

Article

The Slash Half-Normal Distribution Applied to a Cure Rate Model with Application to Bone Marrow Transplantation

Diego I. Gallardo ^{1,*}, Yolanda M. Gómez ^{1,2}, Héctor J. Gómez ³, María José Gallardo-Nelson ²
and Marcelo Bourguignon ⁴

¹ Mathematics Department, Faculty of Engineering, University of Atacama, Copiapó 1530000, Chile

² Medicine Department, Faculty of Medicine, University of Atacama, Copiapó 1530000, Chile

³ Department of Mathematical and Physical Sciences, Faculty of Engineering, Catholic University of Temuco, Temuco 4780000, Chile

⁴ Statistics Department, Federal University of Rio Grande do Norte, Natal 59078-970, Brazil

* Correspondence: diego.gallardo@uda.cl; Tel.: +56-52-2255565

Abstract: This paper proposes, for the first time, the use of an asymmetric positive and heavy-tailed distribution in a cure rate model context. In particular, it introduces a cure-rate survival model by assuming that the time-to-event of interest follows a slash half-normal distribution and that the number of competing causes of the event of interest follows a power series distribution, which defines six new cure rate models. Several properties of the model are derived and an alternative expression for the cumulative distribution function of the model is presented, which is very useful for the computational implementation of the model. A procedure based on the expectation–maximization algorithm is proposed for the parameter estimation. Two simulation studies are performed to assess some properties of the estimators, showing the good performance of the proposed estimators in finite samples. Finally, an application to a bone marrow transplant data set is presented.

Keywords: half-normal distribution; kurtosis; power series; blood and marrow transplantation

MSC: 62F10; 62N02



Citation: Gallardo, D.I.; Gómez, Y.M.; Gómez, H.J.; Gallardo-Nelson, M.J.; Bourguignon, M. The Slash Half-Normal Distribution Applied to a Cure Rate Model with Application to Bone Marrow Transplantation. *Mathematics* **2023**, *11*, 518. <https://doi.org/10.3390/math11030518>

Academic Editor: Stelios Psarakis

Received: 19 December 2022

Revised: 10 January 2023

Accepted: 13 January 2023

Published: 18 January 2023



Copyright: © 2023 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 and Motivation

Survival models that incorporate a proportion of cured subjects are becoming increasingly popular in the analysis of data from clinical trials and other studies, mainly due to advances in medicine. In a cure rate model context, specifically using a competitive risk model framework, it is assumed that the event of interest is produced due to $M \in \mathbb{N}_0$ carcinogenic cells (a latent variable), each of them with an associated survival function $S(\cdot; \lambda)$, where λ is a vector of unknown parameters from a distribution defined on the positive line, and $M = 0$ represents the cure individuals. Different models are assumed for the number of cells (M) and the survival function $S(\cdot; \lambda)$. Specifically for the latter, we can mention Weibull, log-normal, generalized gamma, Birnbaum–Saunders distributions, among others. In particular, we focus on the proposed extension in Olmos et al. [1]. This model (parsimonious, since it only has two parameters) is called a slash half-normal (SHN) distribution. This distribution with a heavy tail has not yet been used in a cure rate model context. The objective of the proposed model is to increase the kurtosis with respect to its half-normal basal distribution, being more useful for modeling data sets that may have heavy tails. A version of the SHN reparameterized in terms of the mean with covariates was presented in Gómez et al. [2].

On the other hand, bone marrow transplantation is an important treatment for more than 50 fatal diseases such as leukemia, lymphoma, solid organ neoplasia, autoimmune diseases, among others. Bone marrow transplantation has unique biological characteristics and associated problems that markedly differ from solid organ transplantation, such as

kidney, liver, or heart (Hardy et al. [3]). The problems are associated with the fact that immunocompetent cells can reject each other, which can cause graft rejection and generate graft-versus-host disease (GVHD). Furthermore, successful engraftment requires a strict match with the major histocompatibility complex (MHC) class II antigen of the donor and the recipient. Added to this is that patients must receive prior treatment with cytotoxic and myeloablative agents to avoid graft rejection (Chinen and Buckley [4], Simpson and Dazzi [5]). For these reasons, bone marrow transplantation is associated with considerable morbidity and mortality.

From a biological point of view, it makes sense to assume a heavy-tailed distribution when following patients over a period of time and following patient survival to assess transplant efficacy. These models can describe in detail the progression of the disease and allow us to make predictions for future clinical trials.

The motivation to introduce the SHN distribution in a cure rate model is given by a real data set application from the European Society for Blood and Marrow Transplantation (EBMT). The EBMT is a nonprofit medical, educational and scientific organization founded more than forty years ago, focused in the field of hematopoietic stem cell transplantation. The EBMT has more than 5000 members from 85 countries. This organization is focused on innovation, research and the advancement in the field of hematopoietic stem cell transplantation. The EBMT hosts a unique patient registry providing a pool of data to researchers all over the world to perform studies and assess new improvements in the field (EBTM [6]). Acute myeloid leukemia (AML) in children is rare in comparison with acute lymphoid leukemia and has overall survival of almost 75% at five years. Patients with high-risk characteristics need to achieve remission and undergo bone marrow transplantation with an overall survival of 70% (de Rooij et al. [7]). Post-transplant relapse is common and involves almost 80% of treated AML, of which 75% occur in the first year after transplantation (Sander et al. [8]).

The data were presented in Fiocco et al. [9] and are available in the `mstate` R package [10], labeled as `ebmt4`. The data set includes 2279 patients transplanted between 1985 and 1998. We considered the time in years from transplantation to relapse as the response variable, which had 84% of censored times. Figure 1 shows the Kaplan–Meier (KM) estimator for this data set and the histogram for the failure times.

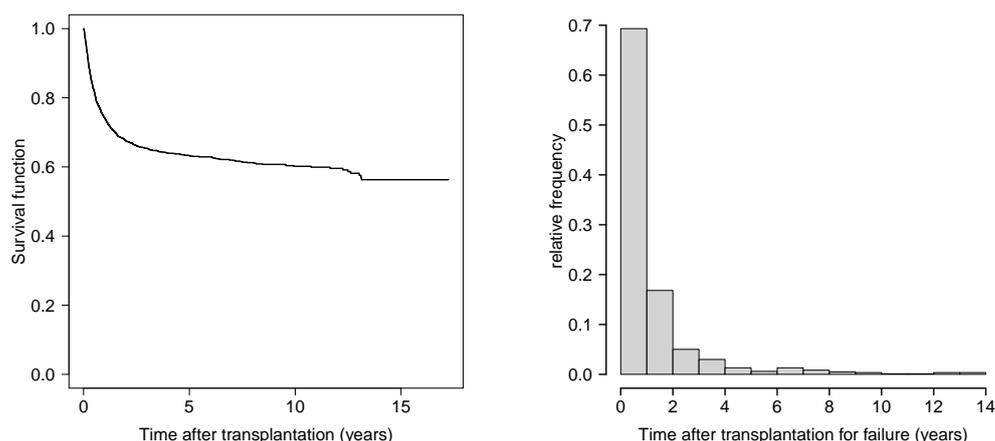


Figure 1. Kaplan–Meier estimator for the `ebmt4` data set (left panel) and histogram for the failure times (right panel).

Note that there is a plateau in the KM estimator, suggesting the presence of cure individuals for this problem. In addition, the histogram of failure times suggest a heavy-tailed distribution for the noncured individuals. Some descriptive statistics for such failure times are mean = 1.14, median = 0.52, standard deviation = 1.84 and kurtosis = 18.71. The high kurtosis for these data, again, suggests the use of a heavy-tailed model for this problem.

Allogeneic stem cell transplantation (allo-HSCT) is a therapeutic approach used for a wide array of malignant and nonmalignant hematologic disorders. Its effectiveness is related to mortality and morbidity. Several complications can occur, including hemorrhages, infections, GVHD and endothelial-damage-related toxicities. There are new immunological and genetic therapies to avoid these complications and reduce mortality (See Horio et al. [11], Radujkovic et al. [12] and Tsai et al. [13]). Other immunodeficiency diseases such as ataxia-telangiectasia can be treated by bone marrow transplantation as mentioned in Sabino Pinho de Oliveira et al. [14]. Further studies with an increased cohort of patients and a longer period of follow-ups are required to validate and monitor the associated risks of transplantation. For this reason, it is important to use a heavy-tailed distribution such as the SHN distribution to study the survival in this type of bone marrow transplant and therapies. We emphasize that, up to this moment, this distribution has not been used in a cure rate model context. In this context, we define the power series–SHN (PS–SHN) class of univariate cure rate models obtained by compounding the SHN and power series distributions. The proposed model defines six new cure rate models as special cases.

The contents in this paper are organized in the following manner. In Section 2, we introduce the SHN distribution and derive some of its distributional properties. In Section 3, we introduce a general cure rate model. In Section 4, we describe a parameter estimation by the maximum likelihood method using the expectation–maximization (EM) algorithm. In Section 5, we present two simulation studies, where the first one is related to assessing the performance of the maximum likelihood estimation (MLE) using the EM algorithm and the second one is the misspecification for the time-to-event for the concurrent causes. In Section 6, we present a real data set on bone marrow transplantation. Finally, some conclusions are included in Section 7.

2. The SHN Distribution

The SHN distribution [1] with parameters $\sigma > 0$ and $\alpha > 0$ (denoted by $SHN(\sigma, \alpha)$) is a model with positive support. Its probability density function (pdf) is given by

$$f_{SHN}(t; \lambda) = \alpha \sqrt{\frac{2^\alpha}{\pi}} \sigma^\alpha t^{-(\alpha+1)} \gamma\left(\frac{t^2}{2\sigma^2}, \frac{\alpha+1}{2}\right), \quad t > 0, \tag{1}$$

where $\lambda = (\alpha, \sigma)$ and $\gamma(x, a) = \int_0^x u^{a-1} e^{-u} du$ is the lower incomplete gamma function. The cumulative distribution function (cdf) of the SHN model is given by

$$F_{SHN}(t; \lambda) = \int_0^t f_{SHN}(w; \lambda) dw = \alpha \sqrt{\frac{2^\alpha}{\pi}} \sigma^\alpha \int_0^t w^{-(\alpha+1)} \gamma\left(\frac{w^2}{2\sigma^2}, \frac{\alpha+1}{2}\right) dw.$$

In the next Proposition, we present an alternative expression for the cdf.

Proposition 1. *The cdf of the SHN distribution is given by*

$$F_{SHN}(t; \lambda) = 2\Phi\left(\frac{t}{\sigma}\right) - 1 - \sqrt{\frac{2^\alpha}{\pi}} \left(\frac{\sigma}{t}\right)^\alpha \Gamma\left(\frac{\alpha+1}{2}\right) G\left(\frac{w^2}{2\sigma^2}, \frac{\alpha+1}{2}\right), \quad t > 0,$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal model and $G(\cdot, a)$ is the cdf of the gamma model with shape and rate parameters equal to a and 1, respectively.

Proof. By definition

$$F_{SHN}(t; \lambda) = \alpha \sqrt{\frac{2^\alpha}{\pi}} \sigma^\alpha \int_0^t w^{-(\alpha+1)} \Gamma\left(\frac{\alpha+1}{2}\right) G\left(\frac{w^2}{2\sigma^2}, \frac{\alpha+1}{2}\right) dw.$$

Using techniques of integration by parts, we have that

$$u = G\left(\frac{w^2}{2\sigma^2}, \frac{\alpha + 1}{2}\right) \Rightarrow du = \frac{w}{\sigma^2} g\left(\frac{w^2}{2\sigma^2}, \frac{\alpha + 1}{2}\right) dw = \frac{w}{\sigma^2 \Gamma\left(\frac{\alpha + 1}{2}\right)} \left(\frac{w^2}{2\sigma^2}\right)^{\frac{\alpha + 1}{2} - 1} \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw$$

$$dv = w^{-(\alpha + 1)} \Rightarrow v = -\frac{w^{-\alpha}}{\alpha},$$

where $g(\cdot, \cdot)$ denotes the density function of the gamma distribution. Therefore,

$$F_{SHN}(t; \lambda) = \alpha \sqrt{\frac{2^\alpha}{\pi}} \sigma^\alpha \Gamma\left(\frac{\alpha + 1}{2}\right) \times \left[-G\left(\frac{t^2}{2\sigma^2}, \frac{\alpha + 1}{2}\right) \frac{t^{-\alpha}}{\alpha} + \frac{1}{\Gamma\left(\frac{\alpha + 1}{2}\right) \alpha \sigma^\alpha} \sqrt{\frac{\pi}{2^\alpha}} \int_0^t \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw \right].$$

and the result follows after identifying the pdf for the normal distribution within the integral. \square

Corollary 1. *The survival function for the SHN model is given by*

$$S_{SHN}(t; \lambda) = 2\Phi\left(-\frac{t}{\sigma}\right) + \sqrt{\frac{2^\alpha}{\pi}} \left(\frac{\sigma}{t}\right)^\alpha \Gamma\left(\frac{\alpha + 1}{2}\right) G\left(\frac{t^2}{2\sigma^2}, \frac{\alpha + 1}{2}\right), \quad t > 0. \quad (2)$$

Note that Corollary 1 provides an alternative way to program the survival function, especially from a computational point of view, because many software tools include the cdf of the standard normal and gamma models and the lower incomplete gamma function.

The SHN distribution has a convenient stochastic representation. If $X \sim HN(\sigma)$ and $Y \sim U(0, 1)$ are independent (i.e., the half-normal and uniform distributions, respectively), then

$$T = \frac{X}{Y^{1/\alpha}} \sim SHN(\sigma, \alpha). \quad (3)$$

The mean, variance, asymmetry ($\sqrt{\beta_1}$) and kurtosis (β_2) coefficients from this distribution are

- $\mathbb{E}(T) = \sqrt{\frac{2}{\pi}} \frac{\alpha \sigma}{\alpha - 1}, \quad \alpha > 1.$
- $\text{Var}(T) = \frac{\alpha \sigma^2 [(\pi - 2)\alpha(\alpha - 2) + \pi]}{\pi(\alpha - 1)^2(\alpha - 2)}, \quad \alpha > 2.$
- $\sqrt{\beta_1} = \frac{\pi \sqrt{2(\alpha - 2)} \left[\frac{4}{\pi} \alpha^2 (\alpha - 2)(\alpha - 3) - (\alpha - 1)^2 (\alpha - 4)(\alpha + 1) \right]}{\sqrt{\alpha} (\alpha - 3) [\alpha(\pi - 2)(\alpha - 2) + \pi]^{3/2}}, \quad \alpha > 3.$
- $\beta_2 = \frac{3\alpha(\alpha - 2)^2(\alpha - 3) [\pi^2(\alpha - 1)^4 - 4\alpha^3(\alpha - 4)] - 4\pi\alpha^2(\alpha - 1)^2(\alpha - 2)(\alpha - 4)(\alpha^2 - 3\alpha + 8)}{\alpha^2(\alpha - 3)(\alpha - 4) [\alpha(\pi - 2)(\alpha - 2) + \pi]^2}, \quad \alpha > 4.$

Note that σ is a scale parameter and α is related mainly with the kurtosis of the model. In fact, as discussed in Olmos et al. [1], low values for α allow us to accommodate distributions with greater kurtosis coefficients. On the other hand, when $\alpha \rightarrow \infty$, the HN distribution is recovered. Besides being a parsimonious model, an interesting property from the SHN distribution is that its tails are heavier than common distributions with positive support, as we enunciate in the following proposition.

Proposition 2. *Suppose a distribution with positive support satisfying the following conditions:*

- (A1) *Its density function can be written in the form $f(t; \lambda) = Ct^a \exp\{-g(t; \lambda)\}$, $t > 0$, where C does not depend on t , $a \in \mathbb{R}$ and $g(t; \lambda) \geq 0, \forall t > 0$.*
- (A2) $\lim_{t \rightarrow \infty} t^{a'} \exp\{-g(t; \lambda)\} = 0, \quad \forall a' > a + 1.$

Then, the SHN model has heavier tails than such a distribution.

Proof. To show that the SHN model has heavier tails than a model with a density function given as in assumption A1, it is enough to prove that

$$\lim_{t \rightarrow +\infty} \frac{f(t; \lambda)}{f_{SHN}(t; \sigma, \alpha)} = 0.$$

Therefore,

$$\begin{aligned} \lim_{t \rightarrow +\infty} \frac{f(t; \lambda)}{f_{SHN}(t; \sigma, \alpha)} &= \lim_{t \rightarrow +\infty} \frac{Ct^a \exp\{-g(t; \lambda)\}}{\alpha\sigma^\alpha \sqrt{\frac{2^\alpha}{\pi}} t^{-(\alpha+1)} \gamma\left(\frac{t^2}{2\sigma^2}, \frac{\alpha+1}{2}\right)} \\ &= \frac{C}{\alpha\sigma^\alpha \sqrt{\frac{2^\alpha}{\pi}}} \lim_{t \rightarrow +\infty} \frac{t^{a+\alpha+1} \exp\{-g(t; \lambda)\}}{\gamma\left(\frac{t^2}{2\sigma^2}, \frac{\alpha+1}{2}\right)}. \end{aligned}$$

As $\lim_{t \rightarrow +\infty} \gamma(t, c') = \Gamma(c'), \forall c' > 0, \alpha > 0$, by assumption A2 and by the continuity of the limit, we have the result. \square

Corollary 2. *If X_1 and X_2 are two random variables with density functions satisfying conditions A1 and A2 in Proposition 2 and $Y = pX_1 + (1 - p)X_2, p \in (0, 1)$, then the SHN model also has heavier tails than the distribution related to Y .*

Remark 1. *Common distributions used in a cure rate model context satisfy the conditions A1 and A2 in Proposition 2. For instance:*

- (a) *If $C = \alpha/\sigma^\alpha, a = \alpha - 1$ and $g(t; \lambda) = (t/\sigma)^\alpha$, the Weibull model is recovered (Chen et al. [15], Rodrigues et al. [16], Pal and Balakrishnan [17]).*
- (b) *If $C = [\Gamma(\alpha)\sigma^\alpha]^{-1}, a = \alpha - 1$ and $g(t; \lambda) = t/\sigma$, the gamma distribution is recovered (Balakrishnan and Pal [18], Wiangnak and Pal. [19], Ortega et al. [20]).*
- (c) *If $C = [\sigma\sqrt{2\pi}]^{-1}, a = -1$ and $g(t; \lambda) = (\log t - \alpha)^2/(2\sigma^2)$, we obtain the log-normal model (Balakrishnan and Pal [21], Gallardo et al. [22]).*
- (d) *If $C = \sqrt{\sigma/(2\pi)} \exp(\sigma/\alpha), a = -3/2$ and $g(t; \lambda) = (\sigma/2)[t/\alpha^2 + 1/t]$, the inverse Gaussian model is recovered. We denote it as $IG(\alpha, \sigma)$.*
- (e) *If $C = \sqrt{\sigma/(2\pi)} \exp(\sigma\alpha), a = -1/2$ and $g(t; \lambda) = (\sigma/2)[t/\alpha^2 + 1/t]$, the density function of the reciprocal of an inverse Gaussian random variable is recovered. Up to this moment, this model has not been used in a cure rate model context. We denote it as $RIG(\alpha, \sigma)$.*
- (f) *If $C = [B(\alpha, \sigma)]^{-1}, a = \alpha - 1$ and $g(t; \lambda) = (\alpha + \sigma) \log(1 + t)$, the density of the beta prime distribution is recovered (Leao et al. [23]).*

Remark 2. *The Birnbaum–Saunders (BS, Birnbaum and Saunders [24,25]) model with density function given by*

$$f(t; \alpha, \sigma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2} \left[\frac{t}{\sigma} + \frac{\sigma}{t} - 2\right]\right) \frac{(t + \sigma)}{2\alpha\sqrt{\sigma t^3}}, \quad t > 0,$$

can be written as a mixture between $IG(\beta, \beta/\alpha^2)$ and $RIG(\beta, 1/(\beta\alpha^2))$ with equal probabilities (see Desmond [26] for details). Therefore, by Corollary 2, the SHN distribution has heavier tails than the BS distribution.

In the next Section, we discuss the use of the SHN model in the presence of long-term survivors.

3. A General Cure Rate Model and the SHN Distribution

To exemplify the use of the SHN distribution in a cure rate model context, we consider a general class of models based on a competitive risk scheme. The reader is referred to Chen et al. [15] for more details. Let M be a random variable denoting the initial number of carcinogenic cells of an individual. Several models have been considered for this scheme, among others, Bernoulli [27], Poisson [15] and negative binomial [16].

Let $W_a, a = 1, \dots, M$ be the random variable representing the time at which the a th concurrent cause produces the event of interest. For instance, in a cancer context, W_a

represents the progressive time for the a th tumor cell. As shown in the Appendix A, a general class of model is produced when it is assumed that M follows a power series distribution (PS, Cancho et al. [28]). As is common in this context, it is also assumed that, conditional on M , the W_a 's are independent identically distributed with a common cumulative distribution function $F(\cdot; \lambda)$ and survival function $S(\cdot; \lambda) = 1 - F(\cdot; \lambda)$, where λ is a vector of unknown parameters. Moreover, it is assumed that the sequence W_1, W_2, \dots , is independent of M . Assuming that only one concurrent cause can produce the of interest, the observable time of the occurrence of the event of interest is given by

$$T = \begin{cases} +\infty & , \text{ if } M = 0, \\ \min(W_1, \dots, W_M) & , \text{ if } M > 0. \end{cases}$$

Under this setting, the population survival function is given by

$$S_{pop}(t; \theta, \lambda) = \frac{A(\theta S(t; \lambda))}{A(\theta)}, \tag{4}$$

where $A(\cdot)$ is related to the PS distribution. Details about the function $A(\cdot)$ for different particular models are presented in Appendix A. From Equation (4), it is immediate that the cure rate of the model is given by $q_0 = a_0 / A(\theta)$ and the survival function for the susceptible individuals is

$$S_{sus}(t; \theta, \lambda) = \frac{a_0 - A(\theta S(t; \lambda))}{a_0 - A(\theta)}. \tag{5}$$

For heterogeneous populations, we consider a set of covariates of dimension $r + 1$ for the i th individual, say $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ir})^\top$ (i.e., considering the intercept term), related to the cure rate term of the model. For the Poisson, binomial and COM-Poisson, such covariates can be introduced as

$$\theta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}), \quad i = 1, \dots, n, \tag{6}$$

where n is the sample size and $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_r)$ is the vector of unknown regression coefficients. For the rest of the discrete models, the covariates can be included as

$$\theta_i = \frac{\exp(\mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i^\top \boldsymbol{\beta})}, \quad i = 1, \dots, n. \tag{7}$$

In the literature, many models have been considered for $S(\cdot; \lambda)$. As discussed in Remark 1, Weibull, gamma and log-normal models are common alternatives, to name a few. However, we remark that, up to this moment, no distribution with heavy tails has been used in this context and in particular, the SHN distribution has not been considered to model the survival function for the concurrent causes. Henceforth, we refer this model as PS-SHN. It is interesting to explore the use of the SHN distribution in a cure rate model context because it is usual that most failures occur until a certain time and from that moment mainly censoring times are observed. However, in diseases, from a certain time, it is considered that the patient is "cured". In practice, and especially in clinical trials with large sample sizes, it may occur that failure times are observed in unusual times, higher than expected.

4. Estimation

In order to estimate the model parameters, we utilize the maximum likelihood (ML) method. Let us consider the situation when the time-to-event is not completely observed and is subject to right-censoring. Let c_i denote the censoring time for the i th individual. We observe $y_i = \min(t_i, c_i)$ and $\delta_i = \mathbb{I}(t_i \leq c_i)$, where $\delta_i = 1$ if t_i is a time-to-event and $\delta_i = 0$

if right-censored, $i = 1, \dots, n$. From n observed vectors $(y_1, \delta_1, \mathbf{x}_1^\top), \dots, (y_n, \delta_n, \mathbf{x}_n^\top)$, the corresponding likelihood function, under noninformative censoring, can be expressed as

$$\begin{aligned} \ell(\boldsymbol{\psi}) &= \sum_{i=1}^n \left\{ \delta_i \log[f_{pop}(y_i; \theta_i, \boldsymbol{\lambda})] + (1 - \delta_i) \log[S_{pop}(y_i; \theta_i, \boldsymbol{\lambda})] \right\} \\ &= \sum_{i=1}^n \left\{ (1 - \delta_i) \log[A(\theta_i S_{SHN}(y_i; \boldsymbol{\lambda}))] - \log[A(\theta_i)] \right. \\ &\quad \left. + \delta_i \left(\log(\theta_i) + \log[f_{SHN}(y_i; \boldsymbol{\lambda})] + \log[A'(\theta_i S_{SHN}(y_i; \boldsymbol{\lambda}))] \right) \right\}, \end{aligned} \tag{8}$$

where $\boldsymbol{\psi} = (\boldsymbol{\beta}^\top, \boldsymbol{\lambda}^\top)^\top$, and $f_{SHN}(\cdot; \cdot)$ and $S_{SHN}(\cdot; \cdot)$ are defined in (1) and (2). Details for the construction of Equation (8) can be seen in Cancho et al. [28] and Gallardo et al. [29].

The MLE of $\boldsymbol{\psi}$, denoted by $\hat{\boldsymbol{\psi}}$, can be obtained by maximizing Equation (8) in relation to $\boldsymbol{\psi}$. However, this procedure can be computationally expensive and it would not necessarily provide a global maximum, especially when the dimension of $\boldsymbol{\beta}$ is high. To facilitate the obtainment of the MLE for $\boldsymbol{\psi}$, we can use the EM algorithm developed in Gallardo et al. [29] for this class of models. The k th iteration of such procedure consists of the following steps:

- E-step: For $i = 1, \dots, n$, compute

$$\tilde{M}_i^{(k)} = (1 - \delta_i) \mathbb{E}\left(M_i; \theta_i^{(k-1)} S_i^{(k-1)}\right) + \delta_i \frac{\mathbb{E}\left(M_i^2; \theta_i^{(k-1)} S_i^{(k-1)}\right)}{\mathbb{E}\left(M_i; \theta_i^{(k-1)} S_i^{(k-1)}\right)},$$

where $\mathbb{E}(M_i; \mu_i)$ and $\mathbb{E}(M_i^2; \mu_i)$ denote the two first moments of the PS distribution with parameter μ_i given by

$$\mathbb{E}(M_i; \mu_i) = \mu_i \frac{\partial \log A(\mu_i)}{\partial \mu_i} \quad \text{and} \quad \mathbb{E}(M_i^2; \mu_i) = \mu_i \frac{\partial \mathbb{E}(M_i; \mu_i)}{\partial \mu_i} + \mathbb{E}^2(M_i; \mu_i).$$

- M1-step: Given $\mathbf{M}^{(k)} = (\tilde{M}_1^{(k)}, \dots, \tilde{M}_n^{(k)})$, find $\boldsymbol{\beta}$ that maximizes $Q_1(\boldsymbol{\beta} \mid \boldsymbol{\psi}^{(k)})$ in relation to $\boldsymbol{\beta}^{(k)}$, where

$$Q_1(\boldsymbol{\beta} \mid \boldsymbol{\psi}^{(k)}) = \sum_{i=1}^n \left\{ \tilde{M}_i^{(k)} \log(\theta_i) - \log[A(\theta_i)] \right\}.$$

- M2-step: Given $\mathbf{M}^{(k)}$, find $\boldsymbol{\lambda}^{(k)}$ that maximizes $Q_2(\boldsymbol{\lambda} \mid \boldsymbol{\psi}^{(k)})$ in relation to $\boldsymbol{\lambda}$, where

$$\begin{aligned} Q_2(\boldsymbol{\lambda} \mid \boldsymbol{\psi}^{(k)}) &= \sum_{i=1}^n \left\{ (\tilde{M}_i^{(k)} - \delta_i) \log \left[2\Phi\left(-\frac{y_i}{\sigma}\right) \right. \right. \\ &\quad \left. \left. + \sqrt{\frac{2^\alpha}{\pi}} \left(\frac{\sigma}{y_i}\right)^\alpha \Gamma\left(\frac{\alpha+1}{2}\right) G\left(\frac{y_i^2}{2\sigma^2}, \frac{\alpha+1}{2}\right) \right] \right. \\ &\quad \left. + \delta_i \left[\log(\alpha) + \frac{\alpha}{2} \log(2) + \alpha \log(\sigma) - (\alpha+1) \log(y_i) \right. \right. \\ &\quad \left. \left. + \log \left[\gamma\left(\frac{y_i^2}{2\sigma^2}, \frac{\alpha+1}{2}\right) \right] \right] \right\}. \end{aligned}$$

The E-, M1- and M2-steps are repeatedly alternated until a suitable convergence rule is satisfied, e.g., the difference in successive values of the estimates is less than a prespecified tolerance. Note that this procedure requires two independent maximization procedures in relation to $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ of dimensions $r + 1$ and 2, respectively, instead of a unique maximization procedure in relation to $\boldsymbol{\psi}$, which has a dimension $(r + 3)$. Standard errors for $\hat{\boldsymbol{\psi}}$ can be

computed based on the Hessian matrix, i.e., $\mathbf{H}(\hat{\boldsymbol{\psi}}) = \partial^2 \ell(\boldsymbol{\psi}) / \partial \boldsymbol{\psi}^\top \partial \boldsymbol{\psi} |_{\boldsymbol{\psi}=\hat{\boldsymbol{\psi}}}$. This matrix can be estimated numerically using, for instance, the package pracma [30] of R [31]. As the estimators obtained via the EM algorithm are MLE, they satisfy the asymptotic distribution

$$\sqrt{n}(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}) \xrightarrow{D} N(\mathbf{0}, [\mathbf{H}(\hat{\boldsymbol{\psi}})]^{-1}), \quad \text{as } n \rightarrow +\infty, \tag{9}$$

where \xrightarrow{D} denotes convergence in distribution. This result allows us to compute an approximated confidence interval for the parameters. Moreover, based on the delta method (see Sen et al. [32] for details), we can also build a confidence interval for any scalar function of $\boldsymbol{\psi}$, say $g(\boldsymbol{\psi})$, using

$$\sqrt{n}[g(\hat{\boldsymbol{\psi}}) - g(\boldsymbol{\psi})] \xrightarrow{D} N(\mathbf{0}, [\nabla g(\boldsymbol{\psi})]^\top [\mathbf{H}(\hat{\boldsymbol{\psi}})]^{-1} \nabla g(\boldsymbol{\psi})), \quad \text{as } n \rightarrow +\infty, \tag{10}$$

where $\nabla g(\boldsymbol{\psi}) = \partial g(\boldsymbol{\psi}) / \partial \boldsymbol{\psi}$ denotes the gradient of $g(\boldsymbol{\psi})$. In particular, a very useful case is $g(\boldsymbol{\psi}) = a_0 / A(\theta_i)$, i.e., the cure rate of the corresponding model. The described estimation procedure was incorporated in the package PScr [33] of R [31].

5. Simulation Studies

In this Section, we present two simulation studies. The first one is related to assessing the performance of the MLE using the EM algorithm discussed in Section 3. The second study is devoted to assessing the performance of the PS-SHN compared with the PS-Weibull model when data are drawn from the PS-Weibull model with some modified scheme. The software tool used for this study was R [31].

5.1. Assessing the Performance of MLE in a Finite Sample Size

In this subsection, we present a simulation study to evaluate the performance of the MLE estimators of the PS-SHN in a finite sample size. We considered three particular distributions belonging to the PS model for the number of concurrent causes: Poisson, logarithmic and geometric. We also considered three sample sizes: 100, 200 and 400. For each individual, we drew two covariates, say x_1 and x_2 , where x_1 was drawn from the Bernoulli distribution with a success probability of 0.5 and x_2 from the standard normal model. For this reason, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ had dimension three. For each particular distribution for the concurrent causes, we considered two scenarios, named high and low cure rates. For the high cure rate, we considered that for β_2 fixed at zero, the values of the two levels for x_1 for the cure rate were 0.7 and 0.45. With such information, the values of β_0 and β_1 could be determined from the equations

$$\exp(-\exp(\beta_0)) = 0.7 \quad \text{and} \quad \exp(-\exp(\beta_0 + \beta_1)) = 0.45$$

for the Poisson case and

$$\frac{1}{A\left(\frac{\exp(\beta_0)}{1+\exp(\beta_0)}\right)} = 0.7 \quad \text{and} \quad \frac{1}{A\left(\frac{\exp(\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}\right)} = 0.45,$$

for the logarithmic and geometric cases. A similar procedure could be applied for the low cure rate case, where we considered that, for β_2 fixed at zero, the values of the two levels for x_1 for the cure rate were 0.5 and 0.25. Summarizing, for the Poisson model, we had $\beta_0 = -1.031$, $\beta_1 = 0.806$ and $\beta_2 = 0$ for the high cure rate case and $\beta_0 = -0.367$, $\beta_1 = 0.693$ and $\beta_2 = 0$ for the lower cure rate case; for the logarithmic distribution we had $\beta_0 = 0.132$, $\beta_1 = 1.588$ and $\beta_2 = 0$ for the high cure rate case and $\beta_0 = 1.366$, $\beta_1 = 2.534$ and $\beta_2 = 0$ for the lower cure rate case; and for the geometric model, we had $\beta_0 = -0.847$, $\beta_1 = 1.048$ and $\beta_2 = 0$ for the high cure rate case and $\beta_0 = 0$, $\beta_1 = 1.099$ and $\beta_2 = 0$ for the lower cure rate case. Finally, we also considered two values for α , 0.75 and 2.5 and two values for σ , 1 and 3. The combinations of distributions for the concurrent causes, sample

size, type of cure rate (high or low), α and σ produced 72 different scenarios. However, we only present the cases for the Poisson and geometric models, although conclusions are similar for the logarithmic distribution. For each, we considered 1000 replicates and computed the MLE based on the EM algorithm described in Section 3. Such results are summarized in Tables 1 and 2 through the average of the estimated bias (bias), the average of the estimated standard error (s.e.) based on the Hessian matrix numerically evaluated at $\psi = \hat{\psi}$, the root of the estimated mean squared error (RMSE) and the 95% coverage probability (CP). As expected, the bias of the estimators was reduced and the terms s.e. and RMSE were closer when the sample size was increased, suggesting the asymptotic consistency of the estimators. Additionally, the terms CP were closer to the nominal values used for their construction when n increased.

Table 1. Simulation study to assess the performance of MLE in the Poisson–SHN model.

Cure Rate	σ	α	Parameter	$n = 100$				$n = 200$				$n = 400$			
				bias	s.e.	RMSE	CP	bias	s.e.	RMSE	CP	bias	s.e.	RMSE	CP
High	1	0.75	$\hat{\beta}_0$	0.023	0.370	0.255	0.978	0.012	0.248	0.200	0.974	0.010	0.177	0.147	0.965
			$\hat{\beta}_1$	−0.019	0.335	0.302	0.961	−0.006	0.234	0.210	0.953	−0.001	0.164	0.145	0.952
			$\hat{\beta}_2$	0.005	0.165	0.144	0.963	0.003	0.116	0.104	0.958	−0.002	0.081	0.074	0.949
			$\hat{\alpha}$	0.294	0.979	0.751	0.975	0.172	0.600	0.505	0.960	0.055	0.376	0.335	0.951
			$\hat{\sigma}$	0.082	0.452	0.372	0.965	0.052	0.321	0.276	0.955	0.015	0.227	0.206	0.952
	2.5	$\hat{\beta}_0$	0.002	0.208	0.180	0.962	0.002	0.145	0.124	0.959	0.001	0.102	0.089	0.956	
		$\hat{\beta}_1$	−0.014	0.316	0.281	0.959	−0.013	0.221	0.192	0.958	−0.008	0.155	0.140	0.953	
		$\hat{\beta}_2$	−0.011	0.155	0.140	0.942	−0.005	0.109	0.103	0.943	−0.003	0.077	0.071	0.951	
		$\hat{\alpha}$	0.287	1.637	1.416	0.929	0.140	1.342	1.091	0.937	0.127	0.995	0.854	0.946	
		$\hat{\sigma}$	−0.039	0.302	0.229	0.933	−0.011	0.211	0.171	0.941	−0.010	0.146	0.131	0.943	
	3	0.75	$\hat{\beta}_0$	−0.140	0.922	0.307	0.962	−0.125	0.738	0.250	0.952	−0.101	0.614	0.212	0.950
			$\hat{\beta}_1$	0.013	0.371	0.327	0.960	0.009	0.258	0.226	0.954	0.003	0.181	0.170	0.953
			$\hat{\beta}_2$	0.006	0.183	0.164	0.948	0.002	0.127	0.111	0.948	0.000	0.089	0.080	0.950
			$\hat{\alpha}$	0.687	1.658	1.304	0.972	0.657	1.389	1.269	0.961	0.584	1.193	1.169	0.955
			$\hat{\sigma}$	−0.269	1.073	0.846	0.957	−0.088	0.892	0.719	0.956	−0.040	0.652	0.602	0.953
2.5	$\hat{\beta}_0$	0.100	0.505	0.263	0.969	0.085	0.339	0.201	0.963	0.078	0.223	0.160	0.958		
	$\hat{\beta}_1$	−0.011	0.325	0.295	0.958	−0.009	0.226	0.201	0.952	−0.008	0.159	0.144	0.951		
	$\hat{\beta}_2$	0.005	0.161	0.151	0.939	0.005	0.112	0.097	0.942	0.003	0.078	0.068	0.949		
	$\hat{\alpha}$	−0.791	1.976	1.583	0.938	−0.549	1.797	1.496	0.941	−0.409	1.490	1.442	0.949		
	$\hat{\sigma}$	−0.641	1.428	0.893	0.919	−0.409	1.237	0.691	0.928	−0.319	0.939	0.592	0.930		
Low	1	0.75	$\hat{\beta}_0$	0.013	0.441	0.296	0.977	0.009	0.285	0.226	0.960	0.005	0.206	0.172	0.954
			$\hat{\beta}_1$	−0.015	0.413	0.369	0.960	−0.013	0.289	0.270	0.952	−0.001	0.202	0.181	0.951
			$\hat{\beta}_2$	0.013	0.198	0.180	0.951	0.008	0.138	0.125	0.950	0.003	0.096	0.088	0.950
			$\hat{\alpha}$	0.293	1.127	0.747	0.972	0.262	0.778	0.632	0.965	0.094	0.459	0.391	0.959
			$\hat{\sigma}$	0.084	0.531	0.385	0.944	0.063	0.392	0.324	0.947	0.037	0.272	0.243	0.949
	2.5	$\hat{\beta}_0$	0.004	0.238	0.210	0.961	0.002	0.163	0.144	0.953	0.001	0.114	0.104	0.951	
		$\hat{\beta}_1$	−0.010	0.384	0.329	0.965	−0.009	0.270	0.239	0.957	−0.004	0.189	0.178	0.953	
		$\hat{\beta}_2$	0.012	0.185	0.166	0.960	0.008	0.129	0.118	0.959	0.004	0.091	0.081	0.951	
		$\hat{\alpha}$	0.175	1.975	1.271	0.924	0.134	1.722	1.144	0.938	0.087	1.198	0.968	0.944	
		$\hat{\sigma}$	−0.057	0.353	0.258	0.912	−0.016	0.244	0.190	0.940	−0.002	0.170	0.146	0.943	
	3	0.75	$\hat{\beta}_0$	−0.151	0.999	0.346	0.941	−0.139	0.805	0.282	0.948	−0.105	0.679	0.227	0.950
			$\hat{\beta}_1$	−0.012	0.464	0.402	0.969	−0.010	0.323	0.279	0.962	−0.007	0.225	0.202	0.958
			$\hat{\beta}_2$	0.012	0.221	0.206	0.943	0.010	0.153	0.140	0.946	0.004	0.107	0.098	0.950
			$\hat{\alpha}$	0.526	1.660	1.336	0.971	0.371	1.378	1.251	0.965	0.294	1.276	1.154	0.959
			$\hat{\sigma}$	−0.467	1.285	0.938	0.929	−0.176	2.077	0.719	0.936	−0.025	1.785	0.619	0.940
2.5	$\hat{\beta}_0$	0.107	0.586	0.295	0.973	0.093	0.405	0.225	0.962	0.086	0.261	0.177	0.959		
	$\hat{\beta}_1$	0.013	0.398	0.365	0.955	0.008	0.276	0.239	0.954	0.000	0.194	0.179	0.951		
	$\hat{\beta}_2$	−0.017	0.192	0.173	0.965	−0.002	0.133	0.119	0.961	0.000	0.093	0.081	0.959		
	$\hat{\alpha}$	−0.912	4.181	1.545	0.934	−0.682	3.910	1.489	0.936	−0.493	3.601	1.462	0.942		
	$\hat{\sigma}$	−0.743	1.490	0.969	0.961	−0.526	1.287	0.764	0.960	−0.379	1.038	0.647	0.958		

Table 2. Simulation study to assess the performance of MLE in the geometric–SHN model.

Cure Rate	σ	α	Estimator	$n = 100$				$n = 200$				$n = 400$			
				bias	s.e.	MSE	CP	bias	s.e.	MSE	CP	bias	s.e.	MSE	CP
High	1	0.75	$\hat{\beta}_0$	0.025	0.556	0.391	0.970	0.008	0.374	0.278	0.983	0.006	0.259	0.219	0.966
			$\hat{\beta}_1$	0.023	0.469	0.424	0.955	0.021	0.328	0.289	0.956	0.001	0.230	0.204	0.955
			$\hat{\beta}_2$	0.004	0.234	0.213	0.954	0.003	0.162	0.150	0.942	0.001	0.114	0.106	0.936
			$\hat{\alpha}$	0.284	1.203	0.741	0.996	0.218	0.796	0.627	0.982	0.109	0.495	0.431	0.961
			$\hat{\sigma}$	0.078	0.546	0.379	0.949	0.070	0.394	0.342	0.950	0.033	0.280	0.251	0.948
	2.5	$\hat{\beta}_0$	−0.014	0.353	0.302	0.959	−0.003	0.241	0.218	0.947	0.003	0.168	0.152	0.949	
		$\hat{\beta}_1$	0.016	0.453	0.403	0.959	0.012	0.315	0.294	0.956	0.008	0.221	0.199	0.941	
		$\hat{\beta}_2$	0.006	0.225	0.200	0.959	0.004	0.156	0.143	0.944	0.002	0.109	0.100	0.943	
		$\hat{\alpha}$	0.242	1.871	1.434	0.854	0.163	1.752	1.270	0.910	0.081	1.332	1.018	0.946	
		$\hat{\sigma}$	−0.078	0.375	0.272	0.911	−0.017	0.266	0.206	0.944	−0.014	0.184	0.153	0.964	
	3	0.75	$\hat{\beta}_0$	−0.153	0.609	0.417	0.989	−0.135	0.488	0.329	0.960	−0.111	0.323	0.262	0.964
			$\hat{\beta}_1$	−0.012	0.510	0.448	0.959	−0.003	0.353	0.310	0.961	−0.001	0.247	0.225	0.946
			$\hat{\beta}_2$	−0.006	0.253	0.226	0.948	−0.004	0.174	0.155	0.959	−0.003	0.121	0.105	0.961
			$\hat{\alpha}$	0.534	1.577	1.389	1.000	0.381	1.402	1.313	1.000	0.194	1.356	1.309	1.000
			$\hat{\sigma}$	−0.490	1.302	0.974	0.910	−0.197	1.153	0.747	0.972	−0.029	0.825	0.638	0.987
2.5	$\hat{\beta}_0$	0.121	0.505	0.385	0.989	0.095	0.380	0.289	0.990	0.083	0.232	0.215	0.985		
	$\hat{\beta}_1$	−0.005	0.460	0.427	0.961	−0.003	0.321	0.279	0.961	−0.002	0.225	0.199	0.957		
	$\hat{\beta}_2$	0.012	0.228	0.196	0.960	0.002	0.159	0.137	0.955	−0.001	0.111	0.099	0.949		
	$\hat{\alpha}$	−0.921	2.029	1.755	0.841	−0.709	1.903	1.570	0.835	−0.569	1.661	1.535	0.819		
	$\hat{\sigma}$	−0.788	1.333	1.004	0.851	−0.555	1.052	0.825	0.886	−0.407	0.899	0.672	0.914		
Low	1	0.75	$\hat{\beta}_0$	0.008	0.477	0.344	0.981	0.003	0.338	0.254	0.969	0.001	0.226	0.194	0.971
			$\hat{\beta}_1$	0.004	0.422	0.384	0.954	0.003	0.294	0.256	0.957	0.003	0.207	0.189	0.951
			$\hat{\beta}_2$	−0.011	0.212	0.197	0.940	−0.006	0.147	0.133	0.953	−0.003	0.103	0.095	0.948
			$\hat{\alpha}$	0.280	1.043	0.702	0.998	0.190	0.693	0.583	0.981	0.097	0.441	0.387	0.957
			$\hat{\sigma}$	0.071	0.480	0.349	0.960	0.059	0.352	0.300	0.960	0.037	0.251	0.217	0.952
	2.5	$\hat{\beta}_0$	0.012	0.309	0.274	0.954	0.010	0.213	0.186	0.956	0.004	0.149	0.136	0.953	
		$\hat{\beta}_1$	0.010	0.409	0.360	0.957	0.007	0.286	0.247	0.964	0.003	0.201	0.180	0.942	
		$\hat{\beta}_2$	−0.007	0.206	0.190	0.961	−0.005	0.143	0.125	0.957	−0.002	0.100	0.084	0.962	
		$\hat{\alpha}$	0.353	1.622	1.301	0.868	0.274	1.438	1.258	0.932	0.269	1.025	0.968	0.939	
		$\hat{\sigma}$	−0.078	0.340	0.260	0.900	−0.040	0.244	0.184	0.956	−0.015	0.167	0.144	0.952	
	3	0.75	$\hat{\beta}_0$	−0.155	0.524	0.379	0.970	−0.128	0.415	0.284	0.973	−0.109	0.266	0.240	0.961
			$\hat{\beta}_1$	0.009	0.448	0.399	0.963	0.005	0.311	0.271	0.962	0.003	0.219	0.187	0.956
			$\hat{\beta}_2$	−0.016	0.223	0.196	0.965	−0.006	0.155	0.136	0.959	−0.003	0.109	0.099	0.950
			$\hat{\alpha}$	0.605	1.893	1.506	1.000	0.481	1.501	1.388	1.000	0.276	1.372	1.292	1.000
			$\hat{\sigma}$	−0.340	1.372	0.894	0.929	−0.089	1.116	0.703	0.966	−0.053	0.773	0.612	0.993
2.5	$\hat{\beta}_0$	0.090	0.607	0.347	0.987	0.074	0.439	0.268	0.990	0.051	0.282	0.201	0.982		
	$\hat{\beta}_1$	0.011	0.415	0.378	0.954	0.005	0.289	0.270	0.946	0.001	0.203	0.180	0.949		
	$\hat{\beta}_2$	−0.005	0.208	0.187	0.952	−0.001	0.145	0.135	0.937	−0.001	0.101	0.094	0.944		
	$\hat{\alpha}$	−0.865	1.780	1.612	0.852	−0.630	1.708	1.517	0.837	−0.449	1.629	1.462	0.838		
	$\hat{\sigma}$	−0.695	1.283	0.927	0.887	−0.492	1.097	0.762	0.905	−0.341	0.732	0.605	0.924		

5.2. Misspecification for the Time-to-Event for the Concurrent Causes

In this subsection, we assessed the performance of the MLE for the Poisson–SHN and geometric–SHN models when data were drawn from an artificial model with heavy tails based on the Weibull distribution.

The generation of the data was very similar to that in the last subsection, considering the following changes:

- The times related to the concurrent causes were drawn from the Weibull model with mean $\mathbb{E}(W) = \mu$ and variance 1. We considered three values for μ : 2.5, 5.0 and 7.5. For the parameterization considered in part a) from Remark 1, this corresponded

to $\alpha = 2.70$ and $\sigma = 7.90$, $\alpha = 5.80$ and $\sigma = 29.16$ and $\alpha = 8.97$ and $\sigma = 62.75$, respectively.

- A determined number of failure times (2% of the sample size, i.e., two observations for $n = 100$, four observations for $n = 200$ and eight observations for $n = 400$), were imputed as $\max(t^{fail}) + U\sigma_t^{fail}$, where $\max(t^{fail})$ and σ_t^{fail} denote the maximum and the standard error of the failure times and $U \sim U(1, 2)$. This scheme artificially simulated a distribution for the time-to-event of the concurrent causes with a heavier tail than the Weibull distribution, but not corresponding to the SHN.

The main goal of this study was to compare the coefficients related to the cure (i.e., the components of the vector β) obtained with the Poisson–SHN and Poisson–Weibull models (when the true model for M is the Poisson distribution) and geometric–SHN and geometric–Weibull models (when the true model is the geometric distribution). All the considered combinations of $\mathbb{E}(W)$, sample size, type of cure (high/low) and model for M (Poisson/geometric), totalizing 36 cases. Each was replicated 1000 times. Table 3 summarizes the bias and RMSE for each combination. Note that, in all the considered cases, the bias and RMSE were lower (or at least the same) for the models considering the SHN rather than the Weibull model. Such differences in favor of the SHN were minimized when the sample size was increased.

Table 3. Misspecification simulation study for Poisson–SHN and geometric–SHN models.

		$n = 100$				$n = 200$				$n = 400$				
		Poisson Case												
Cure Rate	$\mathbb{E}(W)$	Estimator	bias		RMSE		bias		RMSE		bias		RMSE	
			SHN	Weibull	SHN	Weibull	SHN	Weibull	SHN	Weibull	SHN	Weibull	SHN	Weibull
High	2.5	$\hat{\beta}_0$	0.059	0.089	0.394	0.431	0.053	0.084	0.248	0.309	0.051	0.073	0.168	0.221
		$\hat{\beta}_1$	−0.018	−0.066	0.356	0.430	−0.006	−0.049	0.248	0.316	−0.003	−0.048	0.175	0.227
		$\hat{\beta}_2$	−0.002	−0.003	0.176	0.255	−0.002	−0.003	0.122	0.187	−0.001	−0.002	0.086	0.131
	5.0	$\hat{\beta}_0$	−0.109	−0.159	0.344	0.347	−0.077	−0.126	0.225	0.239	−0.058	−0.123	0.134	0.169
		$\hat{\beta}_1$	−0.014	−0.020	0.402	0.773	−0.007	−0.018	0.280	0.601	−0.005	−0.017	0.197	0.451
		$\hat{\beta}_2$	−0.005	−0.006	0.197	0.288	−0.003	−0.004	0.138	0.209	0.000	0.000	0.097	0.146
	7.5	$\hat{\beta}_0$	−0.307	−0.359	0.246	0.297	−0.276	−0.355	0.201	0.210	−0.251	−0.322	0.107	0.159
		$\hat{\beta}_1$	−0.017	−0.025	0.443	0.933	−0.012	−0.015	0.311	0.773	−0.011	−0.013	0.219	0.650
		$\hat{\beta}_2$	0.011	0.012	0.218	0.317	−0.003	−0.003	0.152	0.230	0.001	0.001	0.107	0.160
Low	2.5	$\hat{\beta}_0$	0.087	0.097	0.287	0.314	0.070	0.079	0.181	0.211	0.058	0.063	0.139	0.144
		$\hat{\beta}_1$	−0.014	−0.069	0.437	0.448	−0.012	−0.061	0.306	0.334	−0.007	−0.052	0.215	0.228
		$\hat{\beta}_2$	−0.001	0.000	0.207	0.297	0.000	0.000	0.146	0.215	−0.001	−0.001	0.102	0.155
	5.0	$\hat{\beta}_0$	−0.093	−0.100	0.234	0.255	−0.071	−0.092	0.165	0.176	−0.052	−0.078	0.110	0.121
		$\hat{\beta}_1$	−0.029	−0.048	0.489	0.809	−0.022	−0.025	0.344	0.656	−0.008	−0.012	0.244	0.487
		$\hat{\beta}_2$	−0.003	−0.004	0.231	0.336	−0.001	−0.001	0.164	0.241	−0.001	−0.001	0.116	0.172
	7.5	$\hat{\beta}_0$	−0.204	−0.285	0.197	0.217	−0.194	−0.282	0.133	0.155	−0.192	−0.279	0.102	0.111
		$\hat{\beta}_1$	−0.012	−0.018	0.533	0.893	−0.003	−0.014	0.379	0.742	−0.002	−0.004	0.267	0.583
		$\hat{\beta}_2$	−0.008	−0.008	0.252	0.372	−0.005	−0.006	0.179	0.265	−0.003	−0.003	0.127	0.187

Table 3. Cont.

		<i>n</i> = 100				<i>n</i> = 200				<i>n</i> = 400				
Geometric Case														
High	2.5	$\hat{\beta}_0$	0.071	0.092	0.196	0.244	0.053	0.082	0.163	0.190	0.031	0.048	0.124	0.139
		$\hat{\beta}_1$	−0.105	−0.127	0.209	0.281	−0.083	−0.094	0.153	0.218	−0.054	−0.056	0.148	0.164
		$\hat{\beta}_2$	−0.003	−0.005	0.254	0.435	0.001	0.002	0.176	0.299	0.001	0.001	0.123	0.207
	5.0	$\hat{\beta}_0$	−0.091	−0.129	0.135	0.154	−0.088	−0.119	0.098	0.103	−0.071	−0.106	0.042	0.071
		$\hat{\beta}_1$	−0.089	−0.130	0.352	0.427	−0.080	−0.120	0.287	0.329	−0.057	−0.092	0.172	0.243
		$\hat{\beta}_2$	0.001	0.000	0.274	0.478	−0.001	−0.001	0.192	0.330	0.001	0.001	0.135	0.229
	7.5	$\hat{\beta}_0$	−0.288	−0.333	0.094	0.125	−0.251	−0.311	0.067	0.082	−0.242	−0.271	0.049	0.059
		$\hat{\beta}_1$	−0.055	−0.099	0.493	0.512	−0.054	−0.096	0.314	0.389	−0.042	−0.053	0.290	0.301
		$\hat{\beta}_2$	0.009	−0.010	0.293	0.524	0.005	−0.008	0.204	0.362	0.001	0.001	0.143	0.248
Low	2.5	$\hat{\beta}_0$	−0.075	−0.109	0.245	0.298	−0.062	−0.081	0.219	0.234	−0.046	−0.069	0.147	0.165
		$\hat{\beta}_1$	−0.094	−0.130	0.217	0.245	−0.074	−0.083	0.133	0.192	−0.050	−0.061	0.128	0.138
		$\hat{\beta}_2$	0.012	0.016	0.235	0.397	0.011	0.012	0.163	0.270	0.002	0.003	0.114	0.187
	5.0	$\hat{\beta}_0$	−0.224	−0.259	0.197	0.209	−0.216	−0.237	0.101	0.144	−0.153	−0.203	0.061	0.089
		$\hat{\beta}_1$	−0.108	−0.126	0.303	0.385	−0.097	−0.106	0.251	0.283	−0.084	−0.099	0.146	0.207
		$\hat{\beta}_2$	0.002	0.002	0.253	0.426	0.002	0.002	0.176	0.292	0.001	0.001	0.123	0.201
	7.5	$\hat{\beta}_0$	−0.260	−0.459	0.116	0.150	−0.256	−0.424	0.096	0.106	−0.202	−0.392	0.042	0.073
		$\hat{\beta}_1$	−0.050	−0.112	0.337	0.471	−0.049	−0.104	0.274	0.359	−0.035	−0.091	0.202	0.268
		$\hat{\beta}_2$	−0.002	−0.004	0.269	0.456	−0.001	−0.002	0.187	0.315	−0.001	−0.001	0.131	0.216

6. Application to Bone Marrow Transplantation

In this section, we present a real data set application from the EBMT presented in the introduction section. Additionally, the prophylaxis covariate (dichotomous, 1730 patients with “no” and 549 patients with “yes”) was available.

For the analysis, we considered some existent models in the literature, all of them belonging to the power-series cure rate model, such as the Poisson, logarithmic, negative binomial, Bernoulli, COM-Poisson and polylogarithm models. Additionally, we considered three commons models used for the concurrent causes such as the Weibull, gamma and BS models. Table 4 shows the Akaike information criterion (AIC) [34] and Bayesian information criterion (BIC) criteria [35] for those models in the ebmt4 data. Note that both criteria favored models where the SHN for concurrent causes was used.

Table 4. AIC and BIC criteria for some cure rate models in the literature based on a competing risk scheme for the ebmt4 data set.

Model for <i>M</i>	Weibull	Gamma	BS	SHN
Po	4650.0/4672.9	4683.0/4706.0	4580.8/4603.8	4537.9/4560.8
Lo	4628.7/4651.6	4639.6/4662.6	4628.1/4651.0	4547.3/4570.3
NB	4561.4/4590.1	4547.4/4576.1	4583.0/4611.7	4540.2/4568.8
Bern	4665.1/4688.0	4712.8/4735.8	4577.0/4599.9	4535.8/4558.8
COM-Po	4639.3/4667.9	4659.5/4688.1	4578.9/4607.6	4537.8/4566.5
PL	4612.1/4658.0	4615.9/4661.8	4585.1/4630.9	4545.3/4591.2

Table 5 shows the estimated parameters for the Bern/SHN cure rate model and its standard errors. Note that both parameters related to the cure were significant with 5% of significance. We also considered a residuals analysis in the selected model based on the Cox–Snell residuals [36,37] and the randomized quantile residuals [38]. The first residual was compared with its Nelson–Aalen estimator and the second with the quantile of the standard normal distribution. In both cases, if the model was correct, we expected to observe the identity function for the respective graphs. According to Figure 2, such residuals suggested

that the choice model was correct. Finally, we also computed the Kaplan–Meier (K–M) estimator for patients with and without prophylaxis and the respective point estimation and 95% confidence interval for the survival function for each time. Such confidence interval was based on the delta method discussed in Equation (10). Figure 3 shows those results. Note the closeness of the estimation for the survival function and that provided by the K–M estimator. As expected, patients without prophylaxis had a lower survival function than patients with prophylaxis.

Table 5. Estimates and standard errors (s.e.) for the Bern/SHN cure rate model in the ebmt4 data set.

Parameter	β_0	β_1	α	σ
Estimate	−0.4912	−0.2224	0.3844	0.3120
s.e.	0.0907	0.1030	0.0753	0.0310

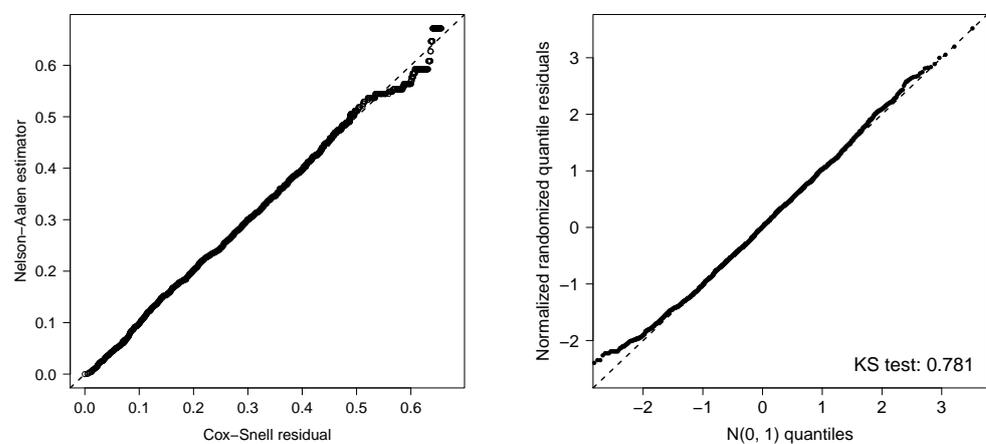


Figure 2. Cox–Snell residuals versus their Nelson–Aalen estimator (left panel) and the randomized quantile residuals versus the N(0,1) quantiles and its p-values for the Kolmogorov–Smirnov test (right panel).

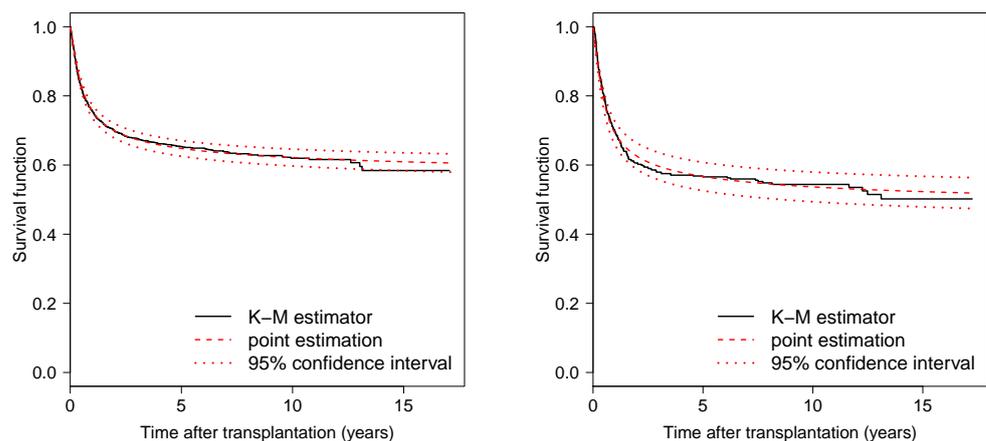


Figure 3. Kaplan–Meier (K–M) estimator for patients with and without prophylaxis and the respective point estimation and 95% confidence interval for the survival function for each time.

7. Conclusions

In this work, we introduced the PS–SHN cure rate model. We provided a mathematical treatment of the new distribution including an alternative expression for the cumulative distribution function. An EM algorithm to compute maximum likelihood estimation procedures in finite samples was discussed. Two simulation studies were performed to assess some properties of the estimators, showing the good performance of the proposed

estimation procedure in finite samples. In the analysis of the bone marrow transplant data, the SHN cure rate model yielded a better fit when compared to other well-known models in the literature, according to the AIC and BIC criteria. We developed an R package to perform inference in the power-series cure rate models. As part of future research, we plan to explore other estimation methods for the proposed cure rate model, for instance, a Bayesian approach. A second line of a future research will be a bias-reduction methodology to estimate the parameters of the proposed regression model.

Author Contributions: Conceptualization, D.I.G. and Y.M.G.; formal analysis, Y.M.G. and H.J.G.; investigation, Y.M.G. and M.J.G.-N.; software, D.I.G. and M.B.; validation H.J.G., M.J.G.-N. and M.B.; writing—original draft preparation H.J.G.; writing—review and editing D.I.G., Y.M.G., M.J.G.-N. and M.B.; Supervision, D.I.G. and Y.M.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: This research received no external funding.

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

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

Appendix A. Parametrization Used for Some Discrete Models

The power series distribution has a probability mass function (p.m.f.) given by

$$P(M = m; \theta) = \frac{a_m \theta^m}{A(\theta)}, \quad m = 0, 1, 2, \dots,$$

where $A(\theta) = \sum_{m=0}^{\infty} a_m \theta^m$ and $\theta \in \Theta$, with Θ the parameter space for θ .

Table A1. Some particular cases of the power series distribution.

	Poisson	Negative Binomial	Binomial	Logarithmic	Polylogarithm	COM-Poisson
a_m	$(m!)^{-1}$	$\binom{m+q-1}{m}$	$\binom{q}{m}$	$(m+1)^{-1}$	$(m+1)^q$	$(m!)^{-q}$
$A(\theta)$	e^θ	$(1-\theta)^{-q}$	$(1+\theta)^q$	$-\frac{\log(1-\theta)}{\theta}$	$\frac{Li_q(\theta)}{\theta}$	$\frac{1}{Z(\theta,q)}$
Θ	$(0, \infty)$	$(0, 1)$	$(0, \infty)$	$(0, 1)$	$(0, 1)$	$(0, \infty)$
Notation	Po(θ)	NB(q, θ)	Bin(q, θ)	Lo(θ)	PL(q, θ)	COM-Po(q, θ)

NOTE: Li denotes the polylogarithm defined as $Li_q(\theta) = \sum_{m=1}^{\infty} m^{-q} \theta^m$ and $Z(\theta, q) = \sum_{m=0}^{\infty} (m!)^{-q} \theta^m$. In the NB, Bin, PL and COM-Po distributions, q is considered known.

Observations:

- NB(1, θ) = Geo(θ), i.e., the geometric distribution.
- Bin(1, θ) = Bern(θ), i.e., the Bernoulli distribution.
- NB(q, θ) = Bin(q, θ), if $-q^{-1} \in \mathbb{N}$ and $0 < \theta/q < 1$.
- $\lim_{q \rightarrow \infty} \text{COM-Po}(q, \theta) = \text{Bern}((1 + \theta)^{-1})$.
- $\lim_{q \rightarrow 0} \text{COM-Po}(q, \theta) = \text{Geo}(1 - \theta)$, if $\theta < 1$.

References

1. Olmos, N.M.; Varela, H.; Gómez, H.W.; Bolfarine, H. An extension of the half-normal distribution. *Stat. Pap.* **2012**, *53*, 875–886. [[CrossRef](#)]
2. Gómez, Y.M.; Gallardo, D.I.; De Castro, M. A regression model for positive data based on the slash half-normal distribution. *Revstat* **2021**, *19*, 553–573.
3. Hardy, R.E.; Ikpeazu, E.V. Bone marrow transplantation: A review. *J. Natl. Med. Assoc.* **1989**, *81*, 518–523.
4. Chinen, J.; Buckley, R.H. Transplantation immunology: Solid organ and bone marrow. *J. Allergy Clin. Immunol.* **2010**, *125*, S324–S335. [[CrossRef](#)]

5. Simpson, E.; Dazzi, F. Bone Marrow Transplantation 1957–2019. *Front. Immunol.* **2019**, *10*, 1246. [CrossRef]
6. The European Society for Blood and Marrow Transplantation (s.f.). The EBMT. Available online: <https://www.ebmt.org> (accessed on 10 December 2022).
7. de Rooij, J.D.; Zwaan, C.M.; van den Heuvel-Eibrink, M. Pediatric AML: From Biology to Clinical Management. *J. Clin. Med.* **2015**, *4*, 127–149. [CrossRef]
8. Sander, A.; Zimmermann, M.; Dworzak, M.; Fleischhack, G.; von Neuhoff, C.; Reinhardt, D.; Kaspers, G.J.; Creutzig, U. Consequent and Intensified Relapse Therapy Improved Survival in Pediatric AML: Results of Relapse Treatment In 379 Patients of Three Consecutive Aml-Bfm Trials. *Leukemia* **2010**, *24*, 1422–1428. [CrossRef]
9. Fiocco, M.; Putter, H.; van Houwelingen, H.C. Reduced-rank proportional hazards regression and simulation-based prediction for multi-state models. *Stat. Med.* **2008**, *27*, 4340–4358. [CrossRef]
10. Putter, H.; de Wreede, L.; Fiocco, M. Mstate: Data Preparation, Estimation and Prediction in Multi-State Models. R package version 2016.0-2. 2016. Available online: <https://CRAN.R-project.org/package=mstate> (accessed on 10 December 2022).
11. Horio, T.; Morishita, E.; Mizuno, S.; Uchino, K.; Hanamura, I.; Espinoza, J.L.; Morishima, Y.; Kodaera, Y.; Onizuka, M.; Kashiwase, K.; et al. Donor Heme Oxygenase-1 Promoter Gene Polymorphism Predicts Survival after Unrelated Bone Marrow Transplantation for High-Risk Patients. *Cancers* **2020**, *12*, 424. [CrossRef]
12. Radujkovic, A.; Kordelas, L.; Bogdanov, R.; Müller-Tidow, C.; Beelen, D.W.; Dreger, P.; Luft, T. Interleukin-18 and Hematopoietic Recovery after Allogeneic Stem Cell Transplantation. *Cancers* **2020**, *12*, 2789. [CrossRef]
13. Tsai, X.C.; Chen, T.T.; Gau, J.P.; Wang, P.N.; Liu, Y.C.; Lien, M.Y.; Li, C.C.; Yao, M.; Ko, B.S. Outcomes of Different Haploidentical Transplantation Strategies from the Taiwan Blood and Marrow Transplantation Registry. *Cancers* **2022**, *14*, 1097. [CrossRef]
14. Sabino Pinho de Oliveira, B.; Putti, S.; Naro, F.; Pellegrini, M. Bone Marrow Transplantation as Therapy for Ataxia-Telangiectasia: A Systematic Review. *Cancers* **2020**, *12*, 3207. [CrossRef]
15. 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]
16. Rodrigues, J.; Cancho, V.G.; de Castro, M.A.; Louzada-Neto, F. On the unification of the long-term survival models. *Stat. Probab. Lett.* **2009**, *79*, 753–759. [CrossRef]
17. Pal, S.; Balakrishnan, N. Likelihood inference for COM-Poisson cure rate model with interval-censored data and Weibull lifetimes. *Stat. Methods Med. Res.* **2017**, *26*, 2093–2113. [CrossRef]
18. Balakrishnan, N.; Pal, S. Likelihood Inference for Flexible Cure Rate Models with Gamma Lifetimes. *Commun. Stat. Theory Methods* **2015**, *19*, 4007–4048. [CrossRef]
19. Wiangnak, P.; Pal, S. Gamma Lifetimes and Associated Inference for Interval Censored Cure Rate Model with COM-Poisson Competing Cause. *Commun. Stat. Theory Methods* **2018**, *47*, 1491–1509. [CrossRef]
20. Ortega, E.M.M.; Cordeiro, G.M.; Hashimoto, E.M.; Suzuki, A.K. Regression models generated by gamma random variables with long-term survivors. *Commun. Stat. Appl. Methods* **2017**, *24*, 43–65. [CrossRef]
21. Balakrishnan, N.; Pal, S. Lognormal lifetimes and likelihood-based inference for flexible cure rate models based on COM-Poisson family. *Comput. Stat. Data Anal.* **2013**, *67*, 41–67. [CrossRef]
22. Gallardo, D.I.; Bolfarine, H.; Pedroso-de-Lima, A.C. An EM algorithm for estimating the destructive weighted Poisson cure rate model. *J. Stat. Comput. Simul.* **2016**, *86*, 1497–1515. [CrossRef]
23. Leao, J.; Bourguignon, M.; Saulo, H.; Santos-Neto, M.; Calsavara, V. The Negative Binomial Beta Prime Regression Model with Cure Rate: Application with a Melanoma Dataset. *J. Stat. Theory Pract.* **2021**, *15*, 63. [CrossRef]
24. Birnbaum, Z.W.; Saunders, S.C. A new family of life distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [CrossRef]
25. Birnbaum, Z.W.; Saunders, S.C. Estimation for a family of life distributions with applications to fatigue. *J. Appl. Probab.* **1969**, *6*, 328–347. [CrossRef]
26. Desmond, A.F. On the relationship between two fatigue models. *IEEE Trans. Reliab.* **1986**, *35*, 167–169. [CrossRef]
27. Berkson, J.; Gage, R. Survival curve for cancer patients following treatment. *J. Am. Stat. Assoc.* **1952**, *47*, 501–515. [CrossRef]
28. 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]
29. 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]
30. Borchers, H.W. Pracma: Practical Numerical Math Functions. R Package Version 2.4.2. 2022. Available online: <https://CRAN.R-project.org/package=pracma> (accessed on 10 December 2022).
31. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022. Available online: <https://www.R-project.org/> (accessed on 10 December 2022).
32. Sen, P.K.; Singer, J.M.; Pedroso-de-Lima, A.C. *From Finite Sample to Asymptotic Methods in Statistics*; Cambridge University Press: New York, NY, USA, 2010.
33. Gallardo, D. PScr: Estimation for the Power Series Cure Rate Model. R Package Version 1.0. 2022. Available online: <https://CRAN.R-project.org/package=PScr> (accessed on 10 December 2022).
34. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
35. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]

36. Conlon, A.S.; Taylor, J.M.; Sargent, D.J. Multi-state models for colon cancer recurrence and death with a cured fraction. *Stat. Med.* **2014**, *33*, 1750–1766. [[CrossRef](#)]
37. Cox, D.R. A general definition of residuals. *J. R. Stat. Soc. Ser. B* **1968**, *30*, 248–275. [[CrossRef](#)]
38. Dunn, P.K.; Smyth, G.K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.