Shahid Chamran University of Ahvaz

Journal of Advanced Mathematical Modeling
ISSN (print): 2251-8088, ISSN (on-line): 2645-6141
Vol. 16 No. 2 (2026), pp. 33-55.
DOI: 10.22055/jamm.2026.49429.2370

A MIXED POISSON WEIGHTED LINDLEY FRAILTY APPROACH TO MODEL BIVARIATE CURE SURVIVAL DATA

MARZIYE ROUZBAHANI, MOHAMMAD REZA AKHOOND* AND RAHIM CHINIPARDAZ

Communicated by: Golam Ali Parham

Abstract. Modeling the dependence structure of bivariate survival data is one of the main issues in survival analysis. Copula and frailty models are widely used to model dependence in such data. The frailty variable is usually assumed to follow a continuous distribution; however, when a cure fraction is present, it can be advantageous to model the frailty as a discrete random variable, with zero frailty representing immunity or cure. We propose a frailty–copula approach induced by a discrete frailty distribution to model bivariate survival data with a proportion of long-term survivors. Specifically, we assume that the frailty follows a mixed Poisson distribution with a mixing distribution following the weighted Lindley distribution. Model parameters are estimated by maximum likelihood, and the performance of the estimators is assessed via a Monte Carlo simulation study. We illustrate the proposed model using data from a retrospective cohort of 143 smokers (tobacco-stained fingers dataset) screened between March 2006 and January 2010 at a 180-bed community hospital in La Chaux-de-Fonds, Switzerland.

1. Introduction

Survival analysis is concerned with the time until an event occurs. Traditionally, it is assumed that every individual in the study population is susceptible to the event under investigation and will eventually experience it if the follow-up period is sufficiently long.

MSC(2020): 62N02, 62H05.

Keywords: Cure rate model, Copula, discrete frailty, Frailty model, Weighted Lindley distribution.

Received: 19 April 2025, Accepted: 16 June 2026.

*Corresponding author.

However, in some situations, this assumption may not hold, as the event will not occur for every subject in the population. For instance, in a cancer study, a fraction of individuals may respond positively to the treatment. These individuals will not experience the event of interest during the follow-up period and exhibit long-term survival; that is, they are considered cured or insusceptible. More generally, there may be two observed lifetimes associated with the same individual, and in some cases, a dependence structure may exist between them. To model the dependence between survival times, one requires a bivariate distribution. Several methods exist for constructing such distributions, each with distinct interpretations and motivations.

One common method is the use of a copula function. Copula models have become a natural and suitable tool for modeling dependence in multivariate data. A copula is a multivariate cumulative distribution function with uniform margins on the [0,1] interval. Different copula functions represent different dependence structures between variables [20]. [23] employed the Archimedean copula family to model the dependence between bivariate lifetime components, assuming a Weibull distribution for the marginals. [16] introduced a bivariate FGM survival copula model for long-term survivors. Their approach assumed a mixture cure rate model for the marginal distribution of each lifetime and a fixed cure fraction for the entire population. Other researchers who have utilized copula functions to model bivariate survival data include [22, 21, 19, 24, 10] and [14].

Frailty models provide another prominent approach. They are extensively used in survival analysis to account for unobserved heterogeneity and dependence among individuals [26]. In these models, frailty is represented as an unobserved random variable that acts multiplicatively on the baseline hazard function. If this heterogeneity is ignored, it can lead to several issues, including an overestimation of the relative hazard rate, biased estimates of regression coefficients, and a tendency for regression parameters to attenuate toward zero. The formal concept of frailty was introduced by [26] in the context of univariate survival data, while [3] had earlier applied a similar idea to address associations in bivariate life tables.

Studies employing frailty models typically assume a non-negative continuous frailty random variable. Frequently utilized frailty distributions include the gamma [26], inverse Gaussian, log-normal, and positive stable distributions [12]. More recently, the weighted

Lindley (WL) distribution has also been proposed for frailty modeling [9]. The WL distribution can be expressed as a two-component mixture of gamma distributions, which facilitates the derivation of desirable statistical properties, such as a closed-form Laplace transform and the efficient generation of random samples [9, 17, 18]. Like other standard frailty distributions, the WL frailty model benefits from a simple Laplace transform. This allows for closed-form expressions of the marginal survival and hazard functions and simplifies parameter estimation [18].

A limitation of continuous frailty distributions is their inability to accommodate a subpopulation with zero risk. A frailty value of zero signifies the existence of a non-susceptible subgroup that will never experience the event of interest, even after an extended follow-up. In studies such as those concerning cancer recurrence, the event of interest might be patient death or tumor relapse. In such contexts, a discrete frailty distribution can be more appropriate. Discrete frailty distributions naturally incorporate a point mass at zero, enabling them to model survival data with a cure fraction. These models can also be derived from proportional hazards models with discrete frailty terms [4].

The frailty model assumes a proportional hazards structure conditional on the random effect Z. Thus, the model is specified by the conditional hazard function $h(t \mid Z) = Zh_0(t)$, where $h_0(t)$ denotes the baseline hazard function. The corresponding conditional survival function is given by :

$$S(t|z) = P(T > t|z) = e^{-zH_0(t)} = S_0(t)^z, \tag{1.1}$$

where $H_0(t) = \int_0^t h_0(u)du$ is the cumulative baseline hazard function and $S_0(t)$ is the baseline survival function.

As previously noted, discrete frailty distributions can be more appropriate in certain contexts. Several authors have developed models in this vein. For instance, [28] employed a discrete compound Poisson process for frailty. [4] proposed models based on the negative binomial, Poisson, and geometric distributions. Furthermore, [5] introduced the extended Conway–Maxwell and hyper-Poisson distributions for discrete frailty modeling, with the latter representing a specific case. More recently, [2] developed multivariate survival models incorporating discrete frailty terms and adopted a Bayesian framework for parameter estimation.

Consider a frailty variable Z that is a non-negative integer-valued random variable with

a probability mass function $P(Z = z)$. The unconditional survival function is then given by

$$S(t) = \sum_{z=0}^{\infty} S_0(t)^z P(Z = z) = E[S_0(t)^z] = G_Z(S_0(t)), \tag{1.2}$$

where $G_Z(\cdot)$ is the probability generating function of Z. Since $\lim_{t\to\infty} S_0(t) = 0$ and $\lim_{t\to\infty} S(t) = G_Z(0) = P(Z = 0) > 0$, the survival function in (1.2) is improper, describing a survival model with cure fraction $P_0 = P(Z = 0) > 0$. For instance, if Z follows Poisson distribution with mean λ and probability generating function $G_Z(t) = \exp(-\lambda(1-t))$, the survival function in (1.2) becomes $S(t) = \exp(-\lambda F_0(t))$, where $F_0(t) = 1 - S_0(t)$ is the baseline cumulative distribution function.

[7, 6, 8] proposed a frailty-copula model for clustered survival data that simultaneously models heterogeneity and dependence. They used frailty terms to account for cluster heterogeneity and Clayton copula to model dependence between failure times. [15] developed a frailty-copula method for recurrent events and terminal events, while [27] proposed frailty-copula models for competing risks data under Gumbel copula. Our main objective is to introduce a bivariate long-term survival model based on the frailty-copula framework of [7, 6, 8]. We consider a mixed Poisson distribution for the frailty variable with mean λW, where $\lambda > 0$ is a real constant and W is a frailty component. [25] used the gamma distribution as a frailty effect. The Weighted Lindley distribution is one of the distributions that, due to possessing increasing and bathtub-shaped hazard functions, has been increasingly used in recent years for modeling survival and reliability data. This distribution also has the characteristic that it can be written as a mixture of two gamma distributions and is essentially a generalized form of the gamma distribution. This grants it greater flexibility in data modeling compared to the Gamma distribution [18]. An advantage of our proposed model is its ability to jointly model heterogeneity and dependence in the presence of cure fractions. The paper is organized as follows: Section 2 describes the weighted Lindley distribution and copula functions, introduces the frailty-copula modeling of bivariate cure survival data, and discusses maximum likelihood estimation. Section 4 presents simulation studies to evaluate the performance of maximum likelihood estimation. Section 5 applies the proposed models to real data, and Section 6 concludes with discussions.

2. Background

2.1. **Weighted Lindley distribution.** The probability density function (pdf) of the WL distribution is given by :

$$f(t) = \frac{\alpha^{\phi+1}}{(\alpha+\phi)\Gamma(\phi)} t^{\phi-1}(1+t)e^{-\alpha t},$$

where $\alpha > 0$ and $\phi > 0$ are shape parameters, and $\Gamma(\alpha) = \int_0^\infty t^{a-1}e^{-t}dt$ is the gamma function. The mean and variance of the WL distribution are given, respectively, by

$$E(T) = \frac{\phi(\alpha+\phi+1)}{\alpha(\alpha+\phi)} \quad and \quad Var(T) = \frac{(\phi+1)(\alpha+\phi)^2 - \alpha^2}{\alpha^2(\alpha+\phi)^2}.$$

According to [1], the mean of the frailty variable is typically assumed to be one when determining the parameters in the frailty distribution. Thus, probability distribution function of the WL distribution is given by :[18]

$$f(w) = \frac{(\sqrt{\phi(\phi+1)})^{\phi+1} w^{\phi-1}(1+w)\exp(-w\sqrt{\phi(\phi+1)})}{(\sqrt{\phi(\phi+1)}+\phi)\Gamma(\phi)}.$$

The Laplace transform of the frailty is given by :

$$L(s) = \left(1 + \frac{s\gamma(\gamma+4)}{2(\gamma+2)}\right)^{-\frac{4}{\gamma(\gamma+4)}-1} \left(1 + \frac{s\gamma}{2}\right), \quad s \in R$$

where $\gamma = 2(\phi + \sqrt{\phi(\phi+1)})^{-1}$ is the variance of the frailty distribution function.

2.2. **Copula function.** The copula model connects two event times by explicitly specifying the dependence structure between their marginal distributions. The joint survival function is given by :

$$S(t_1, t_2) = C_\theta(S_1(t_1), S_2(t_2)),$$

where C_θ is a survival copula and θ describes the degree of dependence between the two events. Among several well-known copulas, we consider the Gumbel copula, defined as

$$C_\theta(u, v) = \exp\left[-\left((-\log u)^{\theta+1} + (-\log v)^{\theta+1}\right)^{\frac{1}{\theta+1}}\right], \quad \theta \geq 0.$$

As $\theta \to 0$, independence between the two random variables is achieved, yielding $C_\theta(u, v) = uv$. The Kendall's tau correlation coefficient for the Gumbel copula is given by $\tau = \frac{\theta}{\theta+1}$.

3. Proposed model

3.1. **Bivariate cure frailty-copula model.** Suppose that T_1 and T_2 are bivariate survival times and let Z_j for $j = 1, 2$ denote the unobserved heterogeneity factors (frailties) associated with T_j. We assume that each Z_j is a discrete random variable following a mixed Poisson distribution with mean $\lambda_j W$, where $\lambda_j > 0$ is a real constant and W is a latent frailty term following a weighted Lindley distribution. The conditional marginal survival function for T_j, given $W = w$, is:

$$S_j(t_j|w) = \sum_{z_j=0}^{\infty} S_{0j}(t_j)^{z_j} \frac{e^{-\lambda_j w}(\lambda_j w)^{z_j}}{z_j!} = \exp\left(-\lambda_j w F_{0j}(t_j)\right), \tag{3.1}$$

where $F_{0j}(t_j) = 1 - S_{0j}(t_j)$ is the baseline cumulative distribution function for T_j. The frailty component W induces correlation between Z_1 and Z_2. Conditional on $W = w$, we express the bivariate survival function using the joint frailty-copula model proposed by [7]:

$$S(t_1, t_2|w) = C_\theta\left(S_1(t_1|w), S_2(t_2|w)\right). \tag{3.2}$$

Substituting (3.1) into (3.2) and applying the Gumbel copula yields the conditional bivariate survival function:

$$S(t_1, t_2|w) = \exp\left[-w\left((\lambda_1 F_{01}(t_1))^{\theta+1} + (\lambda_2 F_{02}(t_2))^{\theta+1}\right)^{\frac{1}{\theta+1}}\right]. \tag{3.3}$$

Using the Laplace transform of the weighted Lindley distribution in Equation (3.3), the unconditional joint survival function is derived as:

$$\begin{aligned} S(t_1, t_2) = &\left(1 + \frac{\gamma(\gamma+4)\left[(\lambda_1 F_{01}(t_1))^{\theta+1} + (\lambda_2 F_{02}(t_2))^{\theta+1}\right]^{\frac{1}{\theta+1}}}{2(\gamma+2)}\right)^{-\frac{4}{\gamma(\gamma+4)}-1} \\ &\times \left(1 + \frac{\gamma\left[(\lambda_1 F_{01}(t_1))^{\theta+1} + (\lambda_2 F_{02}(t_2))^{\theta+1}\right]^{\frac{1}{\theta+1}}}{2}\right). \end{aligned} \tag{3.4}$$

When $\theta = 0$, Equation (3.4) simplifies to the WL cure frailty model. The bivariate survival function (3.4) is improper, with its long-term limit given by:

$$\lim_{t_1,t_2\to\infty} S(t_1,t_2) = p_{00} = \left(1+\frac{\gamma(\gamma+4)\left(\lambda_1^{\theta+1}+\lambda_2^{\theta+1}\right)^{\frac{1}{\theta+1}}}{2(\gamma+2)}\right)^{-\frac{4}{\gamma(\gamma+4)}-1} \times \left(1+\frac{\gamma\left(\lambda_1^{\theta+1}+\lambda_2^{\theta+1}\right)^{\frac{1}{\theta+1}}}{2}\right).$$

The marginal survival functions derived from (3.4) are given by:

$$S_j(t_j) = \left(1+\frac{\gamma(\gamma+4)\lambda_j F_{0j}(t_j)}{2(\gamma+2)}\right)^{-\frac{4}{\gamma(\gamma+4)}-1} \times \left(1+\frac{\gamma\lambda_j F_{0j}(t_j)}{2}\right), \quad j=1,2. \tag{3.5}$$

Equation (3.5) reveals that the marginal survival functions exhibit a cure rate structure. The corresponding cure probabilities are:

$$p_{0j} = \lim_{t_j\to\infty} S_j(t_j) = \left(1+\frac{\gamma(\gamma+4)\lambda_j}{2(\gamma+2)}\right)^{-\frac{4}{\gamma(\gamma+4)}-1} \times \left(1+\frac{\gamma\lambda_j}{2}\right), \quad j=1,2. \tag{3.6}$$

3.2. **Likelihood function.** We now describe the likelihood function for the parameters of the bivariate WL-Gumbel cure model given in Equation (3.4). Let T_{ij} and C_{ij} denote the failure time and censoring time, respectively, for the i-th subject and the j-th outcome ($i=1,\ldots,n$; $j=1,2$). We assume that T_{ij} and C_{ij} are independent. The observed event time is $t_{ij} = \min(T_{ij}, C_{ij})$, with the failure indicator $\delta_{ij} = I(T_{ij} \le C_{ij})$, where $I(\cdot)$ is the indicator function (i.e., $\delta_{ij} = 1$ if t_{ij} is a failure time and 0 if it is censored). We assume T_{ij} follows a Weibull distribution with parameters $\boldsymbol{\varphi}_j = (\alpha_j, r_j)$. Let $\mathbf{x}_i$ be the vector of covariates for the i-th individual, and $\boldsymbol{\beta}_j$ the corresponding vector of regression coefficients. Covariates are incorporated into the model through the cure rate parameter via $\lambda_{ij} = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}_j)$ for $i=1,\ldots,n$ and $j=1,2$. Let $\boldsymbol{\varphi}$ denote the vector of all parameters from the marginal distributions of the bivariate survival times. Therefore, the likelihood function for the parameter vector $(\theta,\gamma,\boldsymbol{\varphi},\boldsymbol{\beta})$ is given by:

$$L(\theta,\gamma,\boldsymbol{\varphi},\boldsymbol{\beta}) = \prod_{i=1}^{n}\left[\left(\frac{\partial^2 S(t_{i1},t_{i2})}{\partial t_{i1}\partial t_{i2}}\right)^{\delta_{i1}\delta_{i2}} \left(-\frac{\partial S(t_{i1},t_{i2})}{\partial t_{i1}}\right)^{\delta_{i1}(1-\delta_{i2})} \times \left(-\frac{\partial S(t_{i1},t_{i2})}{\partial t_{i2}}\right)^{\delta_{i2}(1-\delta_{i1})} \left(S(t_{i1},t_{i2})\right)^{(1-\delta_{i1})(1-\delta_{i2})}\right],$$

where

$$-\frac{\partial S(t_{i1},t_{i2})}{\partial t_{ij}} = \lambda_{ij} f_{0j}(t_{ij})(\lambda_{ij}F_{0j}(t_{ij}))^{\theta} B^{-\frac{4}{\gamma(\gamma+4)}-2} \times \left(-\frac{\gamma}{2}A - \left(-\frac{4}{\gamma(\gamma+4)}-1\right)\frac{\gamma(\gamma+4)}{2(\gamma+2)}\left(1+\frac{\gamma B}{2}\right)\right), \quad j=1,2$$

$$\frac{\partial^2 S(t_{i1},t_{i2})}{\partial t_{i1}\partial t_{i2}} = \frac{\gamma}{2}\lambda_{i1} f_{01}(t_{i1})(\lambda_{i1}F_{01}(t_{i1}))^{\theta}\lambda_{i2} f_{02}(t_{i1})(\lambda_{i2}F_{02}(t_{i2}))^{\theta} B^{-(2\theta+1)} A^{-\frac{4}{\gamma(\gamma+4)}-2} \left(-\theta A + (-\frac{4}{\gamma(\gamma+4)}-1)\frac{\gamma(\gamma+4)B}{2(\gamma+2)}\right) + \left(-\frac{4}{\gamma(\gamma+4)}-1\right)\left(\frac{\gamma(\gamma+4)}{2(\gamma+2)}\right) \lambda_{i1} f_{01}(t_{i1})(\lambda_{i1}F_{01}(t_{i1}))^{\theta}\lambda_{i2} f_{02}(t_{i1})(\lambda_{i2}F_{02}(t_{i2}))^{\theta} B^{-(2\theta+1)} A^{-\frac{4}{\gamma(\gamma+4)}-3} \left(-\theta A(1+\frac{\gamma B}{2}) + (-\frac{4}{\gamma(\gamma+4)}-2)(\frac{\gamma(\gamma+4)B}{2(\gamma+2)})(1+\frac{\gamma B}{2}) + \frac{\gamma B A}{2}\right)$$

where:

$$B = \left[(\lambda_{i1}F_{01}(t_{i1}))^{\theta+1} + (\lambda_{i2}F_{02}(t_{i2}))^{\theta+1}\right]^{\frac{1}{\theta+1}},$$
$$A = 1 + \frac{\gamma(\gamma+4)B}{2(\gamma+2)},$$

$f_{0j}(t_{ij})$ is the baseline density function, and $F_{0j}(t_{ij})$ is the baseline cumulative distribution function for $j = 1, 2$.

We assume a Weibull distribution for the baseline density function $f_{0j}(t_{ij})$. The log-likelihood function for the parameter vector $\boldsymbol{\vartheta} = (\theta, \gamma, \boldsymbol{\varphi}, \boldsymbol{\beta})$ is given by:

$$\ell(\boldsymbol{\vartheta}) = \sum_{i=1}^{n}\delta_{i1}(1-\delta_{i2})\log\left(-\frac{\partial S(t_{i1},t_{i2})}{\partial t_{i1}}\right) + \sum_{i=1}^{n}\delta_{i2}(1-\delta_{i1})\log\left(-\frac{\partial S(t_{i1},t_{i2})}{\partial t_{i2}}\right) + \sum_{i=1}^{n}\delta_{i1}\delta_{i2}\log\left(\frac{\partial^2 S(t_{i1},t_{i2})}{\partial t_{i1}\partial t_{i2}}\right) + \sum_{i=1}^{n}(1-\delta_{i1})(1-\delta_{i2})\log S(t_{i1},t_{i2}).$$

The maximum likelihood estimates (MLEs) of the parameters are obtained by solving the score equations $U(\boldsymbol{\vartheta}) = \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \mathbf{0}$ iteratively using the Newton-Raphson method. Numerical optimization was implemented in R using the `optim` function with the Nelder-Mead algorithm. Successful convergence was verified by ensuring that the convergence code in the `optim` output equaled 0. The R code for model fitting is available from the corresponding author upon request.

The asymptotic properties of the MLEs are used to construct confidence intervals. Under regularity conditions, $\hat{\boldsymbol{\vartheta}}$ follows an asymptotic multivariate normal distribution

with mean $\boldsymbol{\vartheta}$ and variance-covariance matrix $\mathcal{I}^{-1}(\hat{\boldsymbol{\vartheta}})$, where $\mathcal{I}(\hat{\boldsymbol{\vartheta}})$ is the observed Fisher information matrix, defined as:

$$\mathcal{I}(\boldsymbol{\vartheta}) = -\mathbb{E}\left[\frac{\partial^2 \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}^\top}\right].$$

An approximate $100(1-\alpha)\%$ confidence interval for ϑ_i is given by

$$\hat{\vartheta}_i \pm z_{\alpha/2}\, \mathrm{SE}(\hat{\vartheta}_i),$$

where $\mathrm{SE}(\hat{\vartheta}_i) = \sqrt{\mathcal{I}^{ii}}$ is the standard error of $\hat{\vartheta}_i$, $z_{\alpha/2}$ denotes the $100(1-\alpha/2)$-th percentile of the standard normal distribution, and $\mathcal{I}^{ii}$ is the i-th diagonal element of $\mathcal{I}^{-1}$.

4. Simulation study

In this section, we present a simulation study to evaluate the performance of the proposed model under different sample sizes and censoring proportions. We assume a Weibull baseline distribution for the event times T_j, $j = 1, 2$, with shape and scale parameters $\boldsymbol{\varphi}_1 = (1.0, 0.6)$ and $\boldsymbol{\varphi}_2 = (1.5, 0.7)$, respectively.

The proposed model includes the cure rate parameter $\lambda_{ij} = \exp(\beta_{0j} + \beta_{1j} x_i)$ for $i = 1, \ldots, n$ and $j = 1, 2$, with regression coefficients $(\beta_{01}, \beta_{02}) = (1.0, 1.0)$ and $(\beta_{11}, \beta_{12}) = (-1.0, -1.0)$. The covariate x_i was generated from a Bernoulli distribution with success probability 0.5, and the frailty components W_i, $i = 1, \ldots, n$, were generated from a weighted Lindley (WL) distribution with mean 1 and variance 0.7.

Censoring times C_{ij} were generated from a uniform distribution $U(0, \varepsilon)$, where the parameter ε was adjusted to achieve the desired censoring proportions. The observed event times were then generated as follows: First, we generate bivariate correlated uniform random variables $v_{i1}, v_{i2} \sim U(0, 1)$ from the Gumbel copula using the following algorithm:

(1) Generate independent $v_{i1}, v_{i2} \sim U(0, 1)$.
(2) Set $u_{i1} = v_{i1}$.
(3) For the Gumbel copula, solve $\frac{\partial C_\theta(u_{i1}, u_{i2})}{\partial u_{i1}} = v_{i2}$ for u_{i2}.
(4) Calculate the cure probability $p_{ij} = \exp(-\lambda_{ij} w_i)$ for $i = 1, \ldots, n$ and $j = 1, 2$.
(5) For each $j = 1, 2$:
    - If $u_{ij} < p_{ij}$, set $T_{ij} = \infty$ (cured observation).

- Otherwise, generate the failure time as:

$$T_{ij} = \left(-\frac{1}{r_j} \log \left(1 + \frac{\log(u_{ij})}{w_i \lambda_{ij}} \right) \right)^{\frac{1}{\alpha_j}}.$$

(6) Generate censoring times $C_{ij} \sim U(0, \varepsilon)$.

(7) The observed data are:

$$t_{ij} = \min(T_{ij}, C_{ij}), \quad \delta_{ij} = I(T_{ij} \leq C_{ij}).$$

All simulations were conducted using R software with 1,000 Monte Carlo replications. We evaluated the effects of sample size and censoring proportion on estimator performance using sample sizes of $n = 100$ and $n = 200$, and censoring proportions of 30%, 40%, and 50%. Performance was assessed using the following measures:

$$\text{Bias} = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} (\hat{\beta}_i - \beta),$$

$$\text{SD} = \sqrt{\frac{1}{n_{\text{sim}} - 1} \sum_{i=1}^{n_{\text{sim}}} (\hat{\beta}_i - \bar{\hat{\beta}})^2},$$

$$\text{MSE} = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} (\hat{\beta}_i - \beta)^2,$$

$$\text{AB} = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} |\hat{\beta}_i - \beta|,$$

where $n_{\text{sim}} = 1000$ denotes the number of Monte Carlo replications, $\hat{\beta}_i$ represents the parameter estimate from the i-th replication, and $\bar{\hat{\beta}} = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} \hat{\beta}_i$ is the mean of the estimates across all replications.

The performance comparison between the WL-Gumbel cure model and the WL cure frailty model under 50% censoring is summarized in Tables 1 and 2 for Kendall's τ values of 0, 0.3, and 0.6, and sample sizes of 100 and 200.

The results demonstrate that the WL-Gumbel model consistently and substantially outperforms the WL frailty model across all evaluation metrics (Bias, SD, MSE, AB) and for nearly all parameters. Both bias and standard deviation for all parameters in the WL-Gumbel model remain at acceptable levels. Notably, the WL-Gumbel model accurately captures the dependence structure, with estimated τ values closely approximating the

true values of 0.3 and 0.6. In contrast, the WL frailty model exhibits poor and unstable performance, particularly for $n = 100$, with substantially larger biases.

Tables **??** and **??** present the performance of the WL-Gumbel cure model under different censoring rates. As censoring increases from 30% to 40%, we observe increased SD and MSE values for most parameters, indicating reduced precision and greater estimation uncertainty. The bias and absolute bias (AB) also generally increase under higher censoring. These findings suggest that the 40% censoring rate adversely affects both the accuracy and variability of parameter estimates. The model demonstrates significantly better performance under the moderate 30% censoring scenario. Consequently, for highly censored data, parameter estimates should be interpreted with caution, and strategies such as increasing sample size may be necessary to maintain estimation quality.

5. Real data analysis

We applied the proposed WL-Gumbel cure model to the tobacco-stained fingers dataset from a retrospective cohort study. The data were collected from 143 smokers screened between March 2006 and January 2010 at a 180-bed community hospital in La Chaux-de-Fonds, Switzerland, with follow-up for death and hospital admission until June 2014[13].

In this bivariate survival analysis, we define:

- T_1: Time to first hospital readmission for smokers with tobacco-stained fingers, censored in case of death before the study closure date (26% censoring)
- T_2: Overall survival time for patients with tobacco-tar stained fingers (48% censoring)

Figure 1 displays the Kaplan-Meier survival curves for T_1 and T_2, illustrating the survival patterns for both endpoints.

The study tracked several patient characteristics, including age, consumption of unfiltered cigarettes, and comorbidities such as vascular disease, ischemic heart disease, stroke, arterial disease, tobacco-related cancer, other cancers, hypertension, hypercholesterolemia, and harmful alcohol use. For the analysis, we fitted the WL-Gumbel cure model incorporating all covariates through the Poisson rate parameter:

$$\lambda_{ij} = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}_j), \quad i = 1, \ldots, 143, \quad j = 1, 2.$$

TABLE 1. Simulation results for bivariate cure models using WL-Gumbel and WL frailty under the scenario with censoring rates 50%, $n = 100$.

	WL-Gumbel					WL frailty				
Parameter	Mean	Bias	SD	MSE	AB	Mean	Bias	SD	MSE	AB
$r_1 = 2$	2.118	0.118	1.161	1.341	0.865	2.078	0.078	1.357	1.865	0.905
$\alpha_1 = 0.6$	0.611	0.011	0.075	0.005	0.058	0.651	0.051	0.194	0.037	0.098
$r_2 = 1.5$	1.668	0.168	1.076	1.157	0.766	1.695	0.195	1.383	1.912	0.772
$\alpha_2 = 0.7$	0.720	0.020	0.095	0.009	0.075	0.767	0.067	0.233	0.054	0.120
$\beta_{01} = 1.0$	1.475	0.475	0.544	0.295	0.556	1.710	0.710	1.052	1.106	0.794
$\beta_{11} = -1.0$	-1.020	-0.020	0.369	0.136	0.287	-1.104	-0.104	0.635	0.403	0.392
$\beta_{02} = 1.0$	1.495	0.495	0.583	0.339	0.593	1.635	0.635	0.951	0.904	0.746
$\beta_{12} = -1.0$	-1.050	-0.050	0.392	0.153	0.311	-1.089	-0.089	0.559	0.312	0.384
$\tau = 0$	0.011	0.011	0.018	0.000	0.011	-	-	-	-	-
$\gamma = 0.7$	0.665	-0.035	0.290	0.084	0.230	0.872	0.172	0.724	0.524	0.463
$r_1 = 2$	2.111	0.111	1.038	1.077	0.766	1.575	-0.425	1.220	1.488	1.039
$\alpha_1 = 0.6$	0.608	0.008	0.086	0.007	0.065	0.721	0.121	0.095	0.009	0.127
$r_2 = 1.5$	1.647	0.147	0.867	0.751	0.644	1.339	-0.161	1.065	1.134	0.858
$\alpha_2 = 0.7$	0.708	0.008	0.102	0.010	0.079	0.835	0.135	0.116	0.013	0.146
$\beta_{01} = 1.0$	1.425	0.425	0.597	0.356	0.546	2.761	1.761	0.844	0.712	1.762
$\beta_{11} = -1.0$	-0.990	0.010	0.393	0.154	0.307	-1.295	-0.295	0.538	0.289	0.485
$\beta_{02} = 1.0$	1.403	0.403	0.592	0.350	0.539	2.559	1.559	0.865	0.748	1.567
$\beta_{12} = -1.0$	-1.019	-0.019	0.406	0.164	0.319	-1.258	-0.258	0.594	0.352	0.507
$\tau = 0.3$	0.338	0.038	0.113	0.012	0.089	-	-	-	-	-
$\gamma = 0.7$	0.634	-0.066	0.492	0.242	0.390	1.697	0.997	0.419	0.175	0.998
$r_1 = 2$	2.072	0.072	0.851	0.724	0.625	1.969	-0.031	2.240	5.017	1.566
$\alpha_1 = 0.6$	0.590	-0.010	0.072	0.005	0.054	1.052	0.452	0.222	0.049	0.454
$r_2 = 1.5$	1.652	0.152	0.704	0.495	0.537	0.918	-0.582	1.192	1.420	1.079
$\alpha_2 = 0.7$	0.690	-0.010	0.086	0.007	0.065	1.262	0.562	0.253	0.064	0.563
$\beta_{01} = 1.0$	1.308	0.308	0.491	0.241	0.425	5.354	4.354	1.867	3.485	4.354
$\beta_{11} = -1.0$	-0.959	0.041	0.361	0.130	0.279	-1.634	-0.634	1.279	1.635	1.107
$\beta_{02} = 1.0$	1.276	0.276	0.481	0.231	0.405	5.797	4.797	1.813	3.286	4.797
$\beta_{12} = -1.0$	-0.957	0.043	0.371	0.137	0.292	-1.710	-0.710	1.317	1.734	1.166
$\tau = 0.6$	0.679	0.079	0.055	0.003	0.044	-	-	-	-	-
$\gamma = 0.7$	0.542	-0.158	0.442	0.193	0.368	3.256	2.556	0.789	0.622	2.556

In addition to the proposed WL-Gumbel cure model, we fitted two alternative bivariate cure models for comparison:

- **Gamma-Gumbel cure model**: Assuming the frailty component W in Equation (3.3) follows a gamma distribution with mean 1 and variance γ, with Laplace transform $\mathcal{L}_W(s) = (1 + \gamma s)^{-1/\gamma}$. Substituting this into Equation (3.3) yields the

TABLE 2. Simulation results for bivariate cure models using WL-Gumbel and WL frailty under the scenario with censoring rates $50\%, n = 200$.

	WL-Gumbel					WL frailty				
Parameter	Mean	Bias	SD	MSE	AB	Mean	Bias	SD	MSE	AB
$r_1 = 2$	1.987	-0.013	0.842	0.708	0.649	2.076	0.076	0.967	0.935	0.699
$\alpha_1 = 0.6$	0.603	0.003	0.055	0.003	0.043	0.612	0.012	0.066	0.004	0.048
$r_2 = 1.5$	1.585	0.085	0.745	0.555	0.545	1.658	0.158	0.842	0.708	0.615
$\alpha_2 = 0.7$	0.708	0.008	0.065	0.004	0.052	0.717	0.017	0.091	0.008	0.059
$\beta_{01} = 1.0$	1.421	0.421	0.414	0.171	0.466	1.508	0.508	0.488	0.238	0.544
$\beta_{11} = -1.0$	-1.005	-0.005	0.252	0.063	0.203	-1.039	-0.039	0.272	0.073	0.213
$\beta_{02} = 1.0$	1.393	0.393	0.452	0.204	0.453	1.456	0.456	0.555	0.308	0.513
$\beta_{12} = -1.0$	-1.012	-0.012	0.260	0.067	0.206	-1.030	-0.030	0.319	0.102	0.224
$\tau = 0$	0.007	0.007	0.011	0.000	0.007	-	-	-	-	-
$\gamma = 0.7$	0.683	-0.017	0.197	0.038	0.161	0.745	0.045	0.364	0.132	0.254
$r_1 = 2$	2.081	0.081	0.812	0.659	0.607	1.468	-0.532	0.964	0.929	0.932
$\alpha_1 = 0.6$	0.601	0.001	0.061	0.003	0.047	0.714	0.114	0.069	0.005	0.116
$r_2 = 1.5$	1.617	0.117	0.707	0.499	0.524	1.273	-0.227	0.901	0.811	0.755
$\alpha_2 = 0.7$	0.705	0.005	0.074	0.005	0.057	0.829	0.129	0.083	0.007	0.132
$\beta_{01} = 1.0$	1.365	0.365	0.438	0.192	0.438	2.737	1.737	0.734	0.538	1.734
$\beta_{11} = -1.0$	-0.999	0.001	0.275	0.075	0.219	-1.281	-0.281	0.381	0.145	0.381
$\beta_{02} = 1.0$	1.363	0.363	0.470	0.220	0.440	2.552	1.552	0.718	0.515	1.552
$\beta_{12} = -1.0$	-0.994	0.006	0.277	0.076	0.218	-1.250	-0.250	0.375	0.140	0.359
$\tau = 0.3$	0.337	0.037	0.077	0.005	0.060	-	-	-	-	-
$\gamma = 0.7$	0.636	-0.064	0.298	0.088	0.272	1.701	1.001	0.343	0.117	1.001
$r_1 = 2$	2.073	0.073	0.714	0.509	0.520	1.909	-0.091	2.041	4.165	1.517
$\alpha_1 = 0.6$	0.595	-0.005	0.050	0.002	0.040	1.030	0.430	0.159	0.025	0.430
$r_2 = 1.5$	1.645	0.145	0.558	0.311	0.423	0.745	-0.755	0.881	0.776	1.025
$\alpha_2 = 0.7$	0.696	-0.004	0.059	0.003	0.047	1.229	0.529	0.176	0.030	0.529
$\beta_{01} = 1.0$	1.251	0.251	0.355	0.112	0.325	5.244	4.244	1.560	2.443	4.244
$\beta_{11} = -1.0$	-0.984	0.016	0.245	0.060	0.195	-1.570	-0.570	0.947	0.896	0.891
$\beta_{02} = 1.0$	1.235	0.235	0.347	0.120	0.309	5.741	4.741	1.477	2.181	4.741
$\beta_{12} = -1.0$	-0.987	0.013	0.248	0.061	0.196	-1.633	-0.633	0.957	0.915	0.927
$\tau = 0.6$	0.624	0.024	0.022	0.001	0.262	-	-	-	-	-
$\gamma = 0.7$	0.554	-0.146	0.337	0.113	0.298	3.269	2.569	0.584	0.341	2.569

joint survival function:

$$S(t_1, t_2) = \left[1 + \gamma \left((\lambda_1 F_{01}(t_1))^{\theta+1} + (\lambda_2 F_{02}(t_2))^{\theta+1}\right)^{\frac{1}{\theta+1}}\right]^{-\frac{1}{\gamma}}.$$

When $\theta = 0$, this reduces to the standard gamma frailty model. As $\gamma \to \infty$, the model approaches the Gumbel copula without frailty.

TABLE 3. Simulation results for bivariate cure models using WL-Gumbel under the scenarios with different censoring rates, $n = 100$.

	30% Censoring					40% Censoring				
parameter	mean	Bias	SD	MSE	AB	mean	Bias	SD	MSE	AB
$r_1 = 2$	2.060	0.060	0.577	0.332	0.440	2.065	0.065	0.924	0.853	0.699
$\alpha_1 = 0.6$	0.60 1	0.001	0.068	0.004	0.053	0.602	0.002	0.082	0.006	0.063
$r_2 = 1.5$	1.571	0.071	0.481	0.231	0.369	1.614	0.114	0.811	0.657	0.632
$\alpha_2 = 0.7$	0.702	0.002	0.080	0.006	0.062	0.707	0.007	0.099	0.009	0.078
$\beta_{01} = 1.0$	1.327	0.327	0.393	0.154	0.392	1.405	0.405	0.566	0.320	0.504
$\beta_{11} = -1.0$	-1.002	0.002	0.337	0.113	0.263	-0.995	0.005	0.362	0.131	0.281
$\beta_{02} = 1.0$	1.307	0.307	0.394	0.155	0.381	1.405	0.405	0.574	0.329	0.524
$\beta_{12} = -1.0$	-0.968	0.032	0.326	0.106	0.255	-0.989	0.011	0.378	0.142	0.294
$\tau = 0.3$	0.347	0.047	0.093	0.008	0.074	0.345	0.045	0.108	0.012	0.087
$\gamma = 0.7$	0.624	-0.076	0.316	0.099	0.257	0.614	-0.086	0.414	0.171	0.336
$r_1 = 2$	2.067	0.067	0.484	0.234	0.374	2.085	0.085	0.739	0.546	0.554
$\alpha_1 = 0.6$	0.591	-0.009	0.055	0.003	0.043	0.593	-0.007	0.066	0.004	0.050
$r_2 = 1.5$	1.584	0.084	0.381	0.145	0.301	1.600	0.100	0.664	0.440	0.490
$\alpha_2 = 0.7$	0.692	-0.008	0.070	0.005	0.054	0.691	-0.009	0.076	0.006	0.059
$\beta_{01} = 1.0$	1.230	0.230	0.309	0.095	0.289	1.279	0.279	0.410	0.168	0.363
$\beta_{11} = -1.0$	-0.957	0.043	0.294	0.086	0.232	-0.949	0.051	0.333	0.110	0.265
$\beta_{02} = 1.0$	1.216	0.216	0.321	0.103	0.279	1.284	0.284	0.419	0.157	0.376
$\beta_{12} = -1.0$	-0.953	0.047	0.298	0.088	0.236	-0.949	0.051	0.332	0.110	0.264
$\tau = 0.6$	0.679	0.079	0.045	0.002	0.037	0.680	0.080	0.050	0.002	0.041
$\gamma = 0.7$	0.564	-0.136	0.302	0.091	0.270	0.555	-0.145	0.373	0.139	0.314

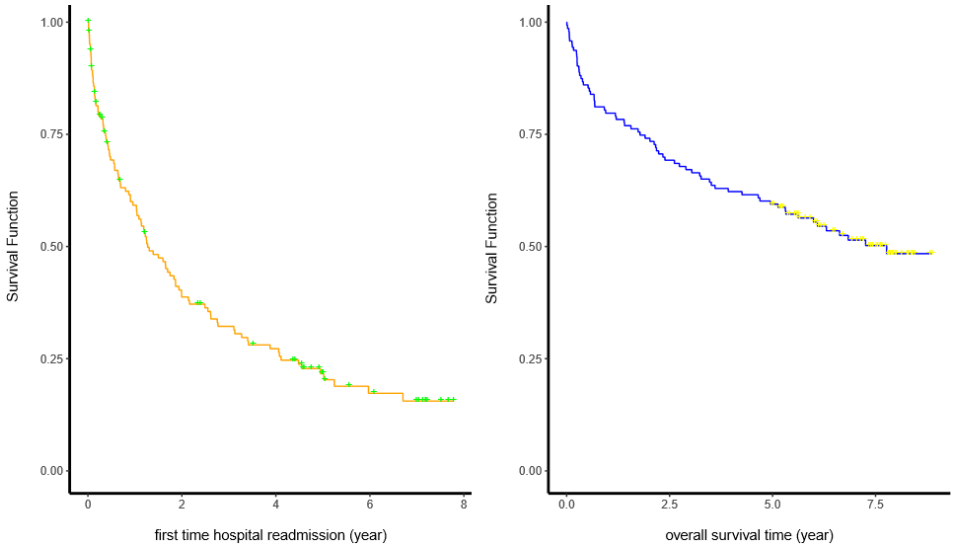


FIGURE 1. Plots of the survival functions estimated by the Kaplan-Meier method for first time hospital readmission and overall survival time.

TABLE 4. Simulation results for bivariate cure models using WL-Gumbel under the scenarios with different censoring rates , $n = 200$.

	30% censoring					40% censoring				
Parameter	Mean	Bias	SD	MSE	AB	Mean	Bias	SD	MSE	AB
$r_1 = 2$	2.014	0.014	0.401	0.160	0.313	2.044	0.044	0.654	0.428	0.506
$\alpha_1 = 0.6$	0.598	-0.002	0.045	0.002	0.035	0.597	-0.003	0.054	0.003	0.043
$r_2 = 1.5$	1.527	0.027	0.335	0.112	0.258	1.559	0.059	0.619	0.383	0.474
$\alpha_2 = 0.7$	0.699	-0.001	0.057	0.003	0.045	0.698	-0.002	0.067	0.004	0.052
$\beta_{01} = 1.0$	1.305	0.305	0.278	0.077	0.327	1.366	0.366	0.385	0.148	0.415
$\beta_{11} = -1.0$	-0.983	0.017	0.231	0.053	0.178	-0.990	0.010	0.265	0.070	0.211
$\beta_{02} = 1.0$	1.315	0.315	0.296	0.087	0.340	1.367	0.367	0.405	0.164	0.424
$\beta_{12} = -1.0$	-0.995	0.005	0.227	0.051	0.175	-0.989	0.011	0.269	0.072	0.214
$\tau = 0.3$	0.342	0.042	0.067	0.004	0.054	0.344	0.044	0.071	0.005	0.058
$\gamma = 0.7$	0.636	-0.064	0.232	0.053	0.194	0.627	-0.073	0.283	0.080	0.235
$r_1 = 2$	2.059	0.059	0.331	0.109	0.260	2.061	0.061	0.556	0.309	0.424
$\alpha_1 = 0.6$	0.586	-0.014	0.039	0.001	0.033	0.585	-0.015	0.044	0.002	0.037
$r_2 = 1.5$	1.572	0.072	0.268	0.071	0.216	1.617	0.117	0.457	0.208	0.361
$\alpha_2 = 0.7$	0.686	-0.014	0.048	0.002	0.040	0.686	-0.014	0.055	0.003	0.044
$\beta_{01} = 1.0$	1.210	0.210	0.219	0.047	0.232	1.231	0.231	0.297	0.088	0.277
$\beta_{11} = -1.0$	-0.940	0.060	0.202	0.040	0.168	-0.930	0.070	0.218	0.051	0.175
$\beta_{02} = 1.0$	1.206	0.206	0.218	0.047	0.228	1.214	0.214	0.283	0.080	0.265
$\beta_{12} = -1.0$	-0.944	0.056	0.202	0.040	0.167	-0.931	0.069	0.222	0.049	0.178
$\tau = 0.6$	0.681	0.081	0.031	0.001	0.029	0.682	0.082	0.033	0.001	0.030
$\gamma = 0.7$	0.565	-0.135	0.212	0.063	0.208	0.541	-0.159	0.267	0.071	0.252

- **Gamma-FGM cure model**: A similar approach using the Farlie-Gumbel-Morgenstern (FGM) copula with gamma frailty.

The conditional bivariate survival function induced by the FGM copula is given by:

$$S(t_1, t_2 \mid w) = \exp\left(-w[\lambda_1 F_{01}(t_1) + \lambda_2 F_{02}(t_2)]\right) \times \left[1 + \theta\left(1 - \exp(-w\lambda_1 F_{01}(t_1))\right)\left(1 - \exp(-w\lambda_2 F_{02}(t_2))\right)\right].$$

Assuming W follows a gamma distribution with mean 1 and variance γ, and integrating over the frailty distribution yields the unconditional bivariate survival function:

$$\begin{aligned} S(t_1, t_2) = (1+\theta)\left[1+\gamma(\lambda_1 F_{01}(t_1)+\lambda_2 F_{02}(t_2))\right]^{-1/\gamma} \\ -\theta\left[1+\gamma(\lambda_1 F_{01}(t_1)+2\lambda_2 F_{02}(t_2))\right]^{-1/\gamma} \\ -\theta\left[1+\gamma(2\lambda_1 F_{01}(t_1)+\lambda_2 F_{02}(t_2))\right]^{-1/\gamma} \\ +\theta\left[1+2\gamma(\lambda_1 F_{01}(t_1)+\lambda_2 F_{02}(t_2))\right]^{-1/\gamma}. \end{aligned}$$

When $\theta = 0$, the Gamma-FGM cure model reduces to the standard gamma frailty model. For further details on frailty-copula cure models, we refer interested readers to [25]. The parameter estimates, standard errors, and 95% confidence intervals for the fitted frailty-copula models (using Weighted Lindley and gamma frailties with Gumbel and FGM copulas) and standard frailty models are reported in Tables 5 and 6. To select the optimal model among all candidates, we employed the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with the comparison results presented in Table **??**. Parameters for the final model were initialized using estimates from simpler marginal models fitted separately for each outcome variable.

The real data analysis revealed distinct covariate patterns for the two event times. For the time to first hospital readmission (T_1), only age showed a statistically significant effect on the cure fraction (Est = 0.039, 95% CI: 0.018, 0.061). In contrast, for the time to tobacco-tar stain appearance (T_2), four covariates demonstrated significant effects: age (Est = 0.061, 95% CI: 0.031, 0.092), tobacco-related cancer (Est = 1.278, 95% CI: 0.021, 2.537), other cancers (Est = 1.476, 95% CI: 0.522, 2.431), and harmful alcohol use (Est = 0.630, 95% CI: 0.017, 1.244).

The positive coefficients for age in both outcomes indicate decreasing cure probabilities with advancing age. For T_2, the positive associations with tobacco-related cancer, other cancers, and harmful alcohol use suggest these factors reduce the probability of cure and increase stain persistence.

These findings have important clinical implications. Advanced age appears to be an independent risk factor for diminished treatment response. The strong association with

cancer history suggests shared pathological mechanisms between carcinogenesis and persistent tar stain formation. Additionally, harmful alcohol use emerges as a significant exacerbating factor for dermal lesion permanence.

According to both AIC and BIC values, the frailty-copula models provided a better fit than the standard frailty models. The WL-Gumbel model demonstrated the lowest AIC and BIC values among all fitted models, justifying the use of this multivariate approach to account for the inherent correlation between the two event times.

6. Discussion

This paper introduces a discrete frailty model for bivariate survival data that incorporates a cure fraction. The proposed model employs a weighted Lindley distribution to account for unobserved heterogeneity and a Gumbel copula to capture dependence between survival times. Parameters are estimated via maximum likelihood estimation.

Our analysis demonstrates that the Weighted Lindley-Gumbel (WL-Gumbel) model provides a superior fit and efficiency compared to alternative models. The flexibility of the Weighted Lindley distribution enables it to accommodate various hazard function shapes and more accurately capture event dependence than Gamma-based alternatives such as Gamma-Gumbel and Gamma-FGM. The WL-Gumbel model produced stable and reliable estimates, as evidenced by relatively low standard errors and reasonable 95% confidence intervals.

Despite these advantages, the model is computationally intensive, which may limit its application to very large datasets. Future research should focus on several promising directions: investigating alternative continuous frailty distributions, exploring different copula functions, examining other discrete frailty distributions, replacing the Weibull baseline with alternative parametric or nonparametric distributions, and extending the frailty-copula framework to other multivariate survival settings. Comparative studies across diverse datasets would further validate the model's performance relative to existing methods.

Acknowledgments

The authors would like to thank the reviewers for their thoughtful comments and efforts toward improving our manuscript.

TABLE 5. ML estimates (MLE), standard errors(SE), 95%CI for the fitted bivariate WL-Gumbel cure, gamma-Gumbel cure and gamma-FGM cure on the tobacco-stained fingers dataset.

Fitted model	WL-Gumbel		Gamma-Gumbel		Gamma-FGM	
parameter	Est(SE)	95% CI	Est(SE)	95% CI	Est(SE)	95% CI
r_1	0.041(0.015)	(0.010, 0.072)	0.042(0.015)	(0.012,0.071)	0.042(0.015)	(0.013, 0.071)
α_1	0.926(0.096)	(0.738, 1.114)	0.923(0.097)	(0.732, 1.114)	0.957(0.095)	(0.770, 1.144)
age	0.039(0.010)	(0.018, 0.061)	0.039(0.011)	(0.017, 0.061)	0.040(0.010)	(0.019, 0.060)
Vascular	-0.123(0.454)	(-1.014,0.76)	-0.127(0.453)	(-1.015, 0.761)	-0.115(0.479)	(-1.055, 0.823)
Ischemic heart	0.233(0.419)	(-0.590, 1.055)	0.243(0.417)	(-0.574, 1.061)	0.208(0.440)	(-0.654, 1.072)
stroke	0.127(0.449)	(-0.752,1.008)	0.125(0.448)	(-0.752, 1.003)	0.228(0.464)	(-0.682, 1.139)
arterial disease	-0.126(0.383)	(-0.878, 0.626)	-0.119(0.382)	(-0.869, 0.629)	-0.119(0.405)	(-0.914, 0.675)
Tobacco cancer	0.956(0.635)	(-0.289, 2.202)	0.932(0.632)	(-0.306, 2.171)	1.090(0.634)	(-0.153, 2.334)
cancer	0.149(0.492)	(-0.814, 1.114)	0.150(0.490)	(-0.811, 1.111)	0.087(0.514)	(-0.921, 1.096)
hypertension	0.048(0.283)	(-0.507, 0.605)	0.050(0.282)	(-0.503, 0.604)	0.085(0.298)	(-0.498, 0.669)
hypercholesterolemia	-0.443(0.293)	(-1.019, 0.132)	-0.442(0.293)	(-1.017, 0.133)	-0.481(0.308)	(-1.085, 0.122)
Harmful alcohol use	0.416(0.260)	(-0.094, 0.927)	0.414(0.260)	(-0.094, 0.924)	0.421(0.293)	(-0.109, 0.952)
Unfiltered cigarettes	0.322(0.281)	(-0.229, 0.874)	0.328(0.280)	(-0.222, 0.878)	0.273(0.001)	(-0.302, 0.849)
r_2	0.005(0.002)	(0.001,0.010)	0.004(0.001)	(0.002,0.006)	0.003(0.001)	(0.001,0.005)
α_2	0.914(0.112)	(0.694, 1.133)	0.909(0.111)	(0.690, 1.128)	0.899(0.103)	(0.696,1.102)
age	0.061(0.015)	(0.031, 0.092)	0.062(0.015)	(0.031, 0.091)	0.051(0.006)	(0.038, 0.064)
Vascular	0.737(0.526)	(-0.294, 1.769)	0.735(0.526)	(-0.296, 1.767)	0.632(0.539)	(-0.425, 1.690)
Ischemic heart	-0.534(0.459)	(-1.435, 0.366)	-0.523(0.458)	(-1.421, 0.374)	-0.428(0.463)	(-1.338, 0.480)
stroke	0.151(0.486)	(-0.802, 1.104)	0.148(0.487)	(-0.805, 1.102)	0.339(0.492)	(-0.625, 1.304)
arterial disease	-0.117(0.454)	(-1.008, 0.773)	-0.109(0.453)	(-0.997, 0.778)	-0.007(0.448)	(-0.887, 0.872)
Tobacco cancer	1.278(0.641)	(0.021, 2.537)	1.252(0.638)	(0.001, 2.503)	1.368(0.628)	(0.137, 2.600)
cancer	1.476(0.487)	(0.522, 2.431)	1.472(0.486)	(0.519, 2.425)	1.375(0.502)	(0.391, 2.359)
hypertension	-0.154(0.329)	(-0.799, 0.490)	-0.149(0.328)	(-0.792, 0.494)	-0.131(0.327)	(-0.774, 0.510)
hypercholesterolemia	-0.526(0.390)	(-1.292, 0.239)	-0.527(0.390)	(-1.292, 0.238)	-0.606(0.391)	(-1.374, 0.160)
Harmful alcohol use	0.630(0.313)	(0.017, 1.244)	0.628(0.312)	(0.016, 1.240)	0.494(0.291)	(-0.077, 1.066)
Unfiltered cigarettes	-0.325(0.335)	(-0.982, 0.331)	-0.317(0.335)	(-0.974, 0.339)	-0.413(0.33)	(-1.066, 0.240)
γ	0.406(0.133)	(0.145, 0.666)	0.407(0.139)	(0.134, 0.679)	0.526(0.156)	(0.220, 0.832)
θ	0.102(0.044)	(0.016, 0.188)	0.103(0.044)	(0.017, 0.189)	0.317(0.313)	(-0.296, 0.930)

TABLE 6. ML estimates (MLE), standard errors(SE), 95% CI for the fitted WL frailty, gamma, Gumbel and FGM copula on tobacco-stained fingers data

Fitted model	WL frailty		Gamma frailty		Gumbel copula		FGM copula	
parameter	Est(SE)	95% CI	Est(SE)	95% CI	Est(SE)	95% CI	Est(SE)	95% CI
r_1	0.274(0.056)	(0.165, 0.384)	0.504(0.071)	(0.364, 0.643)	0.075(0.031)	(0.014,0.13)	0.072(0.036)	(0.002, 0.143)
α_1	1.015(0.089)	(0.838, 1.191)	0.936(0.076)	(0.787, 1.085)	0.591(0.067)	(0.459,0.724)	0.755(0.062)	(0.632, 0.877)
age	0.016(0.005)	(0.006, 0.026)	0.002(0.004)	(-0.006, 0.011)	0.037(0.006)	(0.025, 0.050)	0.031(0.007)	(0.016, 0.046)
Vascular	-0.351(0.474)	(-1.281, 0.579)	0.275(0.479)	(-0.664, 1.21)	-0.309(0.37)	(-1.043, 0.423)	0.243(0.398)	(-0.537, 1.024)
Ischemic heart	0.503(0.427)	(-0.334, 1.341)	-0.228(0.431)	(-1.074, 0.616)	0.546(0.335)	(-0.110, 1.203)	-0.458(0.353)	(-1.151, 0.235)
stroke	0.204(0.454)	(-0.686, 1.094)	-0.341(0.485)	(-1.291, 0.609)	-0.682(0.421)	(-1.507, 0.143)	0.059(0.374)	(-0.675, 0.793)
arterial disease	0.038(0.391)	(-0.728, 0.805)	-0.079(0.398)	(-0.860, 0.700)	0.015(0.294)	(-0.561, 0.593)	-0.046(0.343)	(-0.720, 0.628)
Tobacco cancer	0.437(0.612)	(-0.761,1.637)	-0.414(0.557)	(-1.507, 0.677)	-0.145(0.502)	(-1.130, 0.839)	0.460(0.497)	(-0.515, 1.436)
cancer	0.063(0.481)	(-0.879, 1.007)	1.617(0.500)	(0.636, 2.597)	0.161(0.440)	(-0.701, 1.024)	0.111(0.467)	(-0.804, 1.027)
hypertension	0.418(0.296)	(-0.163,1.001)	0.256(0.299)	(-0.331, 0.843)	-0.037(0.220)	(-0.471, 0.395)	0.190(0.249)	(-0.298, 0.679)
hypercholesterolemia	-0.577(0.290)	(-1.146, -0.008)	0.100(0.306)	(-0.500, 0.701)	-0.180(0.234)	(-0.640, 0.279)	-0.542(0.275)	(-1.081, -0.003)
Harmful alcohol use	-0.041(0.248)	(-0.527, 0.445)	0.387(0.251)	(-0.106, 0.881)	-0.191(0.202)	(-0.588, 0.204)	0.266(0.224)	(-0.173, 0.705)
Unfiltered cigarettes	0.361(0.285)	(-0.197, 0.920)	-0.108(0.002)	(-0.697, 0.479)	0.269(0.227)	(-0.176, 0.714)	-0.050(0.246)	(-0.533, 0.431)
r_2	0.189(0.044)	(0.102, 0.276)	0.005(0.089)	(0.001, 0.009)	0.013(0.006)	(0.001,0.025)	0.005(0.002)	(0.001,0.010)
α_2	1.028(0.121)	(0.789, 1.267)	0.850(0.300)	(0.675, 1.025)	0.435(0.049)	(0.338, 0.533)	0.485(0.055)	(0.377, 0.593)
age	-0.021(0.006)	(-0.034, -0.008)	0.044(0.007)	(0.030, 0.059)	0.047(0.007)	(0.033, 0.061)	0.064(0.006)	(0.051, 0.076)
Vascular	0.727(0.542)	(-0.334, 1.790)	0.719(0.572)	(-0.401, 1.841)	-0.058(0.414)	(-0.870, 0.752)	-0.304(0.477)	(-1.240, 0.631)
Ischemic heart	0.167(0.452)	(-0.718, 1.053)	0.622(0.503)	(-0.364, 1.609)	0.108(0.365)	(-0.608, 0.825)	-0.257(0.439)	(-1.118, 0.604)
stroke	0.756(0.492)	(-0.208, 1.721)	-0.779(0.577)	(-1.911, 0.351)	-0.508(0.442)	(-1.376, 0.359)	-0.641(0.519)	(-1.658, 0.375)
arterial disease	0.091(0.408)	(-0.709, 0.892)	-0.401(0.435)	(-1.253, 0.451)	-0.157(0.335)	(-0.816, 0.500)	-0.222(0.417)	(-1.041, 0.596)
Tobacco cancer	2.111(0.668)	(0.801, 3.421)	1.435(0.796)	(-0.125, 2.997)	0.122(0.604)	(-1.063, 1.307)	0.261(0.754)	(-1.217, 1.741)
cancer	0.692(0.546)	(-0.378, 1.763)	-0.052(0.712)	(-1.448, 1.343)	-0.053(0.534)	(-1.101, 0.994)	0.067(0.689)	(-1.283, 1.417)
hypertension	0.554(0.329)	(-0.091, 1.200)	0.432(0.340)	(-0.235, 1.099)	0.196(0.256)	(-0.305, 0.698)	-0.584(0.283)	(-1.140, -0.027)
hypercholesterolemia	-0.210(0.345)	(-0.888, 0.466)	-0.788(0.368)	(-1.509, -0.066)	0.129(0.261)	(-0.384, 0.642)	0.705(0.296)	(0.124, 1.286)
Harmful alcohol use	0.398(0.281)	(-0.153, 0.950)	0.079(0.300)	(-0.509, 0.667)	0.082(0.232)	(-0.373, 0.537)	0.126(0.258)	(-0.381, 0.633)
Unfiltered cigarettes	-0.166(0.319)	(-0.793, 0.460)	-0.185(0.342)	(-0.857, 0.485)	-0.0486(0.249)	(-0.538, 0.441)	-0.150(0.287)	(-0.713, 0.412)
γ	0.444(0.191)	(0.068, 0.821)	0.549(0.147)	(0.260, 0.837)	-	-	-	-
θ	-	-	-	-	1.152(0.227)	(0.707, 1.596)	-0.118(0.399)	(-0.901, 0.664)

TABLE 7. Comparison of models fitted to the tobacco-stained fingers dataset

Fitted model	AIC	BIC	- log-likelihood
WL-Gumbel	749.974	832.933	346.987
Gamma-Gumbel	750.109	833.068	347.054
Gamma-FGM	752.724	835.684	348.362
WL frailty	812.239	892.236	397.119
Gamma frailty	853.947	933.944	399.973
Gumbel copula	857.064	937.061	401.532
FGM copula	878.876	958.873	412.438

References

[1] Balakrishnan, N. and Peng, Y., 2006. Generalized gamma frailty model. *Statistics in medicine*, 25(16), pp.2797–2816. doi:10.1002/sim.2375.

[2] Cancho, V.G., Suzuki, A.K., Barriga, G.D.C. and Santo, A.P.J.E., 2022. A multivariate survival model induced by discrete frailty. *Communications in Statistics-Simulation and Computation*, 51(11), pp.6572–6590. doi:10.1080/03610918.2020.1806323.

[3] Clayton, D.G., 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1), pp.141–151. doi:10.1007/s10985-010-9151-3.

[4] Caroni, C., Crowder, M. and Kimber, A., 2010. Proportional hazards models with discrete frailty. *Lifetime data analysis*, 16, pp.374–384. doi:10.1093/biomet/65.1.141.

[5] de Souza, D., Cancho, V.G., Rodrigues, J. and Balakrishnan, N., 2017. Bayesian cure rate models induced by frailty in survival analysis. *Statistical methods in medical research*, 26(5), pp.2011–2028. doi:10.1177/0962280217708671.

[6] Emura, T., Matsui, S. and Rondeau, V., 2019. *Survival analysis with correlated endpoints: Joint Frailty-Copula models*. Springer. doi:10.1007/978-981-13-3516-7.

[7] Emura, T., Nakatochi, M., Murotani, K. and Rondeau, V., 2017. A joint frailty-copula model between tumour progression and death for meta-analysis. *Statistical methods in medical research*, 26(6), pp.2649–2666. doi:10.1177/0962280215604510.

[8] Emura, T., Sofeu, C.L. and Rondeau, V., 2021. Conditional copula models for correlated survival endpoints: Individual patient data meta-analysis of randomized controlled trials. *Statistical Methods in Medical Research*, 30(12), pp.2634–2650. doi:10.1177/09622802211046390.

[9] Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K. and Husain, H.A., 2011. A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in simulation*, 81(6), pp.1190–1201. doi:10.1016/j.matcom.2010.11.005.

[10] Ha, I.D., Jeong, J.H. and Lee, Y., 2018. *Statistical modelling of survival data with random effects: h-likelihood approach*. Springer Publishing Company, Incorporated. doi:10.1007/978-981-10-6557-6.

[11] Hougaard, P., 1995. Frailty models for survival data. *Lifetime data analysis*, 1, pp.255–273. doi:10.1007/BF00985760.

[12] Hougaard, P., 1986. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73(2), pp.387–396. doi:10.1093/biomet/73.2.387.

[13] John, G., Louis, C., Berner, A. and Genne, D., 2015. Tobacco stained fingers and its association with death and hospital admission: A retrospective cohort study. *PloS one*, 10(9), e0138211. doi:10.1371/journal.pone.0138211.

[14] Kwon, S., Ha, I.D., Shih, J.H. and Emura, T., 2022. Flexible parametric copula modeling approaches for clustered survival data. *Pharmaceutical Statistics*, 21(1), pp.69–88. doi:10.1002/pst.2153.

[15] Li, Z., Chinchilli, V.M. and Wang, M., 2019. A Bayesian joint model of recurrent events and a terminal event. *Biometrical Journal*, 61(1), pp.187–202. doi:10.1002/bimj.201700326.

[16] Louzada, F., Suzuki, A.K. and Cancho, V.G., 2013. The FGM long-term bivariate survival copula model: modeling, Bayesian estimation, and case influence diagnostics. *Communications in Statistics-Theory and Methods*, 42(4), pp.673–691. doi:10.1080/03610926.2012.725147.

[17] Mazucheli, J., Coelho-Barros, E.A. and Louzada, F., 2016. On the hypothesis testing for the weighted lindley distribution. *Chilean Journal of Statistics*, 7(2), pp.17–27.

[18] Mota, A., Ramos, P.L., Ferreira, P., Tomazella, V. and Louzada, F., 2021. A reparameterized weighted Lindley distribution: properties, estimation and applications. *Revista Colombiana de Estadistica*, 44(1), pp.65–90. doi:10.15446/rce.v44n1.86566.

[19] Nair, N.U., Sankaran, P.G. and John, P., 2018. Modelling bivariate lifetime data using copula. *Metron*, 76, pp.133–153. doi:10.1007/s40300-018-0135-5.

[20] Nelsen, R.B., 2006. *An introduction to copulas*. Springer. doi:10.1007/0-387-28678-0.

[21] Peres, M.V.O., Achcar, J.A. and Martinez, E.Z., 2018. Bivariate modified Weibull distribution derived from Farlie-Gumbel-Morgenstern copula: a simulation study. *Electronic Journal of Applied Statistical Analysis*, 11(2), pp.463–488. doi:10.1285/i20705948v11n2p463.

[22] Prenen, L., Braekers, R. and Duchateau, L., 2018. Investigating the correlation structure of quadrivariate udder infection times through hierarchical Archimedean copulas. *Lifetime Data Analysis*, 24, pp.719–742. doi:10.1007/s10985-017-9411-6.

[23] Romeo, J.S., Tanaka, N.I. and Pedroso-de-Lima, A.C., 2006. Bivariate survival modeling: a Bayesian approach based on copulas. *Lifetime Data Analysis*, 12, pp.205–222. doi:10.1007/s10985-006-9001-5.

[24] Romeo, J.S., Meyer, R. and Gallardo, D.I., 2018. Bayesian bivariate survival analysis using the power variance function copula. *Lifetime data analysis*, 24, pp.355–383. doi:10.1007/s10985-017-9396-1.

[25] Rouzbahani, M., Akhoond, M.R. and Chinipardaz, R., 2024. A new bivariate survival model with a cured fraction: a mixed Poisson frailty–copula approach. *Japanese Journal of Statistics and Data Science*, pp.1–25. doi:10.1007/s42081-023-00240-z.

[26] Vaupel, J.W., Manton, K.G. and Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3), pp.439–454. doi:10.2307/2061224.

[27] Wang, Y.C., Emura, T., Fan, T.H., Lo, S.M.S. and Wilke, R.A., 2020. Likelihood-based inference for a frailty-copula model based on competing risks failure time data. *Quality and Reliability Engineering International*, 36(5), pp.1622–1638. doi:10.1002/qre.2650.

[28] Wienke, A., 2010. *Frailty models in survival analysis*. Chapman and Hall/CRC. doi:10.1201/9781420073911.

Marziye Rouzbahani
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran.
Email: M-roozbahani@stu.scu.ac.ir

Mohammad Reza Akhoond
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran.
Email: Mr.akhoond@scu.ac.ir

Rahim Chinipardaz
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran.
Email: Chinipardaz_r@scu.ac.ir



©2024 Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-NonComertial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).