Journal of Statistics *ISSN-P: 1684-8403*
Volume 28, 2024, pp. 1-17 *ISSN-E: 2957-8981*

A Comparison of Cure Fraction Estimation Methods in Promotion Time Cure Model based on Burr Type XII Distribution

*Ayesha Tahira[1] and Muhammad Yameen Danish[*2]*

Abstract

Survival data from clinical trials often show the proportion of patients with long-term survivors and the standard models like Cox and accelerated failure time model are inappropriate for fitting such data. The two-component mixture cure model is often used for this purpose. However, for the failure time data with proportional hazards structure, the promotion time cure model can be more adaptable than the mixture cure model. The present paper compares semiparametric method with parametric Bayesian and maximum likelihood methods for estimating the proportion of insusceptible patients using log link function and for modeling the failure times of susceptible subjects using Burr-XII distribution. For Bayesian estimation, we use improper uniform prior distributions for regression parameters and vague gamma prior distributions for baseline distribution parameters. Numerical experiments are considered to examine the performance of different methods. It is observed that for small sample sizes, the Bayes method perform better than the parametric and semiparametric methods in terms of biases, mean square errors and empirical variances and for large sample sizes, the performance of the Bayes and maximum likelihood methods is approximately equal. The proposed methods are applied to real data for illustration and motivation.

Keywords

Bayesian method, Cure fraction, Promotion time cure model, Semiparametric maximum likelihood method.

1. Introduction

In survival analysis, it is assumed that if the complete follow-up of the individuals was possible, each would have eventually experienced the event under study. But, in many clinical trials, it is observed that some patients never experience the event of interest for sufficiently long follow-up time. The mixture cure rate model introduced by Boag (1949) and further studied by Berkson and Gage (1952) is usually applied to fit the survival data with cure fraction in the population assumed to be a mixture of insusceptible and susceptible individuals. Chen et al. (1999) proposed an alternative cure rate model that is quite different from the mixture cure rate model and has several novel properties including proportionality of hazards. Chen and Ibrahim (2001) offered some maximum likelihood

[*] Corresponding author
[1]Department of Statistics, Government Viqar-un-Nisa Post Graduate College for women, Tipu Road, Rawalpindi Pakistan.
Email: tahiraaisha1@gmail.com
[2] Department of Statistics, Allama Iqbal Open University Islamabad, Pin 44000, Pakistan.
Email: yameen.danish@aiou.edu.pk

methods for parameter estimation for a novel class of semi-parametric survival models with a cure fraction allowing the covariates to be missing. Tsodikov et al. (2003) addressed the utility of bounded cumulative hazard model in cure rate estimation. Zeng et al. (2006) proposed a general class of transformation models for survival data with a cure fraction. Lambert et al. (2007) extended the parametric non-mixture cure fraction model to incorporate the background mortality. Portier et al. (2017) considered a semiparametric promotion time cure model. Chen and Du (2018) suggested a nonparametric approach to model the covariate effects under the framework of promotion time cure model. D'Andrea et al. (2018) proposed a unified approach with negative binomial distribution for modeling cure rates under Kumaraswamy family of distributions. Wang and Han (2020) considered the regression analysis of the data with a cured subgroup in population using a semiparametric non-mixture cure model. The readers are referred to Balakrishnan and Pal (2015), Gallardo et al. (2016), Coelho-Barros et al. (2017), Abbas et al. (2020), Liu et al. (2020), Leao et al. (2021), Pedrosa-Laza et al. (2022), Botta et al. (2023), Escobar-Bach and Van Keilegom (2023), Ezquerro et al. (2023) for more recent work and extensions in cure models, corresponding baseline distributions and estimation methods.

In this paper we compare three different methos, namely, parametric maximum likelihood method, parametric Bayesian method and semiparametric maximum likelihood method for estimating the proportion of insusceptible patients using log link function and for modeling the failure times of susceptible subjects using the Burr-XII distribution. Although the Burr distribution is relatively less popular lifetime distribution than Weibull and gamma distributions, however, its probability function can assume different shapes corresponding to different combination of its parameter values. The motivation behind this distribution is that much of the region covered by the gamma and lognormal distributions in the skewness-kurtosis plane is also covered by the Burr type XII distribution; see Rodriguez (1977) for detail.

The remainder of the paper is structured as follows. The model is formulated in Section 2 and its maximum likelihood (ML) estimates are presented in Section 3. Section 4 describes the semiparametric ML estimation. The Bayesian estimation of parameters is considered in Section 5. Numerical experiments for the suitability of different techniques are conducted in Section 6 and an application of proposed methods to a real-life example is discussed in Section 7. Finally, we conclude the paper in Section 8.

2. The model and its assumptions

Suppose that for a subject in a group of patients under investigation, the count of carcinogenic cells left active after some medication is a random variable, say N, distributed as Poisson with parameter θ. Further assume that for the ith cell, Z_i represent the random time taken to create a cancer disease, for $i = 1, 2, 3, ...$, according to some probability distribution $F_Z(.)$ independent of the random variable N. Then $T = \min\{Z_i,\ 0 \leq i \leq N\}$ with $P(Z_0 = \infty) = 1$ can be considered as the time to recurrence of cancer for the subject, where Z_0 denotes the time when there is no cancer. The promotion time cure model for the whole population is defined in Chen et al. (1999) as

$$S_p(t) = \exp(-\theta F_Z(t)). \tag{1}$$

The density and hazard functions corresponding to the model in Equation (1) are, respectively, given by

$f_p(t) = \theta f_Z(t)\exp(-\theta F_Z(t))$

and

$h_p(t) = \theta f_Z(t)$.

Although the model provided in Equation (1) is derived for promotion times of carcinogenic cells, it is often used for modeling the failure time data with cure fraction as an alternative to two-component mixture cure model. The Burr-XII distribution is considered to model the failure time and log link function to model the cure fraction. The density function of the Burr-XII distribution with shape parameters α and λ is

$$f_Z(z;\ \alpha, \lambda) = \alpha\lambda z^{\lambda-1}\left(1 + z^{\lambda}\right)^{-\alpha-1};\ \ z > 0, \alpha > 0, \lambda > 0, \tag{2}$$

and the corresponding distribution function is

$$F_z(z;\ \alpha, \lambda) = 1 - \left(1 + z^{\lambda}\right)^{-\alpha}. \tag{3}$$

The log link function describing the effects of covariates on the cure rate parameter θ in Equation (1) is given by $\theta(X) = \exp(X'\beta)$, where X is a $p \times 1$ vector of covariates, β is a $p \times 1$ vector of regression coefficients.

3. Parametric maximum likelihood estimation

We assume two independent sequences of unobservable random variables, one denoting the failure times $(T_1, \ldots,\ T_n)$ and the other representing the censoring times $(C_1,\ \ldots,\ C_n)$ of n participants enrolled in some life testing investigation. We suppose that $f_T(t)$, $f_C(c)$, $F_T(t)$ and $F_C(c)$ are their corresponding density and distribution functions. Instead of observing T_i's and C_i's, we actually view identically distributed random pairs $(Y_1, D_1), (Y_2, D_2), \ldots, (Y_n, D_n)$,where $Y_i = \min(T_i, C_i)$ and $D_i = 1$ if $T_i \le C_i$ and 0 otherwise; $i = 1, 2, \ldots, n$. If $O = \{y_i, d_i, X_i,\ i = 1, 2, \ldots, n\}$ denote the observed data, then the likelihood function is

$$\begin{aligned} l(\gamma, \beta|O) &= \prod_{i=1}^{n}\left[f_p(y_i)\right]^{d_i}\left[S_p(y_i)\right]^{1-d_i} \\ &= \prod_{i=1}^{n}[\exp(X_i'\beta)f_Z(y_i)]^{d_i}\exp(-\exp(X_i'\beta)F_Z(y_i)), \end{aligned} \tag{4}$$

Where, $\gamma = (\alpha, \lambda)$. In this paper we consider only one covariate so that $X = (X_0, X_1)'$ and $\beta = (\beta_0, \beta_1)'$, extension to more than one covariate is straight forward. Replacing $f_Z(.)$ and $F_Z(.)$ from (2) and (3) in the above likelihood function, we have

$$\begin{aligned} l(\gamma, \beta|O) = \prod_{i=1}^{n}\Big[&\exp(X'\beta)\alpha\lambda y_i^{\lambda-1}\left(1 + y_i^{\lambda}\right)^{-\alpha-1}\Big]^{d_i} \\ &\times \exp\Big[-\exp(X'\beta)\Big\{1 - \left(1 + y_i^{\lambda}\right)^{-\alpha}\Big\}\Big]. \end{aligned} \tag{5}$$

The corresponding log-likelihood function, say $L(\gamma, \beta|O) = \ln[l(\gamma, \beta|O)]$, is

$$L(\gamma,\beta|O) = r\beta_0 + \beta_1\sum_{i=1}^{n} d_i x_i + r\ln(\alpha) + r\ln(\lambda) + \lambda\sum_{i=1}^{n} d_i \ln(y_i) - (\alpha+1)\sum_{i=1}^{n} d_i \ln\left(1+y_i^{\lambda}\right) - \sum_{i=1}^{n}\exp(\beta_0+\beta_1 x_i) + \sum_{i=1}^{n}\frac{\exp(\beta_0+\beta_1 x_i)}{\left(1+y_i^{\lambda}\right)^{\alpha}}, \quad (6)$$

where, $r = \sum_{i=1}^{n} d_i$.

The ML estimates are solution to the following likelihood equations.

$$\frac{\partial L}{\partial \alpha} = \frac{r}{\alpha} - \sum_{i=1}^{n} d_i \ln\left(1+y_i^{\lambda}\right) - \sum_{i=1}^{n}\frac{\exp(\beta_0+\beta_1 x_i)\ln\left(1+y_i^{\lambda}\right)}{\left(1+y_i^{\lambda}\right)^{\alpha}} = 0,$$

$$\frac{\partial L}{\partial \lambda} = \frac{r}{\lambda} + \sum_{i=1}^{n} d_i \ln y_i - (\alpha+1)\sum_{i=1}^{n}\frac{d_i y_i^{\lambda}\ln y_i}{1+y_i^{\lambda}} - \alpha\sum_{i=1}^{n}\frac{\exp(\beta_0+\beta_1 x_i) y_i^{\lambda}\ln y_i}{\left(1+y_i^{\lambda}\right)^{\alpha+1}} = 0,$$

$$\frac{\partial L}{\partial \beta_1} = \sum_{i=1}^{n} d_i x_i - \sum_{i=1}^{n} x_i \exp\left(\beta_0+\beta_1 x_i\right) + \sum_{i=1}^{n} x_i \frac{\exp(\beta_0+\beta_1 x_i)}{\left(1+y_i^{\lambda}\right)^{\alpha}} = 0,$$

$$\frac{\partial L}{\partial \beta_0} = r - \sum_{i=1}^{n}\exp(\beta_0+\beta_1 x_i) + \sum_{i=1}^{n}\frac{\exp(\beta_0+\beta_1 x_i)}{\left(1+y_i^{\lambda}\right)^{\alpha}} = 0.$$

Unfortunately, the solution of these nonlinear equations does not exist in closed form, we need some iterative procedure like Newton-Raphson to solve them simultaneously. We use *maxLik* package in R-language to find the ML estimates and associated metrics. Relevant R-code is provided in Appendix. The detail of the package can be seen in Henningsen and Toomet (2011).

4. Semiparametric maximum likelihood estimation

For semiparametric estimation, the distribution function $F_Z(.)$ of the random variable Z is left unspecified and is estimated nonparametrically. The likelihood function in this regard can be written from (4) as

$$l(\beta,F|O) = \prod_{i=1}^{n}\left\{\left[\left(\exp\left(-\theta(X_i)F(y_i)\right)\theta(X_i)F\{y_i\}\right)^{d_i}\left(\exp\left(-\theta(X_i)F(y_i)\right)\right)^{1-d_i}\right]^{I(y_i<\infty)}\left[\exp\left(-\theta(X_i)\right)\right]^{I(y_i=\infty)}\right\}, \quad (7)$$

where, $F\{y_i\}$ denotes the jump size of $F(.)$ at y_i and $F(.)$ is a right continuous distribution function with jumps at event times only. The corresponding log-likelihood function is

$$L(\beta,F|O) = \sum_{i=1}^{n} I(y_i<\infty)[d_i\log(p_i) + d_i\beta' X_i - F(y_i)\exp(\beta' X_i)] - \sum_{i=1}^{n} I(y_i=\infty)\exp(\beta' X_i), \quad (8)$$

where, $p_i = F\{y_i\}$. Let $y_{(1)}, \ldots, y_{(m)}$ denote the event times ordered from smallest to largest and $p_{(1)}, \ldots, p_{(m)}$ the corresponding jump sizes, where m is the number of distinct event times such that $\sum_i^m p_{(i)} = 1$. The constrained log-likelihood function thus is

$$L(\beta, F|O) = \sum_{i=1}^{m} \log p_{(i)} + \sum_{j=1}^{n} d_j \beta' X_j - \sum_{i=1}^{n} I(y_i < \infty) F_i \exp(\beta' X_i) - \sum_{i=1}^{n} I(y_i = \infty) \exp(\beta' X_i) - \sum_{i=1}^{n} I(y_i = \infty)\, e\exp(\beta' X_i) - \tau\left(\sum_{i}^{m} p_{(i)} - 1\right), \quad (9)$$

where, τ is Lagrange multiplier, and

$$F_i = F(y_i) = F_Z(y_i) = \sum\nolimits_{y_{j \le} y_i,\ d_j = 1} p_j.$$

To maximize the likelihood function in (9), some iterative procedure is required. We, however, use *miCoPTCM* package in R-language to find the semiparametric ML estimates and associated metrics. The interested readers are referred to Bertrand (2015) for detail. Relevant R-code is provided in Appendix.

5. Parametric Bayesian estimation

The proposed model contains two baseline distribution parameters and two regression parameters. For the distribution parameters, we take independent gamma priors. For regression parameters, we consider improper uniform priors. The joint prior density can be written as

$$\pi(\alpha, \lambda, b_1, b_0) = \frac{(b_1)^{a_1}}{\Gamma(a_1)} \frac{(b_2)^{a_2}}{\Gamma(a_2)} \alpha^{a_1 - 1} \lambda^{a_2 - 1} \frac{e\exp(-b_1 \alpha)}{(b_3 - a_3)} \frac{\exp(-b_2 \lambda)}{(b_4 - a_4)}. \quad (10)$$

The posterior distribution of the unknown quantities is attained from (4) and (10) as

$$\pi(\gamma, \beta|O) \propto l(\gamma, \beta|O) \times \pi(\alpha, \lambda, b_1, b_0)$$

and the Bayes estimate of inference function, say $U(\alpha, \lambda, \beta_0, \beta_1)$, with respect to squared error loss function as

$$E(U(\alpha, \lambda, \beta_0, \beta_1)|O) = \frac{\int_0^\infty \int_0^\infty \int_0^\infty U(\alpha, \lambda, \beta_0, \beta_1) \pi(\gamma, \beta|O) d\alpha d\lambda d\beta_0 d\beta_1}{\int_0^\infty \int_0^\infty \int_0^\infty \pi(\gamma, \beta|O) d\alpha d\lambda d\beta_0 d\beta_1}. \quad (11)$$

However, explicit expressions for Bayes estimates are not possible, some Markov chain Monte Carlo (MCMC) methods are required to solve the expression in (11) numerically. We, however, use R2OpenBUGS package in R software to find the Bayes estimates and associated metrics. Relevant R codes are provided in the appendix. For further detail about the package *R2OpenBUGS*, see Sturtz et al. (2005).

6. Simulation

A simulation study is performed to observe the behaviour of different estimators under different methods for different sample sizes, different parameters values, different censoring proportions ρ_0 (censoring proportion in control group and ρ_1 (censoring

proportion in treatment group) and for different cure proportions p_0 (cure proportion in control group) and p_1 (cure proportion in treatment group). Specifically, we consider sample sizes of 200, 300, 400; censoring rates from 20% to 70%; cure rates from 10% to 50%; increasing and decreasing hazards rates. For each setting, 500 samples are simulated from the promotion time cure rate model according to two-group clinical trial as follows. For each sample size, the patients are first randomly allocated to one of the treatment arms according to binary covariate $X \sim \text{Bern}(0.5)$, one for treatment group and zero for control group. Their event times are then generated from the promotion time cure rate model using Burr-XII as the baseline distribution function with specified parameters and exponential link function with specified cure rates. Then censoring times are generated from the truncated Burr-XII distribution by setting the parameters in such a way to achieve the desire levels of censoring rates. For Bayesian estimation of parameters, we use improper uniform priors for β_0 and β_1 over the interval (–10, +10), vague gamma prior for the shape parameters α and λ with hyperparameters equal to 1 and 0.01 (this guaranteed the posterior distribution to be proper).

The Bayes estimates (Ests), parametric and semiparametric ML estimates and the corresponding mean square errors (MSEs), empirical variances (EVs) and 95% confidence/credible interval coverage percentages (CPs) based on 500 simulated samples are obtained for each setting and are in Tables 1-5. It is observed that the estimators under each method tends to overestimate the corresponding parameter with smaller amount of bias for semiparametric method. The bias decreases as sample size increases. The rate of depreciation of bias is higher for ML method than the others. It is also observed that the estimators are unbiased in estimating the corresponding parameters as the MSEs and EVs are almost equal. For small sample sizes, the Bayes method perform better than the parametric and semiparametric methods in terms of bias, MSE and empirical variance. For large sample sizes, the performance of the Bayes and ML methods is approximately equal. When comparing the coverage percentages, we note that the parametric and semiparametric ML methods provide better coverage than the Bayesian method. We further observe that the Bayesian method has narrow lengths of 95% credible intervals and low coverage percentages while the ML methods (parametric and semiparametric) have wide lengths of confidence intervals and high coverage percentages. One characteristic that is common to all three methods is that they do not much fluctuate in estimating the cure fractions in both treatment groups.

Table 1: Average values of the estimates and corresponding MSEs, EVs and 95% CPs when $\beta_0 = 0.85$, $\beta_1 = -0.25$, $\alpha = 1.5$, $\lambda = 2$, $p_0 = 0.10$, $p_1 = 0.16$, $\rho_0 = 0.20$ and $\rho_1 = 0.30$.

(*a*)

		PML				Bayes			
n	Stat	β_0	β_1	α	λ	β_0	β_1	α	λ
200	Est	0.877	-0.250	1.493	2.010	0.863	-0.254	1.547	2.020
	EV	0.033	0.030	0.112	0.023	0.032	0.031	0.114	0.022
	MSE	0.034	0.030	0.112	0.023	0.032	0.031	0.116	0.023
	CP	93.600	92.600	90.800	90.000	89.400	90.000	88.400	90.400
300	Est	0.870	-0.244	1.479	1.998	0.861	-0.251	1.520	2.005
	EV	0.017	0.018	0.070	0.014	0.017	0.020	0.073	0.014
	MSE	0.017	0.018	0.070	0.014	0.017	0.020	0.073	0.014
	CP	95.800	95.400	88.800	91.800	93.000	93.000	88.000	90.200
400	Est	0.867	-0.254	1.492	1.994	0.857	-0.256	1.529	1.999
	EV	0.013	0.013	0.058	0.010	0.014	0.013	0.061	0.011
	MSE	0.013	0.013	0.058	0.010	0.014	0.013	0.062	0.011
	CP	93.600	96.000	87.400	90.800	91.200	94.800	84.000	89.800

(*b*)

		PML		SML				Bayes	
n	Stat	p_0	p_1	β_0	β_1	p_0	p_1	p_0	p_1
200	Est	0.095	0.158	0.865	-0.252	0.097	0.162	0.101	0.166
	EV	0.001	0.002	0.029	0.030	0.001	0.002	0.001	0.002
	MSE	0.001	0.003	0.029	0.030	0.001	0.002	0.001	0.002
	CP	93.600	99.800	95.000	92.800	95.000	97.000	89.400	91.200
300	Est	0.095	0.157	0.875	-0.246	0.094	0.157	0.099	0.163
	EV	0.001	0.002	0.020	0.018	0.001	0.002	0.001	0.002
	MSE	0.001	0.002	0.021	0.018	0.001	0.002	0.001	0.002
	CP	95.800	99.800	91.200	95.800	91.200	99.400	93.000	91.600
400	Est	0.095	0.160	0.877	-0.255	0.093	0.159	0.098	0.165
	EV	0.001	0.002	0.020	0.013	0.001	0.002	0.001	0.001
	MSE	0.001	0.002	0.020	0.013	0.001	0.002	0.001	0.001
	CP	93.600	99.400	86.400	95.800	86.400	98.200	91.200	89.400

Table 2: Average values of the estimates and corresponding MSEs, EVs and 95% CPs when $\beta_0 = 0.5,\ \beta_1 = -0.35,\ \alpha = 1.5,\ \lambda = 2,\ p_0 = 0.19,\ p_1 = 0.31,\ \rho_0 = 0.35$ and $\rho_1 = 0.50$.

(*a*)

		PML				Bayes			
n	Stat	β_0	β_1	α	λ	β_0	β_1	α	λ
200	Est	0.499	-0.344	1.549	2.029	00.490	-0.355	1.600	2.040
	EV	0.023	0.036	0.088	0.022	0.022	0.036	0.093	0.023
	MSE	0.023	0.036	0.090	0.023	0.022	0.036	0.102	0.024
	CP	96.600	94.600	95.400	95.400	95.000	93.200	91.800	96.000
300	Est	0.481	-0.336	1.570	2.026	0.477	-0.346	1.607	2.033
	EV	0.016	0.022	0.060	0.014	0.016	0.023	0.063	0.014
	MSE	0.016	0.023	0.065	0.015	0.017	0.023	0.074	0.015
	CP	96.000	96.400	95.600	95.400	94.000	95.800	91.200	94.400
400	Est	0.495	-0.349	1.544	2.027	0.488	-0.356	1.580	2.034
	EV	0.012	0.020	0.050	0.011	0.012	0.020	0.051	0.011
	MSE	0.012	0.020	0.052	0.012	0.012	0.020	0.057	0.012
	CP	94.400	93.200	92.800	96.200	93.800	91.800	90.200	95.000

(*b*)

		PML		SML				Bayes	
n	Stat	p_0	p_1	β_0	β_1	p_0	p_1	p_0	p_1
200	Est	0.195	0.312	0.488	-0.344	0.198	0.316	0.200	0.320
	EV	0.002	0.004	0.022	0.036	0.002	0.004	0.002	0.004
	MSE	0.002	0.004	0.022	0.036	0.002	0.004	0.002	0.004
	CP	96.600	98.000	96.800	94.400	96.800	97.000	95.000	92.000
300	Est	0.200	0.315	0.472	-0.336	0.203	0.319	0.203	0.321
	EV	0.002	0.002	0.015	0.023	0.002	0.003	0.002	0.003
	MSE	0.002	0.002	0.016	0.023	0.002	0.003	0.002	0.003
	CP	96.000	96.000	96.600	96.400	96.600	99.800	94.000	92.600
400	Est	0.195	0.315	0.492	-0.349	0.196	0.316	0.199	0.321
	EV	0.001	0.002	0.015	0.020	0.001	0.002	0.001	0.002
	MSE	0.001	0.002	0.015	0.020	0.001	0.002	0.001	0.002
	CP	94.400	97.000	92.800	93.000	92.800	99.600	93.800	92.000

Table 3: Average values of the estimates and corresponding MSEs, EVs and 95% CPs when $\beta_0 = 0.20, \beta_1 = -0.50, \alpha = 1.5, \lambda = 2, p_0 = 0.29, p_1 = 0.48, \rho_0 = 0.50$ and $\rho_1 = 0.70$.

(*a*)

		PML				Bayes			
n	Stat	β_0	β_1	α	λ	β_0	β_1	α	λ
200	Est	0.203	-0.504	1.552	2.033	0.201	-0.515	1.596	2.042
	EV	0.024	0.037	0.117	0.029	0.025	0.036	0.119	0.030
	MSE	0.024	0.037	0.120	0.030	0.025	0.036	0.128	0.032
	CP	95.600	95.600	95.200	92.000	93.400	95.200	91.400	92.000
300	Est	0.193	-0.503	1.552	2.022	0.189	-0.511	1.594	2.031
	EV	0.019	0.030	0.086	0.019	0.019	0.030	0.086	0.020
	MSE	0.019	0.030	0.089	0.020	0.019	0.030	0.095	0.021
	CP	94.600	93.400	96.600	94.600	92.200	92.000	92.600	93.200
400	Est	0.196	-0.506	1.554	2.020	0.187	-0.513	1.600	2.028
	EV	0.016	0.021	0.074	0.015	0.014	0.022	0.072	0.015
	MSE	0.016	0.021	0.076	0.015	0.014	0.022	0.082	0.016
	CP	95.200	95.400	94.200	94.600	93.800	93.000	91.600	93.600

(*b*)

		PML		SML				Bayes	
n	Stat	p_0	p_1	β_0	β_1	p_0	p_1	p_0	p_1
200	Est	0.295	0.476	0.188	-0.505	0.300	0.481	0.297	0.479
	EV	0.003	0.005	0.024	0.037	0.003	0.005	0.003	0.005
	MSE	0.003	0.005	0.024	0.037	0.003	0.005	0.003	0.005
	CP	95.600	98.000	95.800	95.600	95.800	98.600	93.400	92.400
300	Est	0.298	0.479	0.183	-0.503	0.302	0.483	0.300	0.482
	EV	0.003	0.003	0.019	0.030	0.002	0.003	0.002	0.003
	MSE	0.003	0.003	0.019	0.030	0.003	0.003	0.002	0.003
	CP	94.600	98.200	93.400	93.600	93.400	99.000	92.200	93.400
400	Est	0.297	0.479	0.180	-0.505	0.303	0.485	0.301	0.484
	EV	0.002	0.003	0.015	0.022	0.002	0.002	0.002	0.003
	MSE	0.002	0.003	0.015	0.022	0.002	0.002	0.002	0.003
	CP	95.200	97.800	94.800	95.400	94.800	96.000	93.800	92.200

Table 4: Average values of the estimates and corresponding MSEs, EVs and 95% CPs when $\beta_0 = 0.50$, $\beta_1 = -0.35$, $\alpha = 1.5$, $\lambda = 1.5$, $p_0 = 0.19$, $p_1 = 0.31$, $\rho_0 = 0.35$ and $\rho_1 = 0.50$.

(*a*)

		PML				Bayes			
n	Stat	β_0	β_1	α	λ	β_0	β_1	α	λ
200	Est	0.506	-0.357	1.527	1.513	0.499	-0.367	1.573	1.522
	EV	0.027	0.035	0.096	0.013	0.026	0.036	0.098	0.013
	MSE	0.027	0.035	0.097	0.013	0.026	0.036	0.103	0.014
	CP	93.800	94.600	92.800	94.800	93.600	92.000	90.400	93.800
300	Est	0.524	-0.348	1.475	1.501	0.517	-0.355	1.511	1.508
	EV	0.017	0.025	0.067	0.010	0.017	0.025	0.068	0.010
	MSE	0.018	0.025	0.068	0.010	0.018	0.025	0.068	0.010
	CP	95.800	95.200	90.800	93.000	93.000	93.800	89.200	92.800
400	Est	0.521	-0.359	1.489	1.501	0.516	-0.365	1.518	1.506
	EV	0.014	0.017	0.054	0.007	0.014	0.018	0.055	0.007
	MSE	0.015	0.017	0.054	0.007	0.015	0.018	0.055	0.007
	CP	95.000	95.000	92.000	93.600	93.400	93.800	90.200	93.000

(*b*)

		PML		SML				Bayes	
n	Stat	p_0	p_1	β_0	β_1	p_0	p_1	p_0	p_1
200	Est	0.193	0.314	0.509	-0.357	0.193	0.313	0.198	0.321
	EV	0.003	0.004	0.034	0.035	0.003	0.004	0.002	0.004
	MSE	0.003	0.004	0.034	0.035	0.003	0.004	0.002	0.004
	CP	93.800	95.000	91.400	95.000	91.400	96.000	93.6.000	92.000
300	Est	0.187	0.304	0.543	-0.348	0.183	0.299	0.190	0.310
	EV	0.002	0.003	0.033	0.024	0.003	0.004	0.002	0.003
	MSE	0.002	0.003	0.035	0.024	0.003	0.005	0.002	0.003
	CP	95.800	94.000	88.000	95.400	88.000	98.000	93.000	92.800
400	Est	0.187	0.309	0.543	-0.359	0.183	0.302	0.190	0.314
	EV	0.001	0.002	0.031	0.017	0.002	0.004	0.001	0.002
	MSE	0.001	0.002	0.033	0.017	0.002	0.004	0.001	0.002
	CP	95.000	96.000	89.000	95.200	89.000	97.400	93.400	93.800

Table 5: Average values of the estimates and corresponding MSEs, EVs and 95% CPs when $\beta_0 = 0.50$, $\beta_1 = -0.35$, $\alpha = 1.5$, $\lambda = 2.5$, $p_0 = 0.19$, $p_1 = 0.31$, $\rho_0 = 0.35$ and $\rho_1 = 0.50$.

(a)

		PML				Bayes			
n	Stat	β_0	β_1	α	λ	β_0	β_1	α	λ
200	Est	0.514	-0.368	1.528	2.521	0.506	-0.375	1.573	2.529
	EV	0.025	0.034	0.079	0.033	0.026	0.035	0.083	0.034
	MSE	0.025	0.034	0.080	0.034	0.026	0.035	0.088	0.035
	CP	95.400	93.800	95.400	95.400	93.200	92.600	93.800	94.400
300	Est	0.490	-0.343	1.540	2.523	0.481	-0.346	1.586	2.529
	EV	0.016	0.027	0.056	0.021	0.017	0.028	0.063	0.022
	MSE	0.016	0.027	0.057	0.022	0.017	0.028	0.070	0.023
	CP	95.600	93.400	96.600	95.800	94.200	91.000	93.800	94.800
400	Est	0.503	-0.355	1.517	2.517	0.493	-0.356	1.558	2.520
	EV	0.014	0.018	0.042	0.018	0.016	0.019	0.049	0.018
	MSE	0.014	0.018	0.042	0.018	0.016	0.0189	0.052	0.018
	CP	94.200	94.600	95.600	95.600	92.400	93.600	91.800	94.600

(b)

		PML		SML				Bayes	
n	Stat	p_0	p_1	β_0	β_1	p_0	p_1	p_0	p_1
200	Est	0.191	0.316	0.506	-0.369	0.193	0.318	0.196	0.322
	EV	0.002	0.004	0.025	0.034	0.003	0.004	0.002	0.004
	MSE	0.002	0.004	0.025	0.034	0.002	0.004	0.002	0.004
	CP	95.400	96.000	95.000	93.800	95.000	96.000	93.200	92.600
300	Est	0.197	0.315	0.486	-0.343	0.198	0.316	0.202	0.320
	EV	0.002	0.003	0.016	0.027	0.002	0.003	0.002	0.003
	MSE	0.002	0.003	0.016	0.027	0.002	0.003	0.002	0.003
	CP	95.600	95.800	95.000	93.400	95.000	95.400	94.200	92.000
400	Est	0.193	0.314	0.497	-0.355	0.195	0.316	0.198	0.319
	EV	0.001	0.002	0.014	0.018	0.001	0.002	0.002	0.002
	MSE	0.001	0.002	0.014	0.018	0.001	0.002	0.002	0.002
	CP	94.200	95.000	93.600	94.200	93.600	95.800	92.400	92.000

7. Real data analysis

We fit the promotion time cure model to the data available in Cai (2013) by parametric and semiparametric methods with binary covariate X taking value 1 for autologous treatment group and 0 for allogeneic treatment group. Table 6 reports the estimates of parameters under different methods with standard errors in parentheses. We observe that the estimators of regression coefficients have similar standard errors for all three methods. The same data have been analyzed by different researchers assuming different models and

Table 6: Estimates of parameters β_0, β_1, α, λ and corresponding standard errors (in parentheses) for the bone marrow transplant data.

Method	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\lambda}$
Parametric ML	0.2057 (0.1821)	0.4128 (0.2433)	2.0508 (0.3884)	1.4937 (0.1347)
Bayes	0.2020 (0.1828)	0.4111 (0.2453)	2.0494 (0.4055)	1.4854 (0.1317)
Semiparametric ML	0.2486 (0.1764)	0.4144 (0.2576)		

Table 7: Comparison of cure fractions (in percentage) for different cure rate models and for differ methods of estimation for the real data.

Model	Allogeneic Treatment (p_0)	Autologous Treatment (p_1)	Overall (Average)
Kaplan–Meier survival curve	27.00	19.00	23.00
Parametric PTCM	29.28	15.63	22.45
Bayesian PTCM	29.54	16.22	22.88
Semiparametric PTCM	27.74	14.36	21.05
Semiparametric AFT mixture cure	27.23	19.64	23.44
Parametric AFT mixture cure	26.85	19.53	23.19

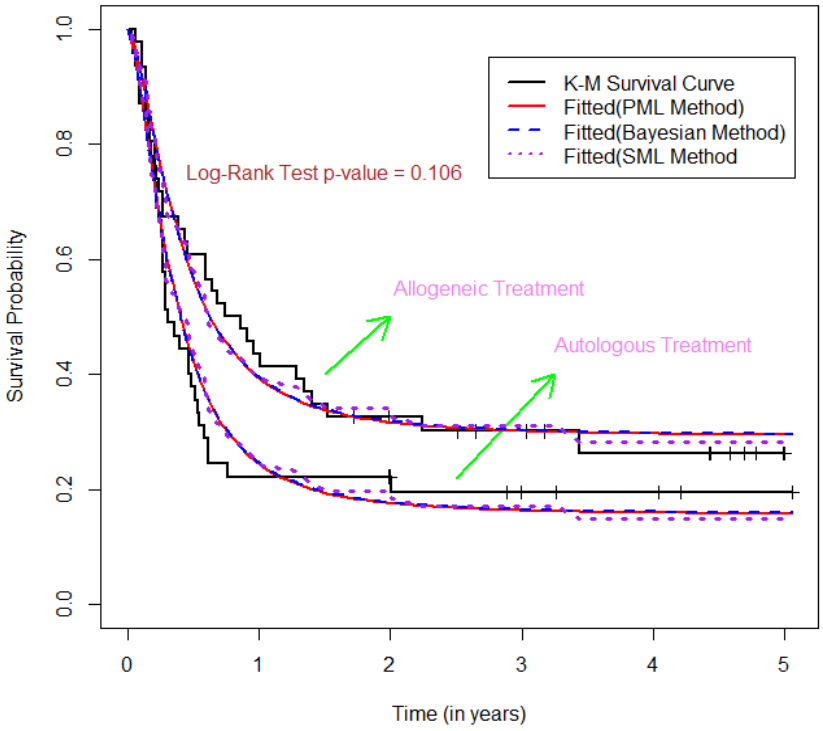

Figure 1: The survival functions fitted to the Kaplan–Meier survival curve for the allogeneic and autologous treatments.

methods, see for example, Maller and Zhou (1996), Peng et al. (2001), Zhang and Peng (2007) and estimated the cure rates for both treatment groups. To make the comparisons more realistic, we average the cure fractions estimated by different methods under different models for the allogeneic and autologous treatments in Table 7. The overall estimate of the cure fraction for the Kaplan–Meier survival estimator method is the average of the cure fractions obtained directly from the Kaplan–Meier survival curve in Figure 1 and it is considered as a standard. We see in Table 7 that the overall estimate of the cure fraction from promotion time cure model under Bayesian method is closest to 23% than the others. To further compare the performance of our methods, we plot the estimated survival function of the promotion time cure model along with the Kaplan–Meier survival curve in Figure 1. It can be noted that the estimated survival functions fit the Kaplan–Meier survival curve very well.

8. Conclusion

Several distributions including exponential, gamma, Weibull, lognormal, log-logistic, etc., are commonly used as the baseline distributions for cure models. It is pointed out that the Burr type XII distribution is not much studied in cure models particularly in the promotion time cure model. However, it is observed that much of the region covered by the gamma and lognormal distributions in the skewness-kurtosis plane is also covered by the Burr type XII distribution. It is seen that the survival data from clinical trials often show the proportion of patients with long-term survivors and the standard models like Cox and accelerated failure time model are inappropriate for fitting such data. The two-component mixture cure model is often used for this purpose. However, for the failure time data with proportional hazards structure, the promotion time cure model can be more adaptable than the mixture cure model. In this study we compare the semiparametric method with the parametric Bayesian and maximum likelihood methods for estimating the proportion of insusceptible patients using log link function and for modeling the failure times of susceptible subjects using Burr-XII distribution. For Bayesian estimation, we use improper uniform prior distributions for regression parameters and vague gamma prior distributions for baseline distribution parameters. Numerical experiments are considered to examine the performance of different methods. One real data application is considered for the illustration of the proposed methods.

Acknowledgement

The authors thank the referees and the associate editor for their useful suggestions to improve the article.

Appendix

R code for parametric ML inference

```
library(maxLik)
LLF <- function(pars, data) {
paras <- pars[1:4]
p1 <- pars[1]
p2 <- pars[2]
p3 <- pars[3]
p4 <- pars[4]
```

```
x <- data[,1]
d <- data[,2]
z <- data[,3]
r <- sum(d)
p1*r+p2*sum(d*z)+r*log(p3)+r*log(p4)+p4*sum(d*log(x))-
(p3+1)*sum(d* log(1+x^p4))-
exp(p1)*sum(exp(p2*z))+exp(p1)*sum(exp(p2*z)/(1+x^p4)^p3)
}
MLfit <- maxLik(LLF, start = c(0.2, 0.40, 2.5, 1.5), data =
data)
est <- MLfit$estimate
SE <- sqrt(diag(vcov(MLfit)))
MB0 <- est[1]
MB1 <- est[2]
MA <- est[3]
ML <- est[4]
Mp0 <- exp(-exp(MB0))
Mp1 <- exp(-exp(MB0+MB1)
MB0; MB1; MA; ML; MF0; MF
SE
```

R code for Bayesian inference

```
model {
      for (i in 1:400)
         {
    pi[i] <- exp(-exp(beta0+beta1*z[i]))
    S[i] <- pow((1+pow(x[i], lambda)), -alpha)
    f[i] <- alpha*lambda*pow(x[i],lambda-
1)*pow((1+pow(x[i], lambda)),
           -alpha-1)
    L[i] <- pow(-log(pi[i])*f[i], d[i])*pow(pi[i],(1-S[i]))
    LogL[i] <- log(L[i])
    zeros [i] <- 0
    zeros[i] ~ dloglik(LogL[i])
         }
      beta0 ~ dunif(-10, 10)
      beta1 ~ dunif(-10, 10)
      alpha ~ dgamma(1, 0.01)
      lambda ~ dgamma(0, 0)
       }
library("R2OpenBUGS")
parameters <- c("beta0", "beta1", "alpha", "lambda")
inits1 <- list(beta0 = 0.15, beta1 = 0.40, alpha = 1.5,
lambda = 1.4)
inits2 <- list(beta0 = 0.16, beta1 = 0.50, alpha = 2.5,
lambda = 1.6)
inits <- list(inits1, inits2)
Bayesfit <- bugs(data, inits, parameters, model.file =
"G:/ptcm(b).R", n.chains = 2, n.iter = 10000, n.burnin =
5000, DIC = T)
```

```
B0 <- Bayesfit$sims.array[,,"beta0"]
B1 <- Bayesfit$sims.array[,,"beta1"]
A <- Bayesfit$sims.array[,,"alpha"]
L <- Bayesfit$sims.array[,,"lambda"]
p01 <- exp(-exp(B0))
p11 <- exp(-exp(B0+B1))
BB0 <- mean(B0)
BB1 <- mean(B1)
Bp0 <- mean(p01)
Bp1 <- mean(p11)
BA <- mean(A)
BL <- mean(L)
BSB0 <- sd(B0)
BSB1 <- sd(B1)
BSF0 <- sd(p01)
BSF <- sd(p11)
BSA <- sd(A)
BSL <- sd(L)
BB0; BB1; BA; BL; Bp0; Bp1
BSB0; BSB1; BSA; BSL; BSp0; BSp1
```

R code for semiparametric ML inference

```
library(miCoPTCM)
vc <- matrix(nrow=2,ncol=2,0)
SMLfit <- PTCMestimBF(formula=Surv(x, d)~z, data=data,
varCov=vc, init=runif(2))
Sest <- SMLfit$coefficients
SSE <- sqrt(diag(SMLfit$vcov))
SB0 <- Sest[1]
SB1 <- Sest[2]
Sp0 <- exp(-exp(SB0))
Sp1 <- exp(-exp(SB0+SB1))
SB0; SB1; Sp0; Sp1
SSE
```

References

1. Abbas, K., Hussain, Z., Rashid, N., Ali, A., Taj, M., Khan, S. A., Manzoor, S., Khalil, U., & Khan, D. M. (2020). Bayesian estimation of Gumbel type-II distribution under type-II censoring with medical applications. *Computational and Mathematical Methods in Medicine*, *2020*, 1-11. https://doi.org/10.1155/2020/1876073
2. Bertrand, A., Legrand, C., Carroll, R. J., De Meester, C., & Van Keilegom, I. (2015). Inference in a survival cure model with mismeasured covariates using a simulation-extrapolation approach, *Biometrika*, *104(*1), 31-50. https://doi.org/10.1093/biomet/asw054
3. Botta, L., Goungounga, J., Capocaccia, R., Romain, G., Colonna, M., Gatta, G., Boussari, O., & Jooste, V. (2023). A new cure model that corrects for increased risk of non-cancer death: analysis of reliability and robustness, and application to real-

life data. *BMC Medical Research Methodology*, *23*(1), 1-19. https://doi.org/10.118/s12874-023-01876-x

4. Berkson, J., & Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association, 47*, 501-515. https://doi.org/10.1080/01621459.1952.10501187
5. Balakrishnan, N., & Pal, S. (2015). Likelihood inference for flexible cure rate models with gamma lifetimes. *Communications in Statistics-Theory and Methods*, *44*(19), 4007-4048. https://doi.org/10.1080/03610926.2014.964807
6. Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of Royal Statistical Society, Series C, 29*, 15-23. https://doi.org/10.1111/j.2517-6161.1949.tb00020.x
7. Cai, C. (2013). Advanced methodology developments in mixture cure models (doctoral dissertation). https://scholarcommons.sc.edu/etd/544/
8. Chen, M. H., Ibrahim, J. G., & Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of Royal Statistical Society, Series C, 94*, 909-19. https://doi.org/10.2307/2670006
9. Chen, M. H., & Ibrahim, J. G. (2001). Maximum likelihood methods for cure rate models with missing covariates. *Biometrics, 57*, 43-52. https://doi.org/10.1111/j.0006-341x.2001.00043.x
10. Chen, T., & Du P. (2018). Promotion time cure rate model with nonparametric form of covariate effects. *Statistics in Medicine, 37*, 1625-1635. https://doi.org/10.1002/sim.7597
11. Coelho-Barros, E. A., Achcar, J. A., & Mazucheli, J. (2017). Cure rate models considering the Burr XII distribution in presence of covariate. *Journal of Statistical Theory and Applications*, *16*(2), 150-164. https://doi.org/10.2991/jsta.2017.16.2.2
12. D'Andrea, A., Rocha, R., Tomazella, V., & Louzada, F. (2018). Negative binomial Kumaraswamy-G cure rate regression model. *Journal of Risk and Financial Management, 11*, 1-14. https://doi.org/10.3390/jrfm11010006
13. Ezquerro, A., Cancela, B., & López-Cheda, A. (2023). On the Reliability of Machine Learning Models for Survival Analysis When Cure Is a Possibility. *Mathematics*, *11*(19), 4150.
14. Escobar-Bach, M., & Van Keilegom, I. (2023). Nonparametric estimation of conditional cure models for heavy-tailed distributions and under insufficient follow-up. *Computational Statistics and Data Analysis, 183*, 107728. https://doi.org/10.1016/j.csda.2023.107728
15. Henningsen, A., & Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, *26*, 443-458. https://doi.org/10.1007/s00180-010-0217-1
16. Lambert, P. C., Thompson, J. R., Weston, C. L., & Dickman, P. W. (2007). Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics, 8*, 576-594. https://doi.org/10.1093/biostatistics/kxl030
17. Liu, Y., Ilyas, M., Khosa, S. K., Muhmoudi, E., Ahmad, Z., Khan, D. M., & Hamedani, G. G. (2020). A flexible reduced logarithmic-X family of distributions with biomedical analysis. *Computational and Mathematical Methods in Medicine, 2020*. https://doi.org/10.1155/2020/4373595
18. Maller, R. A., & Zhou, X. (1996). *Survival Analysis with Long-term Survivors*. 1st Edition. Wiley, Chichester.
19. Pedrosa-Laza, M., López-Cheda, A., & Cao, R. (2022). Cure models to estimate time until hospitalization due to COVID-19: A case study in Galicia (NW

Spain). *Applied Intelligence*, *52*(1), 794-807. https://doi.org/10.1007/s10489-021-02311-8

20. Peng, Y., Dear, K. B. G., & Carriere, K.C. (2001). Testing for the presence of cured patients: a simulation study. *Statistics in Medicine, 20,* 1783-1796. https://doi.org/10.1002/sim.781
21. Portier, F., Ghouch, A. E., & Keilegom, I. V. (2017). Efficiency and bootstrap in the promotion time cure model. *Bernoulli*, *23,* 3437-3468. https://doi.org/10.3150/16-BEJ852
22. Rodriguez, R. N. (1977). A guide to the Burr type XII distributions. *Biometrika*, 64 (1), 129-134. https://doi.org/10.2307/2335782
23. Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical software*, *12*, 1-16. https://doi.org/10.18637/jss.v012.i03
24. Tsodikov, A. D., Ibrahim, J. G., & Yakovlev, A. Y. (2003). Estimating cure rates from survival data. *Journal of the American Statistical Association, 98*, 1063-1078. https://doi.org/10.1198/016214503000001007
25. Wang, X., & Han, B. (2020). Efficient estimation for the non-mixture cure model with current status data. *Statistics*, *54*, 756-777. https://doi.org/10.1080/02331888.2020.1783541
26. Zeng, D., Yin, G., & Ibrahim, J. G. (2006). Semiparametric transformation models for survival data with a cure fraction. *Journal of the American Statistical Association,* 101, 670-684. https://doi.org/10.1198/016214505000001122
27. Zhang, J., & Peng, Y. (2007). A new estimation method for the semiparametric accelerated failure time mixture cure model. *Statistics in Medicine, 26*, 3157-3171. https://doi.org/10.1002/sim.2748