / P(T > t_0) = \exp\{1 - [1 - \log(p)]^{S(t_0; \lambda)}\}. \tag{9}$$

Of course, if $t_0 = 0$ in (9), p_{t_0} reduces to the cure rate p in (5). For comparative purposes, Remark 1 gives expressions for p_{t_0} under the Pocr, Locr, NBcr and Bincr models.

Remark 1. For the Pocr, Locr, NBcr and Bincr models, routine calculation shows that p_{t_0} is given by

$$\begin{aligned} p_{t_0} &= p^{S(t_0; \lambda)}, \text{ for the Pocr model,} \\ p_{t_0} &= \left(1 - (1 - p^{1/q})S(t_0; \lambda)\right)^q, \text{ for the NBcr model,} \\ p_{t_0} &= (1 + (1/p - 1)S(t_0; \lambda))^{-q}, \text{ for the Bincr model and} \\ p_{t_0} &= \frac{-\theta(p)S(t_0; \lambda)}{\log(1 - \theta(p)S(t_0; \lambda))}, \text{ for the Locr model,} \end{aligned}$$

where $\theta(p) = 1 + pW[-\exp(-1/p)/p]$.

4. Simulation Studies

In this section, we present two simulation studies. The main goal of the first study is to assess the performance of the ML estimates for the parameters of the Bellcr model computed with the EM algorithm in Section 3. The second study is devoted to model selection based on Vuong’s test statistic in [21] (see Appendix A) when the true cure rate model is the Bellcr model or may be wrongly specified. The null hypothesis states that the two compared models are undistinguishable. If the null hypothesis is rejected, the model with the highest value of the likelihood function is preferable. We pick a model with three covariates x_1, x_2 and x_3 . Figure 2 shows the scheme to draw the covariates.

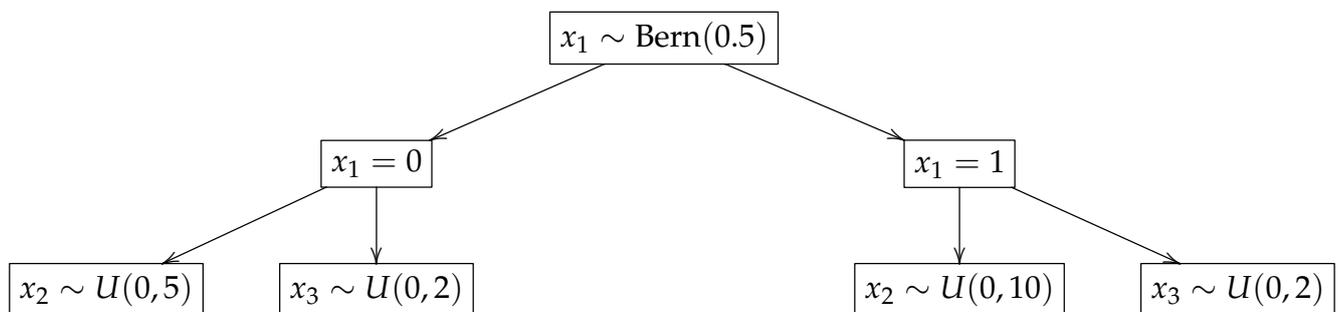


Figure 2. Scheme to draw the covariates in the simulation study. Bern (0.5) denotes the Bernoulli distribution with probability 0.5 and $U(a, b)$ denotes the uniform distribution on (a, b) .

The true value of $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^\top$ is computed from four combinations of values of (x_1, x_2, x_3) and cure rates p_{c1}, \dots, p_{c4} . For (x_1, x_2, x_3) , we choose $(0, 0, 0)$, $(1, 0, 0)$, $(0, 5, 2)$ and $(1, 10, 5)$. Solving the equations in (5), we get $\beta_0 = \text{logit}(p_{c1})$, $\beta_1 = -\text{logit}(p_{c1}) + \text{logit}(p_{c2})$, $\beta_2 = -\text{logit}(p_{c1}) + (2/5)\text{logit}(p_{c2}) + \text{logit}(p_{c3}) - (2/5)\text{logit}(p_{c4})$ and $\beta_3 = 2 \text{logit}(p_{c1}) - \text{logit}(p_{c2}) - 2\text{logit}(p_{c3}) + \text{logit}(p_{c4})$, where $\text{logit}(p) = \log(p) - \log(1 - p)$, for $p \in (0, 1)$. In the studies presented here, cure rates (p_{c1}, \dots, p_{c4}) are $(0.9, 0.8, 0.65, 0.5)$, $(0.8, 0.7, 0.55, 0.4)$ and $(0.7, 0.6, 0.45, 0.3)$, labeled as Cure 1, Cure 2 and Cure 3, respectively. On the other hand, to set the vector $\lambda = (\alpha, \nu)^\top$ we choose values for the expected value $\mathbb{E}(Z)$ and the variance $\text{Var}(Z)$ of the latent time Z in Section 1. For the Wei (α, ν) distribution in Section 3, these conditions imply that $\alpha = \nu\{\log[\Gamma(1 + 1/\nu)] - \log[\mathbb{E}(Z)]\}$ and ν is the solution to the equation $\Gamma(1 + 2/\nu) / \Gamma^2(1 + 1/\nu) = 1 + \text{Var}(Z) / [\mathbb{E}(Z)]^2$. Table 2 showcases the true values of the parameters for the simulations.

Table 2. True values of the parameters in the simulation studies.

Cure Rate	β_0	β_1	β_2	β_3	$(\mathbb{E}(Z), \text{Var}(Z))$	α	ν
Cure 1	2.197	−0.811	−1.024	1.770	(7, 4)	−8.018	3.920
Cure 2	1.386	−0.539	−0.684	1.118	(5, 3)	−5.444	3.165
Cure 3	0.847	−0.442	−0.547	0.843	(2, 1)	−1.712	2.101

The simulations comprise two sample sizes, $n = 200$ and $n = 400$, and two cases for the latent time Z in Section 1, namely, (i) $\mathbb{E}(Z) = 7$ and $\text{Var}(Z) = 4$ and (ii) $\mathbb{E}(Z) = 5$ and $\text{Var}(Z) = 3$. For each combination of sample size, cure rate (p_{c1}, \dots, p_{c4}) and $(\mathbb{E}(Z), \text{Var}(Z))$, we compute $\theta_i = \log[1 - \log(p_i)]$, with p_i as in (5). Next, M_i is drawn from the Bell(θ_i) distribution. If $M_i = 0$, the failure time is $T_i = +\infty$. If $M_i \geq 1$, we draw Z_1, \dots, Z_{M_i} from the Wei(α, ν) distribution and the failure time is $T_i = \min(Z_1, \dots, Z_{M_i})$. The censoring time C_i is sampled from the $U(0, 15)$ distribution. A simulated dataset is formed by $x_i = (1, x_{1i}, x_{2i}, x_{3i})^\top$, $Y_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$, $i = 1, \dots, n$. For the Pocr, Locr, NBcr, Geocr and Berncr models in Section 4.2, the data generation process is similar.

Each scenario was replicated 1000 times. The average proportion of censored times ranges from 56% to 77%. The vectors of covariates $x_i, i = 1, \dots, n$, are kept fixed throughout the replications.

4.1. Estimation

Having in mind the goal of assessing the behavior of the ML estimates of the parameters in the Bellcr model, Table 3 reports the simulated bias for each parameter (Bias), the average of the asymptotic standard errors (SE) based on the covariance matrix in Section 3, the root of the simulated mean squared error (RMSE) and the coverage probability of the normal theory 95% asymptotic confidence intervals (CP). In general, the estimators show a good behavior in all scenarios of our study. We see that Bias is low so that SE and RMSE are close and get closer when the sample size is 400. It is noteworthy that CP is close to the nominal value for both sample sizes, ranging from 0.939 to 0.963. Overall, we see that within the scope of our study, the estimators and the estimation algorithm in Section 3 have a good performance.

4.2. Model Comparison

In this section, we present a simulation study aimed to test the Bellcr model against the Pocr, Locr, NBcr and Geocr models. In our comparisons, the models are compared with the Bellcr model whichever the model generating the data (True model in Tables 4 and 5). The models are tested against the Bellcr model using the Vuong’s test statistic. The nominal significance level is 5%.

In Table 4, when the data generating model is the Bellcr model, rejection rates of the Bellcr model are very low regardless of the model being compared, as expected. When the true model is the Pocr model, the highest rejection rates of the Bellcr model, between 11.4% and 24.4%, correspond to the Pocr and NBcr models. We recall the Pocr model is a limiting case of the NBcr model. In Tables 4 and 5, we see null rejection rates (up to one decimal place) of the Bellcr model against the Berncr model. This is not surprising because the Berncr model (mixture cure model) is too simple to cope with the structure of the data generated in our simulation study.

When the data are drawn from the NBcr ($q = 3$) model, rejection rates in Table 4 are low, as expected in light of Figure 1. The largest rejection rates of the Bellcr model are achieved when the true model is the Locr model in Table 4 (rates between 75.8% and 80.9%) and the Geocr model in Table 5 (rates between 15.8% and 96.2%). In all scenarios in Tables 4 and 5, rejection rates of the Bellcr model are consistent with the patterns in Figure 1.

Table 3. Simulated bias (Bias), average of the asymptotic standard errors (SE), root of the simulated mean squared error (RMSE) and coverage probability of the 95% asymptotic confidence intervals (CP) for the Bellcr model.

Cure Rate	$(\mathbb{E}(Z), \text{Var}(Z))$	Parameter	$n = 200$				$n = 400$			
			Bias	SE	RMSE	CP	Bias	SE	RMSE	CP
Cure 1	(7, 4)	β_0	0.109	0.594	0.659	0.945	0.046	0.403	0.422	0.952
		β_1	−0.003	0.692	0.737	0.949	−0.003	0.471	0.474	0.954
		β_2	−0.067	0.210	0.257	0.963	−0.027	0.138	0.143	0.950
		β_3	0.112	0.396	0.467	0.957	0.049	0.260	0.269	0.955
		α	−0.335	0.823	0.898	0.953	−0.221	0.572	0.647	0.951
		ν	0.164	0.446	0.487	0.951	0.107	0.307	0.339	0.951
	(5, 3)	β_0	0.118	0.532	0.597	0.944	0.050	0.363	0.383	0.948
		β_1	0.006	0.621	0.641	0.958	0.005	0.427	0.451	0.949
		β_2	−0.057	0.183	0.208	0.950	−0.021	0.123	0.132	0.950
		β_3	0.082	0.344	0.366	0.955	0.023	0.232	0.247	0.949
		α	−0.206	0.518	0.563	0.943	−0.124	0.361	0.397	0.946
		ν	0.111	0.321	0.348	0.943	0.073	0.223	0.237	0.945
Cure 2	(7, 4)	β_0	0.043	0.471	0.501	0.945	0.021	0.323	0.323	0.953
		β_1	0.005	0.590	0.600	0.955	−0.003	0.408	0.434	0.953
		β_2	−0.035	0.139	0.147	0.950	−0.015	0.095	0.100	0.950
		β_3	0.058	0.268	0.288	0.941	0.028	0.183	0.191	0.945
		α	−0.276	0.820	0.889	0.940	−0.132	0.569	0.590	0.949
		ν	0.137	0.431	0.458	0.947	0.069	0.298	0.309	0.952
	(5, 3)	β_0	0.051	0.421	0.433	0.948	0.020	0.291	0.288	0.952
		β_1	−0.019	0.534	0.552	0.956	−0.001	0.369	0.390	0.951
		β_2	−0.026	0.124	0.127	0.964	−0.013	0.086	0.087	0.950
		β_3	0.045	0.240	0.248	0.951	0.022	0.165	0.171	0.950
		α	−0.146	0.514	0.534	0.944	−0.078	0.358	0.360	0.953
		ν	0.090	0.311	0.324	0.948	0.051	0.216	0.220	0.952
Cure 3	(7, 4)	β_0	0.005	0.420	0.414	0.957	0.004	0.290	0.300	0.949
		β_1	0.000	0.542	0.541	0.957	−0.003	0.375	0.387	0.950
		β_2	−0.024	0.116	0.119	0.956	−0.011	0.079	0.083	0.953
		β_3	0.050	0.227	0.234	0.951	0.013	0.155	0.160	0.950
		α	−0.192	0.765	0.814	0.957	−0.132	0.534	0.567	0.953
		ν	0.096	0.401	0.421	0.956	0.071	0.279	0.297	0.948
	(5, 3)	β_0	0.013	0.376	0.387	0.954	0.013	0.261	0.275	0.948
		β_1	−0.010	0.490	0.500	0.950	−0.005	0.341	0.336	0.950
		β_2	−0.018	0.104	0.110	0.955	−0.012	0.072	0.074	0.948
		β_3	0.029	0.203	0.211	0.946	0.017	0.141	0.145	0.950
		α	−0.129	0.479	0.515	0.939	−0.061	0.334	0.339	0.953
		ν	0.081	0.289	0.307	0.951	0.035	0.201	0.200	0.950

Table 4. Rejection rate (in %) of the Bellcr model when compared with different models using the Vuong’s test statistic—Part 1.

True Model	$(\mathbb{E}(Z), \text{Var}(Z))$	Model Compared with Bellcr	$n = 200$			$n = 400$		
			Cure Rate			Cure Rate		
			Cure 1	Cure 2	Cure 3	Cure 1	Cure 2	Cure 3
Bellcr	(7, 4)	Pocr	2.7	2.4	2.9	1.8	2.1	1.6
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	4.2	3.3	3.7	3.8	3.6	3.1
		Geocr	0.4	0.3	0.2	0.0	0.0	0.5
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(5, 3)	Pocr	1.7	2.1	2.8	1.0	1.0	1.8
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	3.6	2.9	3.5	2.3	2.2	2.8
		Geocr	0.2	0.6	0.5	0.0	0.0	0.3
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(2, 1)	Pocr	1.3	1.5	2.8	0.7	0.9	0.7
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	2.7	2.7	3.7	2.3	2.0	1.6
		Geocr	0.5	0.7	0.8	0.0	0.0	0.1
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
Pocr	(7, 4)	Pocr	13.6	11.4	11.8	16.6	13.6	14.2
		Locr	0.0	0.0	0.1	0.0	0.0	0.0
		NBcr	16.1	12.4	12.9	19.6	15.5	16.0
		Geocr	0.2	0.1	0.0	0.0	0.0	0.0
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(5, 3)	Pocr	12.9	14.5	11.7	21.7	17.2	15.2
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	14.9	15.3	12.2	24.4	18.6	17.2
		Geocr	0.0	0.2	0.1	0.0	0.0	0.0
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(2, 1)	Pocr	15.6	15.0	13.9	20.9	18.3	15.1
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	17.2	16.4	14.6	22.8	20.1	16.8
		Geocr	0.1	0.0	0.1	0.0	0.0	0.0
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
Locr	(7, 4)	Pocr	6.2	1.6	0.4	9.5	0.0	0.3
		Locr	79.9	79.7	79.6	80.9	80.8	80.8
		NBcr	5.9	5.0	6.0	9.5	2.3	6.3
		Geocr	0.5	3.7	9.7	0.0	4.5	11.3
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(5, 3)	Pocr	6.8	0.8	0.7	6.7	0.6	0.2
		Locr	76.1	75.9	75.8	80.5	80.4	80.3
		NBcr	7.4	3.4	4.8	6.7	3.9	7.9
		Geocr	0.3	4.3	9.4	0.1	4.4	14.7
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(2, 1)	Pocr	5.7	0.9	0.4	7.2	0.3	0.1
		Locr	80.0	79.9	79.7	78.6	78.5	78.6
		NBcr	6.1	4.7	6.7	7.2	4.1	10.9
		Geocr	0.5	6.7	10.7	0.1	6.3	18.2
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0

Table 5. Rejection rate (in %) of the Bellcr model when compared with different models using the Vuong’s test statistic—Part 2.

True Model	$(\mathbb{E}(Z), \text{Var}(Z))$	Model Compared with Bellcr	$n = 200$			$n = 400$		
			Cure Rate			Cure Rate		
			Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
NBcr ($q = 3$)	(7, 4)	Pocr	0.9	2.3	3.4	0.4	1.4	1.6
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	4.8	3.5	5.3	6.5	4.6	3.9
		Geocr	0.8	1.4	0.6	0.3	0.5	0.5
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(5, 3)	Pocr	0.8	2.5	3.0	0.4	0.4	1.5
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	7.2	4.8	4.2	11.2	3.4	4.1
		Geocr	1.7	1.0	0.7	0.5	0.3	0.7
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(2, 1)	Pocr	0.7	1.4	2.8	0.1	1.0	1.7
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	5.0	4.2	3.9	9.8	4.6	4.7
		Geocr	1.0	0.8	0.5	0.6	0.2	0.5
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
Geocr	(7, 4)	Pocr	0.0	0.0	0.0	0.0	0.0	0.0
		Locr	0.4	0.8	0.5	0.1	0.1	0.4
		NBcr	64.7	19.4	13.3	92.4	49.8	34.3
		Geocr	58.7	25.7	18.8	83.5	49.1	38.8
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(5, 3)	Pocr	0.0	0.0	0.0	0.0	0.0	0.0
		Locr	0.0	0.1	0.3	0.0	0.0	0.1
		NBcr	69.8	25.8	15.8	94.8	55.6	36.8
		Geocr	62.5	30.8	22.1	87.4	52.0	39.5
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0
	(2, 1)	Pocr	0.0	0.0	0.0	0.0	0.0	0.0
		Locr	0.0	0.0	0.0	0.0	0.0	0.0
		NBcr	72.2	29.4	17.3	96.2	60.4	40.7
		Geocr	63.9	33.4	24.2	89.6	56.7	42.7
		Berncr	0.0	0.0	0.0	0.0	0.0	0.0

5. Data Analysis

In this section, we conduct an analysis of a dataset on melanoma available at the timereg R package [22]. The study includes 205 patients, with survival times observed after an operation for removal of a malignant melanoma. The observed times vary from 10 to 5565 days (from 0.03 to 15.24 years), with mean and median 5.89 and 5.49 years, respectively, and standard deviation 3.07 years. The dataset comprises 148 censored observations (72.2%), corresponding to patients who died from other causes or were still alive at the end of the study. The dataset includes the covariates ulceration status (x_1) (absent, $n = 115$; present, $n = 90$) and tumor thickness (x_2) (in mm, mean = 2.91, median = 1.94 and standard deviation = 2.96).

We consider the Bell cure rate model for analyzing this dataset. For comparison purposes, we also consider the Pocr, Locr, NBcr, Geocr and Bincr models. We assume the Weibull distribution for the time-to-event for the concurrent causes. First we assess the goodness of fit of the models. For such purpose, we use the Cox-Snell residual, see, e.g., [23] and the normalized quantile residual [24]. In Figure 3, we see that the Bellcr, Pocr, Locr and Geocr models yield a reasonable fit. However, the deviation from the identity line in both residuals plot is evident for the NBcr model. We omit the results for the Bincr model

because the plots are similar to the ones for the NBcr model. For this reason, we disregard the NBcr and Bincr models. The Cox-Snell residuals highlight observations 200–205, which correspond to the largest censored times. On the other hand, the normalized quantile residuals do not suggest the presence of possible poorly fitted observations. Table 6 displays parameter estimates and odds ratio for the Bellcr, Locr, Pocr and Geocr models. At a 5% significance level, the coefficients of ulceration status (β_1) and tumor thickness (β_2) are significant and the negative sign is as expected. Moreover, the estimates of ν indicate that the exponential distribution for the latent times in Section 1 is not adequate.

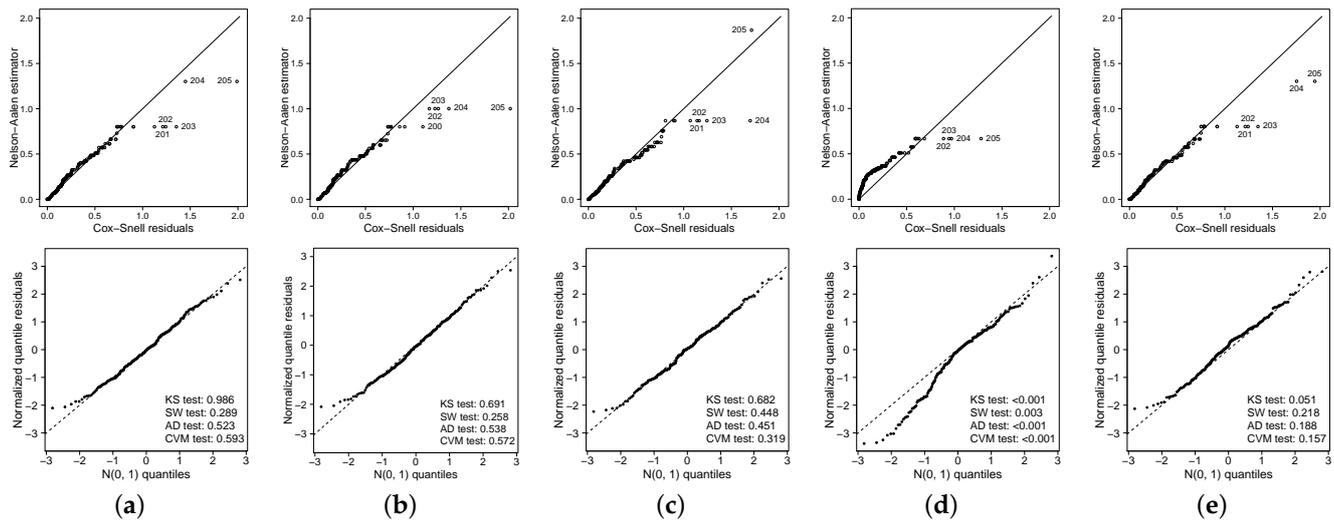


Figure 3. Cox-Snell residuals (upper panels) and normalized quantile residuals plots (lower panels): (a) Bellcr, (b) Pocr, (c) Locr, (d) NBcr and (e) Geocr models.

Table 6. Parameter estimates, standard errors (SE) and maximum value of the log-likelihood function ($\ell(\hat{\psi})$) for different models.

Parameter	Model							
	Bellcr		Pocr		Locr		Geocr	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
β_0	1.8759	0.2449	1.9039	0.3431	1.6771	0.3566	1.8127	0.3501
β_1	−1.4536	0.2843	−1.4814	0.3901	−1.4816	0.3278	−1.4807	0.3569
β_2	−0.1908	0.0416	−0.1960	0.0705	−0.1412	0.0356	−0.1785	0.0537
α	−3.2461	0.3124	−2.9435	0.3411	−4.1863	0.5154	−3.4907	0.3916
ν	1.8287	0.1808	1.7368	0.2178	2.2110	0.2725	1.9132	0.2357
$\exp(\beta_1)$	0.2337	0.0664	0.2273	0.0887	0.2273	0.0745	0.2275	0.0812
$\exp(\beta_2)$	0.8263	0.0344	0.8220	0.0580	0.8683	0.0309	0.8365	0.0449
$\ell(\hat{\psi})$	−206.3		−207.5		−203.7		−205.4	

Note that for all the models, $\exp(\hat{\beta}_1) \approx 0.23$. Therefore, patients with ulceration have approximately their probability of being cured reduced by a quarter in relation with patients without ulceration.

Table 7 shows the Vuong’s statistic for the fitted models. At a 5% significance level, we conclude that the Bellcr model is preferred to the Pocr model. Moreover, there is no significant difference among the Bellcr and the Locr and Geocr models.

Table 7. Vuong’s statistic and p -value when testing the Bell cure rate model against some alternatives.

Model	Pocr	Locr	Geocr
Vuong’s statistic	−2.473	1.238	1.146
p -Value	0.013	0.216	0.252

Figure 4 shows the estimates of the mean of the conditional distribution of $M_i | t_i, \delta_i; \hat{\psi}$, say \tilde{M}_i , for $i = 1, \dots, n$. As previously discussed, values closer to 0 suggest which individuals are cured. Note that the Bellcr and Locr models provide very similar values. For instance, the observation 6 corresponds to a patient who died after 204 days (0.56 years), with ulceration and a tumor thickness of 4.84 cm., providing \tilde{M}_i equal to 2.988, 2.985 and 2.156 for Bellcr, Locr and Geocr models, respectively. Therefore, the three models suggest that this observation has a greater number of carcinogenic cells (a susceptible individual). Similarly, the observation 149 corresponds to a patient who died after 2782 days (7.62 years), without ulceration and a tumor thickness of 1.94 cm., providing \tilde{M}_i equal to 1.076, 1.076 and 1.453 for Bellcr, Locr and Geocr models. This explains why the patient 6 fails in a time considerably lower than patient 149. In a similar way, the observation 203 has a censored time after 4688 days (12.84 years), without ulceration and a tumor thickness of 0.48 cm., resulting in \tilde{M}_i equal to 0.002, 0.002 and 0.015 for the three models. In this case, the models agree that this patients is a potentially cured individual.

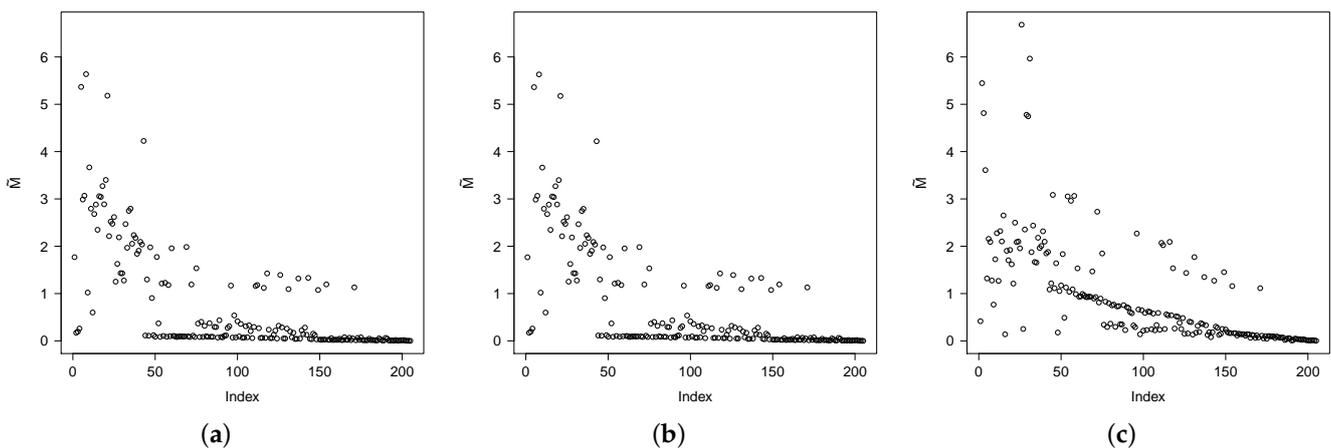


Figure 4. Estimate of the mean of the conditional distribution of $M_i | t_i, \delta_i; \hat{\psi}$ in melanoma data set using (a) Bellcr, (b) Locr and (c) Geocr models.

Finally, Figure 5 presents the width of the 95% confidence intervals for the conditional cure rate in (9) and Figure 6 displays estimates of cure rate under the Bellcr, Locr and Geocr models. In all models, the confidence intervals are computed using the delta method [25] (Theorem 3.4.5) using expressions of the covariance matrix of the estimators in the power series cure rate model presented in Gallardo et al. [14]. For the Locr and Geocr models, the conditional cure rate is computed as in Remark 1. Note that the Bellcr model provides more accurate confidence intervals for some settings in Figure 5. Finally, Figure 6 shows the conditional cure rate in Equation (9) for different in terms of the tumor thickness for $t_0 = 0$ and $t_0 = 5$ years and patients with and without ulceration.

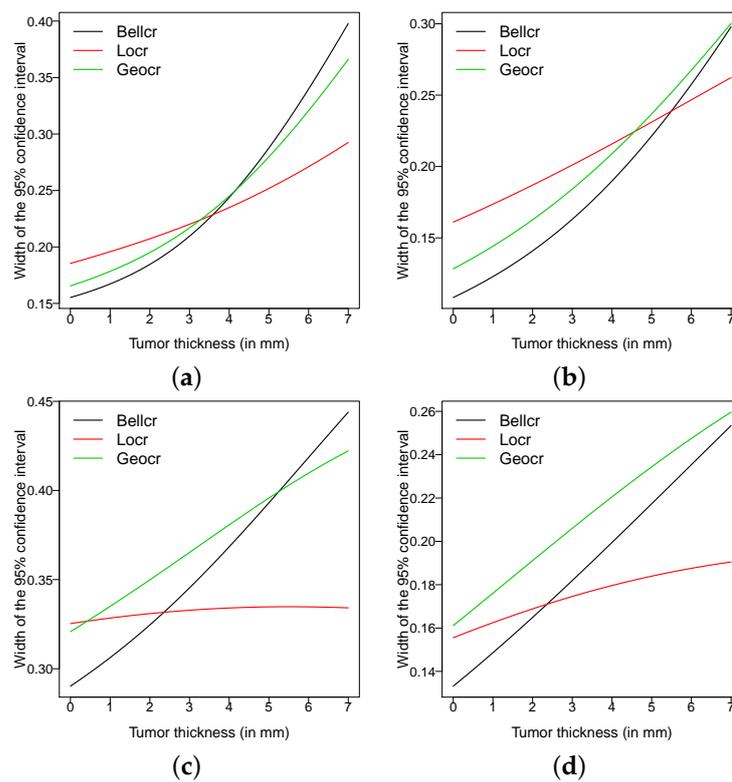


Figure 5. Width of the 95% confidence interval for the conditional cure rate: (a) without ulceration and $t_0 = 0$ year, (b) without ulceration and $t_0 = 5$ years, (c) with ulceration and $t_0 = 5$ years and (d) with ulceration and $t_0 = 10$ years.

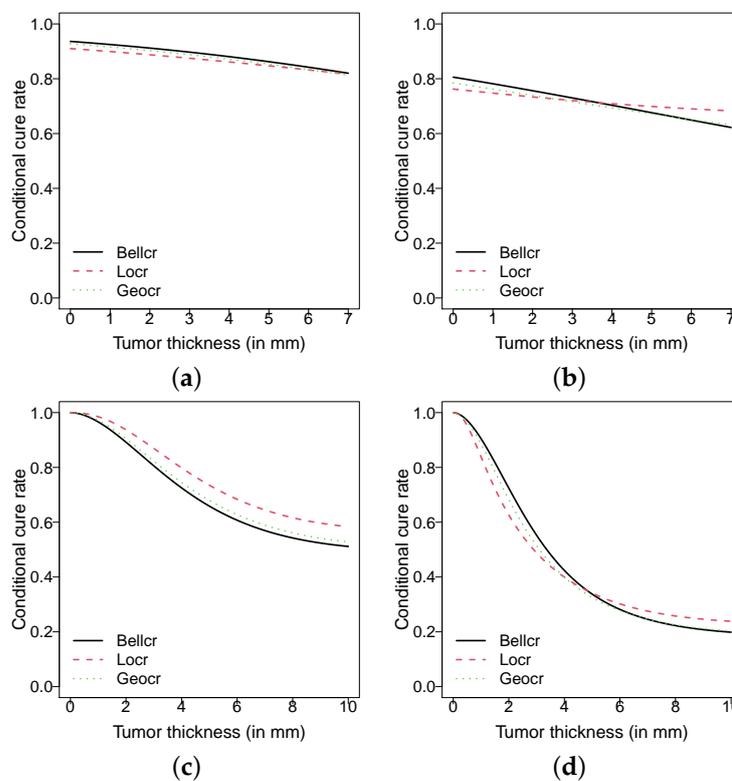


Figure 6. Estimates of the conditional cure rate (p_{t_0} in Equation (9)) for $t_0 = 5$ years: (a) without ulceration and (b) with ulceration. Estimated survival function for a tumor thickness of 10 cms: (c) without ulceration and (d) with ulceration.

6. Conclusions

In this paper, a cure rate model is proposed under the competing risks setup. For the number of competing causes of the event of interest, we posit the Bell distribution. This is a one-parameter distribution that accommodates overdispersed counts. [26] emphasized that the Poisson distribution represents a strong assumption when modelling the number of competing causes. Therefore, compared to the two-parameter negative binomial distribution, the proposed model is more parsimonious. The cure rate is a parameter linked to covariates, facilitating the comparison with other models. Parameter estimates are computed through iterative steps of the EM algorithm. In order to compare different models, the test statistic proposed in [21] is implemented. In our simulation studies, the estimation method and the test statistic have a good performance.

A dataset on melanoma is analyzed using the proposed model as well as some models from the literature. Extensions of the proposed model to accommodate other types of censoring (e.g., interval censoring) might be a theme for future work.

Author Contributions: Conceptualization, D.I.G. and M.d.C.; Formal analysis, D.I.G. and M.d.C.; Investigation, D.I.G., M.d.C. and H.W.G.; Methodology, D.I.G. and M.d.C.; Software, D.I.G. and M.d.C.; Supervision, D.I.G., M.d.C. and H.W.G.; Validation, D.I.G., M.d.C. and H.W.G.; All of the authors contributed significantly to this research article. All authors have read and agreed to the published version of the manuscript.

Funding: The research of H.W. Gómez was supported by SEMILLERO UA-2021 project, Chile. The work of M.d.C. was partially funded by CNPq, Brazil.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in Section 5 is duly referenced.

Acknowledgments: We thank the editors and the anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Model Selection

Now we present the Vuong's statistic [21] to test the hypothesis H_0 : The data are generated from the Bellcr model against H_1 : The data are generated from an alternative cure rate model. Let $\hat{\psi}_0$ and $\hat{\psi}_1$ denote the ML estimators of the model parameter ψ under H_0 and H_1 , respectively. The test statistic is given by $V = n^{-1/2}\xi/\omega$, where $\xi = \sum_{i=1}^n \log[f_i(\hat{\psi}_0)/g_i(\hat{\psi}_1)]$ and $\omega = \{n^{-1} \sum_{i=1}^n \log^2[f_i(\hat{\psi}_0)/g_i(\hat{\psi}_1)] - \xi^2\}^{1/2}$, with

$$f_i(\psi_0) = f_{pop}(t_i; p, \lambda)^{\delta_i} S_{pop}(t_i; p, \lambda)^{1-\delta_i},$$

where S_{pop} and f_{pop} are given in (3) and (4), respectively and $g_i(\psi_1)$ denotes the contribution of the i -th individual in the likelihood function under H_1 (Pocr, Locr, NBcr, Geocr or Bincr model).

References

1. Boag, J.W. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *J. R. Stat. Soc. B* **1949**, *35*, 15–53. [[CrossRef](#)]
2. Berkson, J.; Gage, R.P. Survival curve for cancer patients following treatment. *J. Am. Stat. Assoc.* **1952**, *47*, 501–515. [[CrossRef](#)]
3. Yakovlev, A.Y.; Tsodikov, A.D. *Stochastic Models of Tumor Latency and Their Biostatistical Applications*; World Scientific: Singapore, 1996.
4. Rodrigues, J.; Cancho, V.G.; de Castro, M.; Louzada-Neto, F. On the unification of the long-term survival models. *Stat. Probab. Lett.* **2009**, *79*, 753–759. [[CrossRef](#)]
5. Rodrigues, J.; de Castro, M.A.; Cancho, V.G.; Balakrishnan, N. COM-Poisson cure rate survival model and an application to a cutaneous melanoma data. *J. Stat. Plan. Inference* **2009**, *139*, 3605–3611. [[CrossRef](#)]
6. Cancho, V.G.; Louzada, F.; Ortega, E.M. The power series cure rate model: An application to a cutaneous melanoma data. *Commun. Stat. Simul. Comput.* **2013**, *42*, 586–602. [[CrossRef](#)]

7. Gallardo, D.I.; Gómez, H.W.; Bolfarine, H. A new cure rate model based on the Yule-Simon distribution with application to a melanoma data set. *J. Appl. Stat.* **2017**, *44*, 1153–1164. [[CrossRef](#)]
8. Gallardo, D.I.; Gómez, Y.M.; De Castro, M. A flexible cure rate model based on the polylogarithm distribution. *J. Stat. Comput. Simul.* **2018**, *88*, 2137–2149. [[CrossRef](#)]
9. Gallardo, D.I.; Gómez, Y.M.; Gómez, H.W.; De Castro, M. On the use of the modified power series family of distributions in a cure rate model context. *Stat. Methods Med. Res.* **2020**, *29*, 1831–1845. [[CrossRef](#)] [[PubMed](#)]
10. Cancho, V.G.; Macera, M.A.C.; Suzuki, A.K.; Louzada, F.; Zavaleta, K.E.C. A new long-term survival model with dispersion induced by discrete frailty. *Lifetime Data Anal.* **2020**, *26*, 221–244. [[CrossRef](#)]
11. Leao, J.; Bourguignon, M.; Gallardo, D.I.; Rocha, R.; Tomazella, V. A new cure rate model with flexible competing causes with applications to melanoma and transplanted data. *Stat. Med.* **2020**, *24*, 3272–3284. [[CrossRef](#)] [[PubMed](#)]
12. Castellares, F.; Lemonte, A.J.; Santos, M.A.C. On the Nielsen distribution. *Braz. J. Probab. Stat.* **2020**, *1*, 90–111. [[CrossRef](#)]
13. Noack, A. A class of random variables with discrete distributions. *Ann. Math. Stat.* **1950**, *21*, 127–132. [[CrossRef](#)]
14. Gallardo, D.I.; Romeo, J.S.; Meyer, R. A simplified estimation procedure based on the EM algorithm for the power series cure rate model. *Commun. Stat. Simul. Comput.* **2017**, *46*, 6342–6359. [[CrossRef](#)]
15. Tsodikov, A.D.; Ibrahim, J.G.; Yakovlev, A.Y. Estimating cure rates from survival data: An alternative to two-component mixture models. *J. Am. Stat. Assoc.* **2003**, *98*, 1063–1078. [[CrossRef](#)] [[PubMed](#)]
16. Cancho, V.G.; De Castro, M.; Dey, D.K. Long-term survival models with latent activation under a flexible family of distributions. *Braz. J. Probab. Stat.* **2013**, *27*, 585–600. [[CrossRef](#)]
17. Chen, M.-H.; Ibrahim, J.G.; Sinha, D. A new Bayesian model for survival data with a surviving fraction. *J. Am. Stat. Assoc.* **1999**, *94*, 909–919. [[CrossRef](#)]
18. Corless, R.M.; Gonnet, G.H.; Hare, D.E.G.; Jeffrey, D.J.; Knuth, D.E. On the Lambert W function. *Adv. Comput. Math.* **1996**, *5*, 329–359. [[CrossRef](#)]
19. R Development Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
20. Gilbert, P.; Varadhan, R. numDeriv: Accurate Numerical Derivatives. R Package Version. Available online: <http://optimizer.r-forge.r-project.org/> (accessed on 3 July 2021).
21. Vuong, Q.H. Likelihood Ratio Tests for Model Selection and non-nested Hypotheses. *Econometrica* **1989**, *57*, 307–333. [[CrossRef](#)]
22. Scheike, T.H.; Zhang, M.J. Analyzing competing risk data using the R timereg package. *J. Stat. Softw.* **2011**, *38*, 1–15. [[CrossRef](#)]
23. Conlon, A.S.C.; Taylor, J.M.G.; Sargent, D.J. Multi-state models for colon cancer recurrence and death with a cure fraction. *Stat. Med.* **2014**, *33*, 1750–1766. [[CrossRef](#)]
24. Dunn, P.K.; Smyth, G.K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244.
25. Sen, P.K.; Singer, J.M. *Large Sample Methods in Statistics: An Introduction with Applications*; Chapman & Hall: New York, NY, USA, 1993.
26. Tournoud, M.; Ecochard, R. Promotion time models with time-changing exposure and heterogeneity: Application to infectious diseases. *Biom. J.* **2008**, *50*, 395–407. [[CrossRef](#)] [[PubMed](#)]