Environ Ecol Stat (2017) 24:151–173
DOI 10.1007/s10651-017-0364-4

A copula-based inference to piecewise exponential models under dependent censoring, with application to time to metamorphosis of salamander larvae

Takeshi Emura[1] · **Hirofumi Michimae**[2]

Received: 18 January 2016 / Revised: 29 December 2016 / Published online: 10 February 2017

Abstract In ecology and evolutionary biology, controlled animal experiments are often conducted to measure time to metamorphosis which is possibly censored by the competing risk of death and the follow-up end. This paper considers the problem of estimating the survival function of time-to-event when it is subject to dependent censoring. When the censorship is due to competing risks, the traditional assumption of independent censorship may not be satisfied, and hence, the usual application of the Kaplan–Meier estimator yields a biased estimation for the survival function of the event time. This paper follows an assumed copula approach (Zheng and Klein in Biometrika 82(1):127–138, 1995) to adjust for dependence between the event time of interest and the competing event time. While the literature on an assumed copula approach has mostly focused on semiparametric settings, we alternatively consider a parametric approach with piecewise exponential models for fitting the survival function. We develop maximum likelihood estimation under the piecewise exponential models with an assumed copula. A goodness-of-fit procedure is also developed, which touches upon the identifiability issue of the copula. We conduct simulations to examine the performance of the proposed method and compare it with an existing semiparametric method. The method is applied to real data analysis on time to metamorphosis

Handling Editor: Pierre Dutilleul.

Electronic supplementary material The online version of this article (doi:10.1007/s10651-017-0364-4) contains supplementary material, which is available to authorized users.

✉ Takeshi Emura
takeshiemura@gmail.com

[1] Graduate Institute of Statistics, National Central University, Taoyuan, Taiwan

[2] Department of Clinical Medicine (Biostatistics), School of Pharmacy, Kitasato University, Tokyo, Japan

for salamander larvae living in Hokkaido, Japan (Michimae et al. in Evol Ecol Res 16:617–629, 2014).

Keywords Bivariate survival analysis · Copula-graphic estimator · Goodness-of-fit test · Kendall's tau · Survival analysis

1 Introduction

Competing risks data appear in many scientific fields, where subjects experience multiple event types and the interest lies in the time up to events (Crowder 2001, 2012; Klein and Moeschberger 2003). By the definition of competing risks, these event types are mutually exclusive in that each subject exhibits only one of the multiple events. Usually, statistical analyses for competing risks data are performed without assuming independence between event times. Competing risks data are popular, especially in biological research involving the experimentally designed follow-up for human and animal subjects (Andersen et al. 2002; Chap 8 of Kalbfleisch and Prentice 2002).

When analyzing animal subjects' time to event data arising from ecology and evolutionary biology, competing risks arise when subjects under study experience only one of multiple events. For instance, in a controlled follow-up experiment of salamander larvae, time to metamorphosis is unavailable if death comes earlier than metamorphosis (Michimae et al. 2014). Conversely, death occurring after metamorphosis no longer has the intended meaning as "death of larvae". Due to the possible dependence between metamorphosis and death, the Kaplan–Meier estimator (Kaplan and Meier 1958) for the survival function may be biased. For analyzing such data, the standard recommendation is the use of the sub-distribution function (Crowder 2001, 2012; Klein and Moeschberger 2003; Bakoyannis and Touloumi 2012) also known as the cumulative incidence function. Hereafter, we focus on the case that an event time (denoted by T, such as time to metamorphosis) is subject to a single competing risk (the competing event time denoted by U, such as time to death).

In a situation where a pair of event times (T, U) is subject to competing risks, the marginal survival function $S_T(t) = \Pr(T > t)$ is unidentifiable from data (Tsiatis 1975). The identifiability problem can be remedied by imposing some assumptions on the joint distribution of (T, U). The traditional assumption is the independence between T and U. With this strong assumption, the marginal survival function $S_T(t)$ is estimated by the Kaplan–Meier estimator treating U as an independent censoring variable. However, the independence assumption is rarely true in many biological settings.

An assumed copula model on (T, U) makes $S_T(t)$ identifiable (Zheng and Klein 1995), where a copula only specifies the form of dependency but does not specify the forms of the marginal survival functions. With an assumed copula, Zheng and Klein (1995) proposed a semiparametric estimator of $S_T(t)$, called the copula-graphic (CG) estimator that is an extension of the Kaplan–Meier estimator. The asymptotic properties of the CG estimator were well-understood with aid of a martingale technique (Rivest and Wells 2001). De Uña-Álvarez and Veraverbeke (2013, 2017)

generalized the CG estimator to allow for additional independent right-censoring and left-truncation.

Adopting the idea of an assumed copula, semiparametric estimators for the marginal survival and hazard functions are derived in the presence of covariates (Braekers and Veraverbeke 2005; Chen 2010; Emura et al. 2015; Emura and Chen 2016). Theoretical and numerical properties of these estimators are studied within the same references.

An assumed copula approach has also been adopted to perform parametric analyses. However, compared to the semiparametric approaches, the literature is scarcer for the parametric approaches. A maximum likelihood estimator (MLE) under the Weibull marginal models for T and U is considered by Escarela and Carriere (2003). The MLE under a more general log-location-scale model is studied by Hsu et al. (2016) in the reliability analysis of a series system.

Here in our paper, we aim to contribute to the parametric approaches for the copula-based competing risks data analysis by developing maximum likelihood inference under a new parametric model. With an assumed copula model for dependence as in Zheng and Klein (1995), we adopt piecewise exponential models for the two marginal distributions as in Staplin et al. (2015). However, our proposed copula model is different from the model of Staplin et al. (2015) who specify the dependence through the conditional density for $(U|T = t)$. The likelihood function under the model of Staplin et al. (2015) involves some numerical integrations of the joint density of (T, U) while our likelihood function has an explicit form (Sect. 4.2). Another important advantage of our approach over Staplin et al. (2015) is that a copula parameter (say, α) is transformed to Kendall's tau, which is convenient when performing the sensitivity analysis (Sect. 4.3).

This paper is organized as follows. Section 2 describes the background on competing risks data analysis. Section 3 reviews piecewise exponential models. Section 4 describes our proposed inference procedures. Section 5 conducts simulations and Sect. 6 performs a real data analysis on salamander larvae. Section 7 concludes.

2 Competing risks framework

We consider a situation where a pair of event times (T, U) is subject to competing risks. Let T be the event time of interest (e.g., time to metamorphosis) and U be the competing event time (e.g., time-to-death). The independence between T and U is not assumed. Let M be a fixed censoring time point (Type I censoring). The event time T is observed if $T \leq U$ and $T < M$. Conversely, the event time U is observed if $U < T$ and $U < M$. What we actually observe is the first-occurring event time $t = \min(T, U, M)$, the event indicator $\delta = \mathbf{I}\{T \leq U\}$, and the censoring indicator $\rho = \mathbf{I}\{\min(T, U) < M\}$, where $\mathbf{I}\{\cdot\}$ is the indicator function. Note that δ is observed only when $\rho = 1$.

In the standard approach to competing risks analysis, the *sub-distribution function* (also known as *cumulative incidence function*), $H(t) = \Pr(T \leq t, T \leq U)$, is often the target for estimation (Crowder 2001, 2012; Klein and Moeschberger 2003; Bakoyannis and Touloumi 2012). This is the proportion of events occurring before time t and can be estimated nonparametrically under the above data structure (Klein and Moeschberger 2003).

In an alternative approach to competing risks analysis, the marginal survival function $S_T(t) = \Pr(T > t)$ is assessed. If T and U were independent, one could use the Kaplan–Meier estimator for $S_T(t)$. More generally, if the copula of (T, U) is known, a consistent estimator of $S_T(t)$ is obtained, e.g., by the generalized copula graphic (GCG) estimator (De Uña-Álvarez and Veraverbeke 2013). When calculating the estimates of $S_T(t)$, the form of the copula is assumed to be known. This is because statistical inference of the copula is inherently difficult under the competing risks setting (Sect. 4.3). For this reason, estimates from competing risks data are typically supplied with sensitivity analysis (Rivest and Wells 2001; Chen 2010; De Uña-Álvarez and Veraverbeke 2013, 2017; Staplin et al. 2015).

Our motivating example is a biological study on time to metamorphosis for salamander larvae living in Hokkaido, Japan (Michimae et al. 2014). A subset of the original data is given in Online Resource A. The event time of interest is days from placing the salamander larvae in the cage to the completion of metamorphosis. If larval death occurs before metamorphosis, the event time is censored. To account for the effect of dependent censoring, Michimae et al. (2014) assessed time to metamorphosis by using a sub-distribution function (or cumulative incidence function). Alternatively, we aim to assess the time to metamorphosis by using a survival function. This is mainly because survival functions are common tools for biologists dealing with animal subjects; see case studies of Kuparinen et al. (2008) for the common frog and Fieberg and DelGiudice (2011) for the female deer.

3 Piecewise exponential model

This section reviews the piecewise exponential model and introduces relevant notations. A good introduction to the piecewise exponential model is found in the book of Lawless (2003).

We define the piecewise exponential models by following Lawless (2003). Let $0 = a_0 < a_1 < \cdots < a_m = M$ be a knot sequence, where m is the number of knots and M is the follow-up end (Type I censoring time). Assume that the hazard function for T in an interval $(a_{j-1}, a_j]$ is a constant,

$$h_T(t; \boldsymbol{\theta}) = e^{\theta_j}; \quad a_{j-1} < t \leq a_j, \quad j = 1, \ldots, m,$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ are unknown parameters. The survival function is

$$S_T(t; \boldsymbol{\theta}) = \exp\left\{ -e^{\theta_j}(t - a_{j-1}) - \sum_{k=1}^{j-1} e^{\theta_k}(a_k - a_{k-1}) \right\}, \qquad t \in (a_{j-1}, a_j], \quad (1)$$

where $\sum_{k=1}^{0}(\cdot) \equiv 0$. The survival function is also derived by using the lack of memory among the exponentially distributed events across intervals,

$$S_T(t; \boldsymbol{\theta}) = \Pr(T > t | T > a_{j-1}) \times \Pr(T > a_{j-1} | T > a_{j-2}) \cdots \times \Pr(T > a_1).$$

The probability density function is

$$f_T(t;\boldsymbol{\theta}) = e^{\theta_j} \exp\left\{-e^{\theta_j}(t-a_{j-1}) - \sum_{k=1}^{j-1} e^{\theta_k}(a_k - a_{k-1})\right\}, \qquad t \in (a_{j-1}, a_j].$$

The usefulness of the piecewise exponential model is due to a mathematically tractable framework for statistical inference and to a flexibility to approximate accurately any model by increasing the number of knots (Friedman 1982). The model is also straightforwardly extended to include a vector of covariates (Friedman 1982; Lawless 2003).

In practice, the true knot sequence is unknown and needs to be chosen by users. In general, the chosen knot sequence may not coincide with the true one. If the true hazard function is smooth, the true knot sequence is not unique or does not exist. Even for these circumstances, the piecewise exponential model can still reasonably approximate the underlying model by letting $m \to \infty$ and $\max_j(a_j - a_{j-1}) \to 0$. This is because the maximum likelihood estimate for each interval $(a_{j-1}, a_j]$ is consistent for the average hazard in the interval (Friedman 1982). This feature allows one to apply the piecewise exponential as a realistic approximation to any application with less concern for the knot specification.

Staplin et al. (2015) imposed the piecewise exponential model on the competing event time (dependent censoring time) distribution of U in addition to the event time distribution of T. To save notations, we use the same knot sequence $0 = a_0 < a_1 < \cdots < a_m = M$ as the event time. The knot sequences of the event time and dependent censoring time can be different in practice. Then, the hazard for U in an interval $(a_{j-1}, a_j]$ is a constant, and hence

$$h_U(u;\boldsymbol{\gamma}) = e^{\gamma_j}; \quad a_{j-1} < u \le a_j, \quad j = 1, \ldots, m,$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)$ are unknown parameters. The corresponding survival function is

$$S_U(u;\boldsymbol{\gamma}) = \exp\left\{-e^{\gamma_j}(u-a_{j-1}) - \sum_{k=1}^{j-1} e^{\gamma_k}(a_k - a_{k-1})\right\}, \quad u \in (a_{j-1}, a_j]. \quad (2)$$

The probability density function is defined as $f_U(u;\boldsymbol{\gamma}) = -dS_U(u;\boldsymbol{\gamma})/du$.

Example 1 Consider a piecewise exponential model on three intervals (0, 1], (1, 2], and (2, 3], which are set by $a_1 = 1$, $a_2 = 2$, and $a_3 = M = 3$ with $m = 3$. Then, the survival function is

$$S_T(t;\boldsymbol{\theta}) = \begin{cases} \exp\{-e^{\theta_1} t\} & \text{if } t \in (0, a_1], \\ \exp\{-e^{\theta_2}(t-a_1) - e^{\theta_1} a_1\} & \text{if } t \in (a_1, a_2], \\ \exp\{-e^{\theta_3}(t-a_2) - e^{\theta_2}(a_2-a_1) - e^{\theta_1} a_1\} & \text{if } t \in (a_2, a_3]. \end{cases}$$

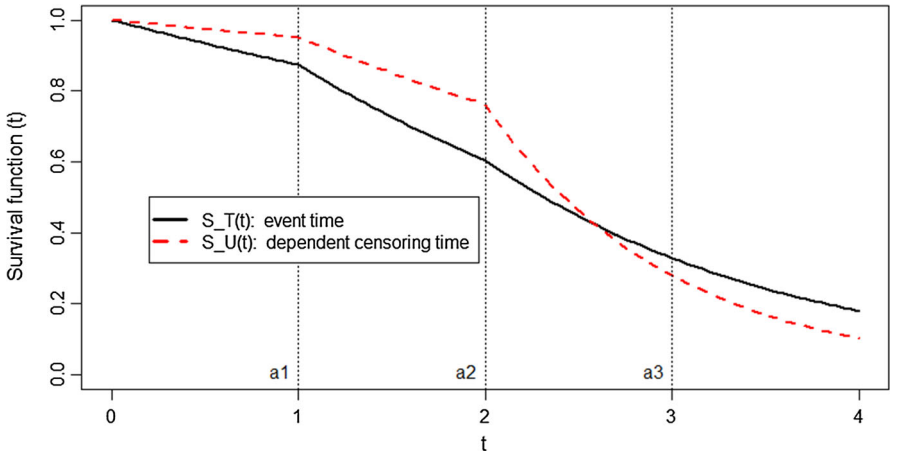

Fig. 1 Survival functions for piecewise exponential models under a knot sequence $a_1 = 1$, $a_2 = 2$, and $a_3 = 3$. The *solid* (*black*) *line* depicts $S_T(t; \theta_1, \theta_2, \theta_3)$ under three hazards $e^{\theta_1} = 0.135$, $e^{\theta_2} = 0.368$, and $e^{\theta_3} = 0.601$ ($\theta_1 = -2$, $\theta_2 = -1$, and $\theta_3 = -0.5$). The *dashed* (*red*) *line* depicts $S_U(t; \gamma_1, \gamma_2, \gamma_3)$ under three hazards $e^{\gamma_1} = 0.05$, $e^{\gamma_2} = 0.223$, and $e^{\gamma_3} = 1$ ($\gamma_1 = -3$, $\gamma_2 = -1.5$, and $\gamma_3 = 0$) (Color figure online)

We set three hazards $e^{\theta_1} = 0.135$, $e^{\theta_2} = 0.368$, and $e^{\theta_3} = 0.601$ ($\theta_1 = -2$, $\theta_2 = -1$, and $\theta_3 = -0.5$) in the intervals. In this model, we set $M = 3$ as the follow-up end (Type I censoring time). For dependent censoring time, we consider three hazards $e^{\gamma_1} = 0.05$, $e^{\gamma_2} = 0.223$, and $e^{\gamma_3} = 1$ ($\gamma_1 = -3$, $\gamma_2 = -1.5$, and $\gamma_3 = 0$) in the same intervals. Figure 1 shows the two survival functions for event time and dependent censoring time. We see that the survival functions are mixtures of three exponential survival functions on the three intervals $(0, a_1]$, $(a_1, a_2]$, and $(a_2, a_3]$.

4 Proposed methods

4.1 Copula-based survival model

A copula is a bivariate distribution function for a pair of uniformly distributed random variables on $[0, 1]$ (Nelsen 2006). Let $C_\alpha : [0, 1]^2 \mapsto [0, 1]$ be a family of one-parameter copulas, where α is a dependence parameter.

Some examples of copulas include the Clayton copula (1978)

$$C_\alpha(v, w) = (v^{-\alpha} + w^{-\alpha} - 1)^{-1/\alpha}, \quad \alpha \geq 0,$$

and the Joe copula (1993)

$$C_\alpha(v, w) = 1 - \{(1 - v)^\alpha + (1 - w)^\alpha - (1 - v)^\alpha (1 - w)^\alpha\}^{1/\alpha}, \quad \alpha \geq 1.$$

The Clayton copula has a lower tail dependence while the Joe copula has an upper tail dependence. Hence these two copulas capture quite different dependence structures and supplement each other in statistical modeling. The independence copula

$C(v, w) = vw$ is obtained as the limit $\alpha \to 0$ under the Clayton copula and $\alpha \to 1$ under the Joe copula. The degree of dependence increases as α departs from these limits.

An important property of copulas is that the dependence parameter α has the one-to-one correspondence with Kendall's tau

$$\tau(\alpha) = 4 \int_0^1 \int_0^1 C_\alpha(v, w) C_\alpha(dv, dw) - 1.$$

The Clayton copula has a form $\tau(\alpha) = \alpha/(\alpha + 2)$ while the Joe copula has an integral form

$$\tau(\alpha) = 1 - 4 \int_0^\infty \frac{t}{\alpha^2} \{1 - \exp(-t)\}^{2/\alpha - 2} \exp(-2t) dt.$$

In both cases, $\tau(\alpha)$ is a monotone increasing function of α with $\tau(\infty) = 1$.

We consider a model for the joint survival function of T and U given as

$$\Pr(T > t, U > u) = C_\alpha\{S_T(t; \boldsymbol{\theta}), S_U(u; \boldsymbol{\gamma})\}, \tag{3}$$

where $S_T(t; \boldsymbol{\theta})$ and $S_U(u; \boldsymbol{\gamma})$ follow the piecewise exponential models (1) and (2), respectively. Escarela and Carriere (2003) considered the model (3) where both $S_T(t; \boldsymbol{\theta})$ and $S_U(u; \boldsymbol{\gamma})$ are Weibull survival functions. Since the copula in the model (3) captures the dependence structure for the joint survival function (rather than the joint distribution function), the tail dependence of the copula is rotated 180 degrees. For instance, the Clayton copula yields an upper tail dependence while the Joe copula yields a lower tail dependence between T and U. Figure 2 compares the scatter plots for (T_i, U_i), $i = 1, \ldots, 1000$, between the Clayton copula and Joe copula under the piecewise exponential models of Example 1. The two copulas have the same Kendall tau $\tau(\alpha) = 0.5$, but show remarkably different dependence patterns in the upper and lower tails.

4.2 Maximum likelihood estimation

For the maximum likelihood estimator (MLE) under the model (3) to be well-defined, we impose a mild assumption that all the derivatives $C_\alpha^{[i,j]}(v, w) = \partial^{(i+j)} C_\alpha(v, w)/\partial v^i \partial w^j$ exist for $(i, j) = (1, 0), (0, 1), (2, 0), (1, 1), (0, 2), (3, 0), (2, 1), (1, 2), (0, 3)$. We give the explicit forms for $C_\alpha^{[i,j]}(v, w)$ under the Clayton and Joe copulas in Online Resource B. Let

$$D_{\alpha,1}(v, w) = C_\alpha^{[1,0]}(v, w)/C_\alpha(v, w), \quad D_{\alpha,2}(v, w) = C_\alpha^{[0,1]}(v, w)/C_\alpha(v, w),$$

which are related to cause-specific hazards (crude hazards) under dependent censoring (Rivest and Wells 2001; Chen 2010; Emura and Chen 2016). For instance, the Clayton copula gives

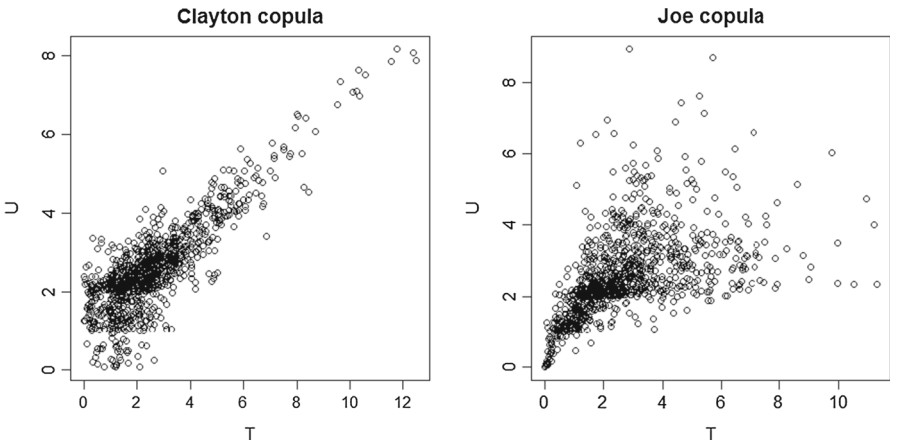

Fig. 2 Scatter plots for (T_i, U_i), $i = 1, \ldots, 1000$ under the model (3). The data are generated from the Clayton (*left*) and Joe (*right*) copulas with the same Kendall tau, $\tau(\alpha) = 0.5$ under the piecewise exponential models of Example 1. Algorithms for generating the data are given in Sect. 5

$$D_{\alpha,1}(v, w) = v^{-\alpha-1}(v^{-\alpha} + w^{-\alpha} - 1)^{-1}, \quad D_{\alpha,2}(v, w) = w^{-\alpha-1}(v^{-\alpha} + w^{-\alpha} - 1)^{-1}$$

and the Joe copula gives

$$D_{\alpha,1}(v, w) = \frac{A_\alpha(v, w)^{1/\alpha-1}\{1 - (1 - w)^\alpha\}(1 - v)^{\alpha-1}}{1 - A_\alpha(v, w)^{1/\alpha}},$$
$$D_{\alpha,2}(v, w) = \frac{A_\alpha(v, w)^{1/\alpha-1}\{1 - (1 - v)^\alpha\}(1 - w)^{\alpha-1}}{1 - A_\alpha(v, w)^{1/\alpha}},$$

where $A_\alpha(v, w) = (1 - v)^\alpha + (1 - w)^\alpha - (1 - v)^\alpha(1 - w)^\alpha$.

Let T_i be event time and U_i be dependent censoring time for $i = 1, \ldots, n$. Observed data consist of $t_i = \min\{T_i, U_i, M\}$, $\delta_i = \mathbf{I}\{T_i \le U_i\}$, and $\rho_i = \mathbf{I}\{\min(T_i, U_i) < M\}$, where $\mathbf{I}\{\cdot\}$ is the indicator function. Given observations (t_i, δ_i, ρ_i), $i = 1, \ldots, n$, the log-likelihood function is

$$\begin{aligned}\ell(\boldsymbol{\theta}, \boldsymbol{\gamma}) = &\sum_{i=1}^{n} \rho_i \delta_i \left[\log f_T(t_i; \boldsymbol{\theta}) + \log D_{\alpha,1}\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}\right] \\ &+ \sum_{i=1}^{n} \rho_i (1 - \delta_i) \left[\log f_U(t_i; \boldsymbol{\gamma}) + \log D_{\alpha,2}\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}\right] \\ &+ \sum_{i=1}^{n} \log C_\alpha\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}, \end{aligned} \tag{4}$$

where f_T and f_U are the probability density functions as defined in Sect. 3. Under the case of the independence copula, we have $D_1(v, w) = v^{-1}$ and $D_2(v, w) = w^{-1}$. Hence, Eq. (4) reduces to the log-likelihood under the independent risks.

We let $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})$ be the MLE that maximizes $\ell(\boldsymbol{\theta}, \boldsymbol{\gamma})$. It suffices to find $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})$ that holds $\partial \ell(\boldsymbol{\theta}, \boldsymbol{\gamma})/\partial(\boldsymbol{\theta}, \boldsymbol{\gamma}) = \mathbf{0}$ with negative definiteness of the converged Hessian matrix

$$H(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}}) = \left. \frac{\partial^2 \ell(\boldsymbol{\theta}, \boldsymbol{\gamma})}{\partial(\boldsymbol{\theta}, \boldsymbol{\gamma}) \partial(\boldsymbol{\theta}, \boldsymbol{\gamma})^{\mathrm{T}}} \right|_{(\boldsymbol{\theta}, \boldsymbol{\gamma}) = (\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})}.$$

For instance, the "nlm()" function in R offers a convenient and reliable Newton-type optimization algorithm, where the score vector and Hessian matrix are internally calculated from the log-likelihood function defined by users. Recommended starting values are $\theta_1 = \cdots = \theta_m = \gamma_1 = \cdots = \gamma_m = 0$. In the "nlm()" function, the matrix $H(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})$ is automatically obtained by specifying the option "nlm(,hessian=TRUE)".

Remark In the log-likelihood of Staplin et al. (2015), the terms corresponding to $D_{\alpha,1}\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}$, $D_{\alpha,2}\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}$, and $C_\alpha\{S_T(t_i; \boldsymbol{\theta}), S_U(t_i; \boldsymbol{\gamma})\}$ in Eq. (4) must be calculated by doing some numerical integrations of the joint density of T_i and U_i. The necessity of the numerical integrations is attributed to their model specification in terms of the joint density $f_{U|T}(u|T = t) f_T(t)$. In contrast, we made our model specification in terms of the copula function between T_i and U_i in Eq. (3). Our approach only requires partial derivatives of the copula function that are technically easier than the numerical integrations. Dependence between T_i and U_i may arise as a consequence of ignoring covariates (environmental, clinical or genetic factors), which is often the case of meta-analysis (Emura et al. 2015, 2017). In such a case, the copula model arises naturally (Emura and Chen 2016).

4.3 Identifiability and sensitivity analysis

For any underlying distribution on (T, U), Tsiatis (1975) constructed the independent random variables T^* and U^* which satisfy

$$(\min(T, U), \mathbf{I}(T \leq U)) \stackrel{d}{=} (\min(T^*, U^*), \mathbf{I}(T^* \leq U^*)),$$

where $\stackrel{d}{=}$ implies the equality of the distribution. The above expression states that observed data cannot distinguish the underlying model from the independent model, a phenomenon known as "nonidentifiability".

The nonidentifiability of Tsiatis (1975) was derived under the non-parametric setting. The nonidentifiability may not occur if the model of (T, U) is restricted to some parametric or semi-parametric classes. Model parameters were claimed to be identifiable in some bivariate parametric classes with one- or two-parameter margins (David and Moeschberger 1978; Basu and Ghosh 1978), parametric Weibull regression model (Escarela and Carriere 2003), and the semi-parametric Cox regression model (Heckman and Honore 1989). However, although the parameters are identifiable, estimation of the dependence parameter remains difficult or unrealistic (Chen 2010; Hsu et al. 2016).

Therefore, it is natural to imagine the difficulty of estimating α in the model (3), though all the model parameters may probably be identifiable. Accordingly, when

performing maximum likelihood inference based on $\ell(\boldsymbol{\theta}, \gamma)$, it is more effective to assume that the copula form and the copula parameter α are known. Then, we suggest a sensitivity analysis that tries a range of possible α and see how the results change. This approach has been well-established in semiparametric models (Rivest and Wells 2001; Chen 2010; De Uña-Álvarez and Veraverbeke 2013) and the piecewise exponential model (Staplin et al. 2015). Further explanation of the sensitivity analysis is given in the data analysis.

4.4 Standard error and confidence interval

We apply the asymptotic normality of the MLE to obtain the standard error (SE) and the confidence interval (CI) for target parameters. For instance, the SE of $\hat{\theta}_j$ is $SE(\hat{\theta}_j) = \sqrt{[-H^{-1}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})]_{\theta_j}}$, $j = 1, \ldots, m$. The $(1-\beta) \times 100\%$ confidence interval for θ_j is $\hat{\theta}_j \pm Z_{\beta/2} \times SE(\hat{\theta}_j)$, where Z_p is the p-th upper quantile for $N(0, 1)$. The delta method is used to obtain the SE of $S_T(t; \hat{\boldsymbol{\theta}})$ as

$$SE\{S_T(t; \hat{\boldsymbol{\theta}})\} = \sqrt{\left\{\partial_{\theta_1,\ldots,\theta_j} S_T(t; \hat{\boldsymbol{\theta}})\right\}^{\mathrm{T}} \times \left[-H^{-1}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\gamma}})\right]_{\theta_1,\ldots,\theta_j} \times \left\{\partial_{\theta_1,\ldots,\theta_j} S_T(t; \hat{\boldsymbol{\theta}})\right\}},$$

for $t \in (a_{j-1}, a_j]$, where

$$\partial_{\theta_1,\ldots,\theta_j} S_T(t; \boldsymbol{\theta}) = -\left[e^{\theta_1} a_1, e^{\theta_2}(a_2 - a_1), \ldots, e^{\theta_{j-1}}(a_{j-1} - a_{j-2}), e^{\theta_j}(t - a_{j-1})\right]^{\mathrm{T}} S_T(t; \boldsymbol{\theta}).$$

Similarly, the $(1 - \beta) \times 100\%$ CI for $S_T(t; \hat{\boldsymbol{\theta}})$ is $S_T(t; \hat{\boldsymbol{\theta}}) \pm Z_{\beta/2} \times SE\{S_T(t; \hat{\boldsymbol{\theta}})\}$.

4.5 Goodness-of-fit test

As the proposed maximum likelihood method strongly relies on the model assumptions, we supplement the method with a formal goodness-of-fit test for testing

$$H_0 : \Pr(T > t, U > u) = C_\alpha\{S_T(t; \boldsymbol{\theta}), S_U(u; \boldsymbol{\gamma})\}, \quad \exists(C_\alpha, \boldsymbol{\theta}, \boldsymbol{\gamma}).$$

Since our target parameter is S_T, it is intuitive to assess the goodness-of-fit in terms of the distance between a semiparametric estimator ($\hat{S}_T$) and the proposed parametric estimator of $S_T(t)$. Particularly, the Cramér–von-Mises type statistics is

$$C = \int_0^\infty \left[\sqrt{n}\{\hat{S}_T(t) - S_T(t; \hat{\boldsymbol{\theta}})\}\right]^2 dF_n(t) = \sum_i \delta_i \rho_i \left\{\hat{S}_T(t_i) - S_T(t_i; \hat{\boldsymbol{\theta}})\right\}^2,$$

where $F_n(t) = \sum_{i=1}^n \delta_i \rho_i \mathbf{I}\{t_i \le t\}/n$. A large value of C indicates a possible misspecification in one of the three parametric forms C_α, $S_T(\cdot; \boldsymbol{\theta})$, and $S_U(\cdot; \boldsymbol{\gamma})$.

As we mentioned in Sect. 3, the important advantage of the piecewise exponential model is the ability to approximate accurately any model by a careful choice of the knot

sequence. It follows that there may be little concern or interest to test the goodness-of-fit for the models on $S_T(\cdot;\boldsymbol{\theta})$ and $S_U(\cdot;\boldsymbol{\gamma})$. This implies that a large value of C can be reasonably interpreted as a misspeficication of C_α.

We particularly pick up a semiparametric estimator of $S_T(t)$, called the generalized copula-graphic (GCG) estimator (De Uña-Álvarez and Veraverbeke 2013), as it can be applied to the present data structure. The GCG estimator is derived under the Archimedean copula model

$$\Pr(T > t, U > u) = S(t,u) = \phi_\alpha^{-1}\left[\{\phi_\alpha\{S_T(t)\} + \phi_\alpha\{S_U(u)\}\right],$$

where ϕ_α is a generator function (Nelsen 2006). The GCG estimator of $S_T(t)$ is

$$\hat{S}_T(t) = \phi_\alpha^{-1}\left\{-\frac{1}{n}\sum_{i=1}^{n}\phi_\alpha'\{\bar{H}_n(t_i)\}\mathbf{I}(t_i \le t)\delta_i\rho_i\right\},$$

where $\bar{H}_n(t) = 1 - \sum_{i=1}^{n}\mathbf{I}(t_i \le t)\rho_i/n$.

Under the Clayton copula with the generator function $\phi_\alpha(t) = (t^{-\alpha} - 1)/\alpha$, the form is

$$\hat{S}_T(t) = \left\{1 + \frac{\alpha}{n}\sum_{i=1}^{n}\bar{H}_n(t_i)^{-\alpha-1}\mathbf{I}(t_i \le t)\delta_i\rho_i\right\}^{-1/\alpha}.$$

Under the Joe copula with the generator function $\phi_\alpha(t) = -\log\{1-(1-t)^\alpha\}$, the form is

$$\hat{S}_T(t) = 1 - \left(1 - \exp\left[-\frac{\alpha}{n}\sum_{i=1}^{n}\frac{\{1-\bar{H}_n(t_i)\}^{\alpha-1}}{1-\{1-\bar{H}_n(t_i)\}^\alpha}\mathbf{I}(t_i \le t)\delta_i\rho_i\right]\right)^{1/\alpha}.$$

To test the hypothesis H_0 with level α, one can apply the parametric bootstrap (Efron and Tibshirani 1993). Let B be a large integer. Then, we perform:

The goodness-of-fit test with parametric bootstrap

Step 1 Generate independent and identically distributed pairs of observations $(t_i^{(b)}, \delta_i^{(b)}, \rho_i^{(b)})$, $i = 1,\ldots,n$, $b = 1,2,\ldots,B$, under the estimated model $\Pr(T > t, U > u) = C_\alpha\{S_T(t;\hat{\boldsymbol{\theta}}), S_U(u;\hat{\boldsymbol{\gamma}})\}$.
Step 2 Compute the bootstrap Cramér–von-Mises statistic $C^{(b)}$ using data $\{(t_i^{(b)}, \delta_i^{(b)}, \rho_i^{(b)}); i = 1,\ldots,n\}$ for each $b = 1,2,\ldots,B$.
Step 3 Reject H_0 with level α if the Cramér–von-Mises statistics C is greater than the $100\times(1-\alpha)$ percent point of $\{C^{(b)}, b = 1,2,\ldots,B\}$.

4.6 Knot selection

In practice, the knot sequence $a_1 < \cdots < a_m$ and the number m are unknown and must be determined by users. The use of the percentiles, such as 33th, 67th, and 100th

percentiles of observed event times, often leads to unreasonable results, especially if the data contain ties or are sparsely scattered. In addition, it is not a valid maximum likelihood approach to treat $a_1 < \cdots < a_m$ and m as unknown parameters to be optimized (Kalbfleisch and Prentice 1973). Alternatively, Kalbfleisch and Prentice (1973) suggested selecting the knot independently from the data and used equally spaced knots in their data analysis.

Rather than fully data-driven routines, the following simple strategy works well in the present setting. We construct an equally spaced knot sequence with $a_j - a_{j-1} = a_1$, $j = 2, \ldots, m$. If an interval $(a_{j-1}, a_j]$ does not contain any observed event time, we set $\theta_j = -\infty$ to be known. Then, maximum likelihood inference is performed for those intervals which contain at least one event times. One may choose the number m such that the displayed survival function facilitates biological interpretations for researchers.

5 Simulations

We performed Monte Carlo simulations to examine the performance of our approach and to compare ours with the GCG estimator. We examined the accuracy of the proposed estimators, standard errors and confidence intervals when the knots are known. We also examine the case where the knots are misplaced and the number of knots increases with sample sizes. Finally, we examined the performance of the proposed goodness-of-fit test.

5.1 Simulation designs

The copula model (3) allows a simple scheme to generate a pair of T_i and U_i. In the first step, we generate a pair (V_i, W_i) from a copula C_α in the following way:

Algorithm: Generate a pair (V_i, W_i) from a copula C_α

Step1 Generate $V_i, V_{i1} \sim U(0, 1)$.
Step2 For the Clayton copula, set $W_i = [V_i^{-\alpha}\{V_{i1}^{-\alpha/(1+\alpha)} - 1\} + 1]^{-1/\alpha}$; For the Joe copula, set W_i as the solution of $V_{i1} = A_\alpha(V_i, W_i)^{1/\alpha - 1}\{1-(1-W_i)^\alpha\}(1-V_i)^{\alpha-1}$, where $A_\alpha(v, w) = (1-v)^\alpha + (1-w)^\alpha - (1-v)^\alpha(1-w)^\alpha$.

In all the simulations, we assume that α is known at the value $\tau(\alpha) = 0.5$, which corresponds to $\alpha = 2$ under the Clayton copula and $\alpha = 2.856$ under the Joe copula (Sect. 4.1). In the next step, we use inverse transformations $T_i = S_T^{-1}(V_i; \boldsymbol{\theta})$ and $U_i = S_U^{-1}(W_i; \boldsymbol{\gamma})$ under piecewise constant hazards with $m = 3$ knots, $a_1 = 1, a_2 = 2$, and $a_3 = M = 3$ (Example 1). The explicit forms of the inverse transformations are

$$
\begin{aligned}
&S_T^{-1}(v; \theta_1, \theta_2, \theta_3) \\
&= \begin{cases}
-e^{-\theta_2}\log(v) & \text{if } 0 > \log(v) \geq -e^{\theta_1}a_1, \\
a_1 - e^{-\theta_2}\{\log(v) + e^{\theta_1}a_1\} & \text{if } -e^{\theta_1}a_1 > \log(v) \geq -e^{\theta_1}a_1 - e^{\theta_2}(a_2 - a_1), \\
a_2 - e^{-\theta_3}\{\log(v) + e^{\theta_1}a_1 + e^{\theta_2}(a_2 - a_1)\} & \text{if } -e^{\theta_1}a_1 - e^{\theta_2}(a_2 - a_1) > \log(v).
\end{cases}
\end{aligned}
$$

Table 1 The misplaced knots sequences where m grows with n

	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}
$n=200, m=5$	0.0	0.6	1.2	1.8	2.4	3												
$n=400, m=11$	0.0	0.3	0.6	0.8	1.1	1.4	1.6	1.9	2.2	2.5	2.7	3						
$n=600, m=17$	0.0	0.2	0.4	0.5	0.7	0.9	1.1	1.2	1.4	1.6	1.8	1.9	2.1	2.3	2.5	2.7	2.8	3

The true knot sequence is $a_1 = 1$, $a_2 = 2$, and $a_3 = M = 3$ with $m = 3$
These sequences are generated by rounding the results of R commands, *seq(0,3,length=6)*, *seq(0,3,length=12)*, and *seq(0,3,length=18)*

and $S_U^{-1}(w; \gamma_1, \gamma_2, \gamma_3)$ in a similar fashion. Parameters for the piecewise exponential hazards $(\theta_1, \theta_2, \theta_3, \gamma_1, \gamma_2, \gamma_3)$ are chosen to yield different observed event proportions.

Scenario (i) $\theta_1 = -2$, $\theta_2 = -1$, $\theta_3 = -0.5$,$\gamma_1 = -3$, $\gamma_2 = -1.5$, $\gamma_3 = 0$: T is moderately censored as $\Pr(T \le U, T < M) = 0.46$ (Clayton copula) and $\Pr(T \le U, T < M) = 0.37$ (Joe copula).

Scenario (ii) $\theta_1 = -2$, $\theta_2 = -2$, $\theta_3 = -2$,$\gamma_1 = -3$, $\gamma_2 = -1.5$, $\gamma_3 = 0$: T is heavily censored as $\Pr(T \le U, T < M) = 0.22$ (Clayton copula) and $\Pr(T \le U, T < M) = 0.13$ (Joe copula).

After a pair (T_i, U_i) is generated, we set $t_i = \min\{T_i, U_i, M\}$, $\delta_i = \mathbf{I}\{T_i \le U_i\}$, and $\rho_i = \mathbf{I}\{t_i < M\}$. Based on observations (t_i, δ_i, ρ_i), $i = 1, \ldots, n$, we calculate the estimators for the survival function $S_T(t; \boldsymbol{\theta})$ in three different ways using:

(a) Piecewise exponential model with known knots; $a_1 = 1, a_2 = 2$, and $a_3 = M = 3$ with $m = 3$
(b) Piecewise exponential model where the knots are misplaced and the number of knots increases (Table 1)
(c) GCG estimator

Based on 500 repetitions, we compare the performance of the three estimators.

5.2 Results for the proposed estimator

Table 2 shows the simulation results under the Clayton copula. If the knots are known, the estimates appear to be nearly unbiased for $S_T(1; \boldsymbol{\theta})$, $S_T(2; \boldsymbol{\theta})$, and $S_T(3; \boldsymbol{\theta})$. The standard deviation (SD) of the estimates decreases as the sample size increases from $n = 200$ to 600. The values of the SE are all very close to the values of the SD. Accordingly, the resulting 95% confidence intervals give satisfactory coverage rates.

A close inspection of Table 2 reveals some patterns for the SD. First, the SD for $S_T(3; \boldsymbol{\theta})$ is larger than the SD for $S_T(1; \boldsymbol{\theta})$. The reason is that $S_T(3; \boldsymbol{\theta})$ involves three parameters, say $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, which makes more variation than estimating $S_T(1; \boldsymbol{\theta})$ involving only θ_1. Second, Scenario ii) gives the larger SD for $S_T(3; \boldsymbol{\theta})$ than Scenario i) does. This is due to the heavier censoring percentage of Scenario ii), yielding more uncertainty in estimation. However, Scenario ii) gives the smaller SD for $S_T(2; \boldsymbol{\theta})$

Table 2 Simulation results for the proposed estimators and generalized copula-graphic (GCG) estimator under the Clayton copula with $\alpha = 2$ based on 500 Monte Carlo replications

		Piecewise (known knots)				Piecewise (misplaced knots)		GCG estimator	
		Mean	SD	SE	95% Cov	Mean	SD	Mean	SD
Scenario (i) $\boldsymbol{\theta} = (-2, -1, -0.5)$, $\boldsymbol{\gamma} = (-3, -1.5, 0)$; $\Pr(T \le U, T < M) = 0.46$, $\Pr(U < T, U < M) = 0.32$									
$n = 200$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.873	0.024	0.024	0.956	0.846	0.023	0.872	0.024
	$S_T(2; \boldsymbol{\theta}) = 0.605$	0.604	0.037	0.037	0.944	0.585	0.034	0.602	0.038
	$S_T(3; \boldsymbol{\theta}) = 0.330$	0.330	0.042	0.041	0.952	0.336	0.043	0.328	0.045
$n = 400$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.873	0.017	0.017	0.956	0.861	0.017	0.872	0.017
	$S_T(2; \boldsymbol{\theta}) = 0.605$	0.603	0.025	0.026	0.960	0.595	0.025	0.602	0.026
	$S_T(3; \boldsymbol{\theta}) = 0.330$	0.330	0.029	0.029	0.950	0.330	0.031	0.328	0.031
$n = 600$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.872	0.014	0.014	0.948	0.864	0.014	0.872	0.014
	$S_T(2; \boldsymbol{\theta}) = 0.605$	0.603	0.021	0.022	0.952	0.597	0.021	0.602	0.022
	$S_T(3; \boldsymbol{\theta}) = 0.330$	0.330	0.024	0.024	0.944	0.330	0.026	0.329	0.026
		Mean	SD	SE	95% Cov	Mean	SD	Mean	SD
Scenario (ii) $\boldsymbol{\theta} = (-2, -2, -2)$, $\boldsymbol{\gamma} = (-3, -1.5, 0)$; $\Pr(T \le U, T < M) = 0.22$, $\Pr(U < T, U < M) = 0.51$									
$n = 200$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.873	0.024	0.024	0.956	0.871	0.022	0.872	0.024
	$S_T(2; \boldsymbol{\theta}) = 0.763$	0.763	0.032	0.033	0.952	0.758	0.031	0.761	0.033
	$S_T(3; \boldsymbol{\theta}) = 0.666$	0.667	0.050	0.048	0.940	0.664	0.054	0.666	0.058
$n = 400$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.873	0.017	0.017	0.958	0.872	0.016	0.872	0.017
	$S_T(2; \boldsymbol{\theta}) = 0.763$	0.762	0.023	0.023	0.950	0.761	0.023	0.761	0.023
	$S_T(3; \boldsymbol{\theta}) = 0.666$	0.666	0.035	0.034	0.946	0.666	0.040	0.665	0.040
$n = 600$	$S_T(1; \boldsymbol{\theta}) = 0.873$	0.872	0.014	0.014	0.950	0.872	0.014	0.872	0.014
	$S_T(2; \boldsymbol{\theta}) = 0.763$	0.762	0.019	0.019	0.948	0.761	0.019	0.761	0.020
	$S_T(3; \boldsymbol{\theta}) = 0.666$	0.665	0.029	0.028	0.936	0.665	0.033	0.664	0.034

Mean (SD) = the mean (standard deviation) of the estimates, SE = the mean of the SEs, 95% Cov = the empirical coverage rate of the 95% confidence intervals

than Scenario i) does. This phenomenon occurs since the value of $S_T(2;\boldsymbol{\theta}) = 0.763$ under Scenario ii) is closer to 1 than the value $S_T(2;\boldsymbol{\theta}) = 0.605$ under Scenario i). These patterns are commonly observed in many available estimators for survival functions.

Table 2 also shows the simulation results under the misplaced knot sequences. In Scenario i), the estimates exhibit systematic biases, especially for the small sample size ($n = 200$). Nevertheless, the biases tend to vanish as the sample sizes increase from $n = 200$ to $n = 600$. This is because the approximation to the true survival function gets improved as the he number of knots increases with the sample sizes. In Scenario ii), the estimates appear to be almost unbiased even under the misplaced knots. This phenomenon is due to the homogeneous parameter values in all intervals (i.e., $\theta_1 = -2$, $\theta_2 = -2$, and $\theta_3 = -2$). In this case, the true survival function can be correctly approximated by any knot sequence.

Table 3 summarizes simulation results under the Joe copula. The patterns of the results are very similar to those under the Clayton copula. All the parameters are nearly unbiasedly estimated with satisfactory performances on the SE and confidence interval. Even under the misplaced knot sequences, the estimates are virtually unbiased under Scenario ii) due to the homogeneous parameter setting (i.e., $\theta_1 = -2.5$, $\theta_2 = -2.5$, and $\theta_3 = -2.5$). Also in Scenario i), the biases due to the misplaced knots are mitigated by increasing the sample size (i.e., increasing the number of knots).

5.3 Comparison with the GCG estimator

Tables 2 and 3 compare the proposed estimator with the GCG estimator. In all parameter settings, the proposed estimator shows the advantage over the GCG estimator in terms of the bias and SD, but the difference is very modest. The results are somewhat surprising since the GCG estimator performs competitively without requiring any parametric forms for marginal distributions. The comparison between the two estimators will be discussed further in Sect. 7.

5.4 Performance of the goodness-of-fit test

The performance of the proposed goodness-of-fit test (Sect. 4.5) is assessed via simulations. In this study, we set the null hypothesis to be the Clayton copula while we generated data from four different copulas (Clayton, Joe, Plackett, and independent). The parameter α in the Clayton, Joe and Plackett copulas is assumed known at the value $\tau(\alpha) = 0.5$. For each repetition, we calculate the Cramér–von-Mises statistic and test the goodness-of-fit hypothesis (the null hypothesis of the Clayton copula with $\alpha = 2$) based on $B = 500$ bootstraps. To examine Type I error rate and power of the test, we count the number of rejections under levels 0.01, 0.05, or 0.10 during 500 repetitions.

Table 4 shows the performance of the goodness-of-fit test. When the data are generated from the Clayton copula (null hypothesis), the rejection rates are in good agreement with all the pre-specified levels. In addition, the mean of the Cramér–von-

Table 3 Simulation results for the proposed estimators and generalized copula-graphic (GCG) estimator under the Joe copula with $\alpha = 2.856$ based on 500 Monte Carlo replications

		Piecewise (known knots)				Piecewise (misplaced knots)		GCG estimator	
		Mean	SD	SE	95% Cov	Mean	SD	Mean	SD
Scenario (i) $\boldsymbol{\theta} = (-2, -1.5, -1)$, $\boldsymbol{\gamma} = (-3, -1.5, -1.5)$; $\Pr(T \le U, T < M) = 0.37$, $\Pr(U < T, U < M) = 0.19$									
$n = 200$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.921	0.018	0.018	0.938	0.903	0.018	0.919	0.018
	$S_T(2;\boldsymbol{\theta}) = 0.737$	0.737	0.030	0.030	0.946	0.724	0.027	0.734	0.030
	$S_T(3;\boldsymbol{\theta}) = 0.510$	0.510	0.033	0.035	0.952	0.511	0.033	0.507	0.033
$n = 400$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.921	0.013	0.013	0.946	0.913	0.013	0.920	0.013
	$S_T(2;\boldsymbol{\theta}) = 0.737$	0.736	0.021	0.021	0.954	0.730	0.020	0.735	0.021
	$S_T(3;\boldsymbol{\theta}) = 0.510$	0.510	0.024	0.025	0.962	0.510	0.024	0.509	0.024
$n = 600$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.920	0.011	0.011	0.964	0.915	0.011	0.920	0.011
	$S_T(2;\boldsymbol{\theta}) = 0.737$	0.736	0.018	0.017	0.950	0.732	0.017	0.735	0.018
	$S_T(3;\boldsymbol{\theta}) = 0.510$	0.510	0.019	0.020	0.962	0.510	0.019	0.510	0.019
		Mean	SD	SE	95% Cov	Mean	SD	Mean	SD
Scenario (ii) $\boldsymbol{\theta} = (-2.5, -2.5, -2.5)$, $\boldsymbol{\gamma} = (-3, -1.5, -1.5)$; $\Pr(T \le U, T < M) = 0.13$, $\Pr(U < T, U < M) = 0.29$									
$n = 200$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.922	0.018	0.018	0.928	0.920	0.017	0.919	0.018
	$S_T(2;\boldsymbol{\theta}) = 0.849$	0.849	0.025	0.024	0.928	0.847	0.022	0.847	0.027
	$S_T(3;\boldsymbol{\theta}) = 0.782$	0.782	0.031	0.031	0.952	0.781	0.031	0.780	0.032
$n = 400$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.921	0.013	0.013	0.944	0.921	0.012	0.920	0.013
	$S_T(2;\boldsymbol{\theta}) = 0.849$	0.848	0.017	0.017	0.934	0.848	0.017	0.848	0.018
	$S_T(3;\boldsymbol{\theta}) = 0.782$	0.782	0.022	0.022	0.940	0.782	0.022	0.781	0.022
$n = 600$	$S_T(1;\boldsymbol{\theta}) = 0.921$	0.920	0.011	0.011	0.964	0.920	0.011	0.920	0.011
	$S_T(2;\boldsymbol{\theta}) = 0.849$	0.848	0.015	0.014	0.934	0.848	0.015	0.848	0.015
	$S_T(3;\boldsymbol{\theta}) = 0.782$	0.782	0.018	0.018	0.952	0.782	0.018	0.782	0.018

Mean (SD) = the mean (standard deviation) of the estimates, SE = the mean of the SEs, 95% Cov = the empirical coverage rate of the 95% confidence intervals

Table 4 Simulation results for the goodness-of-fit test for the Clayton copula

Underlying copula	Sample size	Test statistics $E[C]$	Bootstrap statistics $E[C^{(\cdot)}]$	Rejection rate at level 0.01	Rejection rate at level 0.05	Rejection rate at level 0.10
Clayton(correct)	$n = 200$	0.024	0.025	0.010	0.030	0.088
	$n = 400$	0.023	0.023	0.014	0.044	0.112
	$n = 600$	0.022	0.022	0.004	0.036	0.092
Joe(incorrect)	$n = 200$	0.029	0.028	0.016	0.040	0.104
	$n = 400$	0.034	0.027	0.042	0.108	0.192
	$n = 600$	0.039	0.026	0.080	0.196	0.328
Plackett(incorrect)	$n = 200$	0.024	0.026	0.014	0.034	0.066
	$n = 400$	0.024	0.025	0.022	0.056	0.096
	$n = 600$	0.025	0.024	0.024	0.074	0.138
Independence(incorrect)	$n = 200$	0.037	0.036	0.012	0.066	0.134
	$n = 400$	0.039	0.032	0.016	0.090	0.180
	$n = 600$	0.042	0.032	0.046	0.132	0.240

Results are based on 500 Monte Carlo replications under $\theta_1 = -2$, $\theta_2 = -1$, $\theta_3 = -0.5$, $\gamma_1 = -3$, $\gamma_2 = -1.5$, and $\gamma_3 = 0$
$E[C]$ = The mean of the Cramér–von Mises test statistics C
$E[C^{(\cdot)}]$ =The mean of the averaged bootstrap Cramér–von Mises test statistics $\sum_{b=1}^{B} C^{(b)}/B$, where B =500

Mises statistics C is fairly close to the mean of the averaged bootstrap Cramér–von Mises statistics $\sum_{b=1}^{B} C^{(b)}/B$. Hence, the bootstrap approximation to the null distribution of the test statistic works well.

When data are generated from the Joe copula, the rejection rates are systematically higher than the specified levels (Table 4). The rejection rates increase as the sample sizes increase from $n = 200$ to 600. Indeed, the Joe copula and Clayton copula show remarkably different dependence patterns even though they have the same Kendall's tau (Fig. 2). However, the rejection rate is still 0.196 (at level 0.05) even under $n = 600$, showing about 20% power to reject the null hypothesis. The modest power is also observed when the data are generated from the independence model. The rejection rate is 0.132 (at level 0.05) under $n = 600$, showing about 13% power to reject the null hypothesis.

When data are generated from the Plackett copula, the rejection rates does not differ much from the specified Type I error rates (Table 4). However, as the sample sizes increase from 200 to 600, rejection rates are slowly increasing up to the rejection rate 0.074 (at level 0.05). This implies the poor ability of distinguishing between the Clayton copula and Plackett copula.

Overall, the proposed goodness-of-fit test offers a good control for Type I error rate and exhibits modest power. In practice, the test can reject the null hypothesis only when the underlying copula structure is remarkably different from the null hypothesis, or when the sample size is very large. The power properties are natural due to the inherent difficulty of identifying the dependence model from the competing risks data.

6 Data analysis

6.1 Salamander data

Michimae et al. (2014) conducted a designed follow-up study for salamander larvae living in Hokkaido, Japan. At the beginning of the follow-up, a randomization was used to divide the larvae into three experimental groups of morphologies "broad head" ($n = 30$), "anti-predator" ($n = 30$), and "control" ($n = 30$). The event time of interest is days from placing the salamander larvae in the experimental cage to the completion of metamorphosis. During the follow-up, the presence of metamorphosis was examined, and the time to metamorphosis was recorded as the first day that exhibits metamorphosis.

As pointed out by Michimae et al. (2014), their collected data comprise the competing risks setting. Time to metamorphosis is censored if death occurs before metamorphosis; time to death is censored if metamorphosis occurs before death. Although death might still be observed even after metamorphosis, time to death measured after metamorphosis may lose the original meaning. In other words, our target outcome is "death in a larval stage", which becomes undefined after larvae become adult salamanders. This confounding effect is not only restricted to our animal experiment but also seen in clinical trials involving human subjects. For instance, a medical intervention at the time of cancer relapse may confound the original interpretation of overall survival measured from a randomization (Pazdur 2008; Buyse et al. 2011; Emura et al. 2015).

Statistical analyses of Michimae et al. (2014) were based on the estimated sub-distribution functions (i.e., estimated probability of metamorphosis in the presence of death over time). They compared the estimated sub-distribution functions between the three experimental groups to see how the experimental condition affects the metamorphic timing. They concluded that the broad head group had shorter incidence for metamorphosis than the other two groups did.

Instead of their comparison in terms of the difference of the sub-distribution functions between the three groups, our proposed method offers a comparison in terms the survival functions for time to metamorphosis. Survival function would be more straightforwardly understood by many biologists due to its popularity (Kuparinen et al. 2008).

We remove the control group that has no dependently censored (dead) subjects. Our goal is to estimate the two survival functions of time to metamorphosis for the broad head group ($n = 30$ with 4 death) and the anti-predator group ($n = 30$ with 8 death), separately. According to the range of the observed time to metamorphosis and the sample sizes, we choose the following equally spaced knot sequences: $a_1 = 25$, $a_2 = 50$, and $a_3 = M = 75$ for the broad head group; $a_1 = 33.3$, $a_2 = 66.6$, and $a_3 = M = 100$ for the anti-predator group. This choice turns out to facilitate a biological interpretation for comparing the metamorphic timing between two groups.

6.2 Results

Initially, we performed the proposed goodness-of-fit test (Sect. 4.5) for the piecewise exponential model with the Clayton copula with $\alpha = 2$ ($\tau = 0.5$). By fitting this

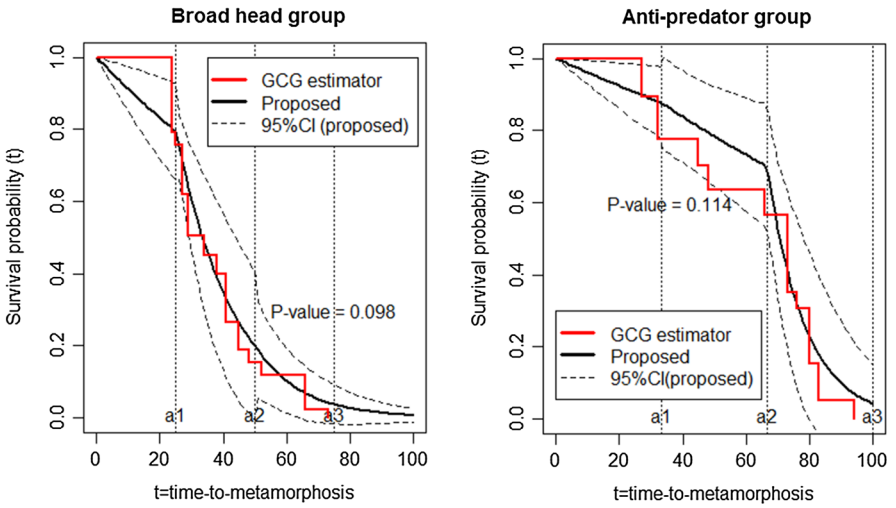

Fig. 3 Estimated survival functions based on the salamander data. The Clayton copula with $\alpha = 2\,(\tau = 0.5)$ is fitted to estimate the survival function using the proposed estimator (*black line*) and the GCG estimator (*red line*). The *left panel* shows the broad head group ($n = 30$) and the *right panel* shows the anti-predator group ($n = 30$). 95% confidence intervals are based on the normal approximation and P-values of the goodness-of-fit test are based on the parametric bootstrap method. The knot sequences are; $a_1 = 25$, $a_2 = 50$, and $a_3 = M = 75$ for the broad head group; $a_1 = 33.3$, $a_2 = 66.6$, and $a_3 = M = 100$ for the anti-predator group (Color figure online)

model, the estimated survival function was in good agreement with the estimated survival function with the GCG estimator (Fig. 3). The tests for a goodness-of-fit hypothesis based on the distance between the two estimated survival functions was not rejected at 5% level (P-value = 0.098 for the broad head group; P-value = 0.114 for the anti-predator group). Therefore, there is little evidence against the piecewise exponential models with the Clayton copula with $\alpha = 2$ ($\tau = 0.5$).

More generally, the goodness-of-fit test with the Clayton model was not rejected in the range $0.1 \leq \alpha \leq 2.5$ ($0.05 \leq \tau \leq 0.56$) with 5% level, but the test with the independence copula was rejected for both the broad head group and anti-predator group (P-value <0.05). This result gives some evidence against the independence assumption between time-to-metamorphosis and time to death. Hereafter we adopt the Clayton copula and examine the sensitivity of the results in the range $0.1 \leq \alpha \leq 2.5$.

We also fitted the Joe copula with $\alpha = 0.2856$ ($\tau = 0.5$) and then obtained very similar estimated survival functions as those under the Clayton copula in Fig. 3. However, the numerical algorithm for maximizing the log-likelihood becomes more sensitive to the choice of the initial values, which is especially inconvenient to perform the bootstrap replications. In this respect, we center our analysis on the Clayton copula.

We compared the estimated survival functions between the broad head group and anti-predator group with piecewise exponential models under the Clayton copula. The comparison is performed in terms of a sensitivity analysis in the range $0.1 \leq \alpha \leq 2.5$ (Fig. 4). Clearly, the broad head group has lower survival probability than the anti-predator group in all the range of α, providing convincing evidence of shorter time

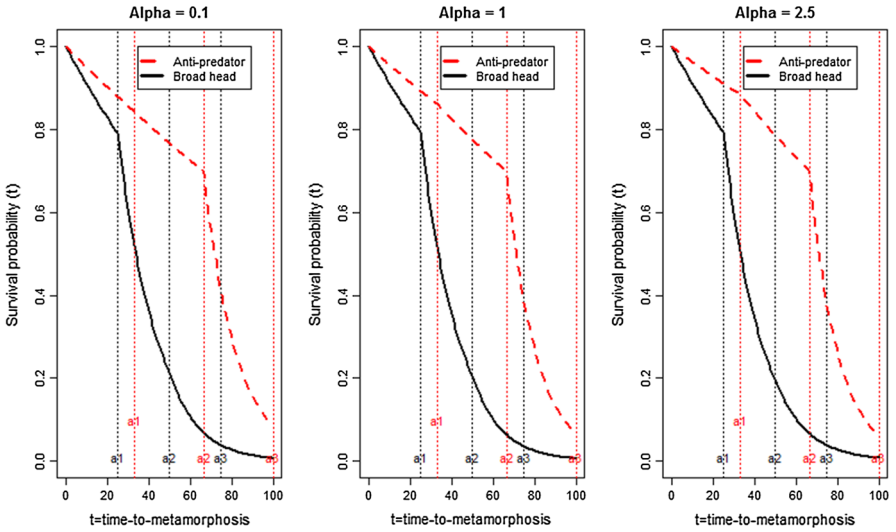

Fig. 4 Estimated survival functions with piecewise exponential models based on the salamander data. The dependence model is the Clayton copula with $\alpha = 0.1$ (*left*, $\tau = 0.05$), $\alpha = 1$ (*center*, $\tau = 0.33$), and $\alpha = 2.5$ (*right*, $\tau = 0.56$). The *solid* (*black*) *line* depicts the estimated survival function for the broad head group ($n = 30$) and the *dashed* (*red*) *line* depicts the estimated survival function for the anti-predator group ($n = 30$) (Color figure online)

to metamorphosis in the broad head group. This conclusion agrees with the results of Michimae et al. (2014) who utilized the sub-distribution functions for the comparison.

A close look to Fig. 4 reveals that the two survival functions for the broad head and anti-predator groups share some similar characteristics. In both groups, the survival function is initially flat and then drops steeply after $a_1 = 25$ (days) in the broad head group and $a_2 = 66.6$ (days) in the anti-predator group. These time points may be interpreted as "change points", where physiological mechanisms in larvae changes from an inactive state to an active state for metamorphosis (Rose 2005). Accordingly, one can conclude that the experimental intervention for the broad head group expedites the change point by 41.6 days (66.6-25), which appears to be substantial amount. Note that such a conclusion about the change point cannot be easily derived by the GCG estimator.

7 Discussion and conclusion

Our study has demonstrated that piecewise exponential models can be useful for analyzing time-to-event data in the presence of dependent censoring. The present approach differs from the commonly employed approach based on sub-distribution functions (or cumulative incidence functions). We utilize an assumed copula model (Zheng and Klein 1995) as a way to adjust for the effect of dependent censoring. Under the piecewise exponential models with an assumed copula, we developed maximum likelihood inference methods, including point and interval estimations and a goodness-of-fit test. The simulations show that the proposed methods exhibit desirable sampling properties, guaranteeing that valid statistical inference for a survival function is possible.

Escarela and Carriere (2003) and Hsu et al. (2016) developed copula-based models for estimating a parametric survival function in the presence of competing risks. They estimated the copula parameter α by the MLE, but the estimates may have a large sampling variation. Instead, we fixed a few values of α and performed the sensitivity analysis (Rivest and Wells 2001; Chen 2010; De Uña-Álvarez and Veraverbeke 2013; Staplin et al. 2015). Staplin et al. (2015) also used the sensitivity analysis under piecewise exponential models but did not utilize copulas for dependent censoring. The idea of the sensitivity analysis is to examine the data analysis results under differing amount of α which determines the degree of dependence between event time and dependent censoring time. A remarkable feature of our approach is that the copula parameter has a simple relationship with Kendall's tau without being influenced by marginal distributions. We have demonstrated the usefulness of the sensitivity analysis through our real data analysis of the salamander larvae (Sect. 6).

Using simulations, we compared our approach with the existing semiparametric estimator (GCG estimator). The results showed that our estimator exhibits only a minor advantage in terms of the reduction in bias and standard deviation. However, this small gain in bias and standard deviation would not be served as a major advantage of our approach. It is rather an interesting phenomenon that the semiparametric GCG estimator for a survival function is nearly efficient with little loss of bias and variation over our parametric estimator. While the asymptotic properties of the GCG estimator are deeply investigated by De Uña-Álvarez and Veraverbeke (2013), closed form expressions for the interval estimations (the standard error and confidence intervals) have not been available. They proposed the bootstrap-based interval estimation, but its numerical validity remains to be checked. Thus, in terms of interval estimation, our approach provides some advantage over the GCG estimator according to the numerical and theoretical support of the proposed interval estimations. Another advantage of the proposed estimator over the GCG estimator is the applicability of non-Archimedean copulas, such as Gaussian, t- and FGM copulas (Nelsen 2006). For instance, the FGM copula is an interesting choice as it offers an explicit form of Kendall's tau and an analytically tractable likelihood function under the piecewise exponential model. However, the GCG estimator is undefined under the FGM copula.

To supplement the strong parametric assumptions made on the underlying model in our approach, we have devised a formal goodness-of-fit test with aid of the parametric bootstrap. The proposed test is associated with the distance between model-based estimator and model-free estimator of a survival function. In this way, the result of the goodness-of-fit test is not only expressed as a formal decision rule (reject or accept), but also interpreted through the graphical comparison of two estimated survival functions. We have explained how the validity of a chosen parametric model is formally and graphically justified though our real data analysis of the salamander larvae (Sect. 6). In the simulations, where we focused on the goodness-of-fit test for a copula, a good control of Type I error rate and moderate statistical power were observed. This demonstrated the ability to distinguish the underlying copula from a misspecified copula. However, due to the inherent difficulty of identifying the dependence structure between two competing event times, the test can reject the null hypothesis only when a large number of samples are available or the true model is remarkably different from the null model.

An important topic that we did not discuss in this paper is the issue of left-truncation (Lawless 2003; Klein and Moeschberger 2003). Consideration for left-truncation is essential in animal experiments if the time scale of event is age, where the inference focuses on the age-specific hazard (Fieberg and DelGiudice 2011). In this case, left-truncation corresponds to entry age, and the available samples are restricted to those who experience an event after entry. In our salamander data, the entry times (time at placing the salamander larvae in the experimental cage) for all subjects are regarded as the time origin, and so there is no issue for left-truncation. The modified expression of the log-likelihood under independent left-truncation is often easily obtained under parametric models (Lawless 2003; Klein and Moeschberger 2003). De Uña-Álvarez and Veraverbeke (2017) modified their semiparametric GCG estimator to account for left-truncation. In this respect, all the proposed estimation and goodness-of-fit procedures can, in principle, be modified to account for left-truncation. A more challenging but interesting issue is to account for "dependent" left-truncation in competing risks analysis (Bakoyannis and Touloumi 2015). The model of dependent truncation demands another copula model between event time and left-truncation time (Chaieb et al. 2006; Emura and Wang 2012; Emura and Murotani 2015).

Electronic supplementary material: Supplementary Materials include Online Resource A (the salamander data) and Online Resource B (Derivatives of copulas).

Acknowledgements The authors thank the editor and two anonymous reviewers for their helpful comments that improved the paper. This work was supported by Ministry of Science and Technology, Taiwan (MOST 103-2118-M-008-MY2 and 105-2118-M-008-001-MY2 for T. Emura) and Grants-in-Aid for a Research Fellow of the Japan Society for the Promotion of Science (no. 23570036 for H. Michimae).

References

Andersen PK, Abildstrom SZ, Rosthøj S (2002) Competing risks as a multi-state model. Stat Methods Med Res 11(2):203–215

Bakoyannis G, Touloumi G (2012) Practical methods for competing risks data: a review. Stat Methods Med Res 21(3):257–272

Bakoyannis G, Touloumi G (2015) Impact of dependent left truncation in semiparametric competing risks methods: a simulation study. Commun Stat Simul Comput. doi:10.1080/03610918.2015.1030415

Basu AP, Ghosh JK (1978) Identifiability of the multinormal and other distributions under competing risks model. J Multivar Anal 8:413–429

Braekers R, Veraverbeke N (2005) A copula-graphic estimator for the conditional survival function under dependent censoring. Can J Stat 33:429–447

Buyse M, Sargent DJ, Saad ED (2011) Survival is not a good outcome for randomized trials with effective subsequent therapies. J Clin Oncol 29(35):4719–4720

Chaieb LL, Rivest LP, Abdous B (2006) Estimating survival under a dependent truncation. Biometrika 93:665–669

Chen YH (2010) Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. J R Stat Soc Ser B 72:235–251

Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65:141–151

Crowder MJ (2001) Classical competing risks. CRC Press, Boca Raton

Crowder MJ (2012) Multivariate survival analysis and competing risks. CRC Press, Boca Raton

David HA, Moeschberger ML (1978) The theory of competing risks, vol 39. Griffin, London

De Uña-Álvarez J, Veraverbeke N (2013) Generalized copula-graphic estimator. Test 22(2):343–360

De Uña-Álvarez J, Veraverbeke N (2017) Copula-graphic estimation with left-truncated and right-censored data. Statistics. doi:10.1080/02331888.2016.1274898
Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, London
Emura T, Wang W (2012) Nonparametric maximum likelihood estimation for dependent truncation data based on copulas. J Multivar Anal 110:171–188
Emura T, Murotani K (2015) An algorithm for estimating survival under a copula-based dependent truncation model. TEST 24(4):734–751
Emura T, Chen YH (2016) Gene selection for survival data under dependent censoring: a copula-based approach. Stat Methods Med Res 25(6):2840–2857
Emura T, Nakatochi M, Murotani K, Rondeau V (2015) A joint frailty-copula model between tumour progression and death for meta-analysis. Stat Methods Med Res. doi:10.1177/0962280215604510
Emura T, Nakatochi M, Matsui S, Michimae H, Rondeau V (2017) Personalized dynamic prediction of death according to tumour progression and high-dimensional genetic factors: meta-analysis with a joint model. Stat Methods Med Res. doi:10.1177/0962280216688032
Escarela G, Carriere JF (2003) Fitting competing risks with an assumed copula. Stat Methods Med Res 12(4):333–349
Fieberg J, DelGiudice GD (2011) Estimating age-specific hazards from wildlife telemetry data. Environ Ecol Stat 18(2):209–222
Friedman M (1982) Piecewise exponential models for survival data with covariates. Ann Stat 10:101–113
Heckman JJ, Honore BE (1989) The identifiability of the competing risks models. Biometrika 76:325–330
Hsu TM, Emura T, Fan TH (2016) Reliability inference for a copula-based series system life test under multiple type-I censoring. IEEE Trans Reliab 65(2):1069–1080
Joe H (1993) Parametric families of multivariate distributions with given margins. J Multivar Anal 46:262–282
Kalbfleisch JD, Prentice RL (1973) Marginal likelihoods based on Cox's regression and life model. Biometrika 60(2):267–278
Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data. Wiley, Hoboken
Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53(282):457–481
Klein JP, Moeschberger ML (2003) Survival analysis techniques for censored and truncated data. Springer, New York
Kuparinen A, O'Hara RB, Merilä J (2008) Probabilistic models for continuous ontogenetic transition processes. PLoS ONE 3(11):e3677
Lawless JF (2003) Statistical models and methods for lifetime data, 2nd edn. Wiley, New York
Michimae H, Tezuka A, Emura T, Kishida O (2014) Environment-dependent trade-offs and phenotypic plasticity in metamorphic timing. Evol Ecol Res 16:617–629
Nelsen RB (2006) An introduction to copulas. Springer series in statistics, 2nd edn. Springer, New York
Pazdur R (2008) Endpoints for assessing drug activity in clinical trials. Oncologist 13:19–21
Rivest LP, Wells MT (2001) A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. J Multivar Anal 79:138–155
Rose CS (2005) Integrating ecology and developmental biology to explain the timing of frog metamorphosis. Trends Ecol Evolut 20(3):129–135
Staplin ND, Kimber AC, Collett D, Roderick PJ (2015) Dependent censoring in piecewise exponential survival models. Stat Methods Med Res 24(3):325–341
Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. PNAS USA 72:20–22
Zheng M, Klein JP (1995) Estimates of marginal survival for dependent competing risks based on an assumed copula. Biometrika 82(1):127–138

Takeshi Emura is Associate Professor at the Graduate Institute of Statistics, National Central University, Taiwan. His current research focuses on statistical methodologies in survival analysis with biomedical applications, especially based on copula models. He teaches courses on statistical inference, survival analysis, and quality control at the graduate school level (both Master and Ph.D. levels).

Hirofumi Michimae is Assistant Professor at the Department of Clinical Medicine (Biostatistics), School of Pharmacy, Kitasato University, Japan. His current research focuses on biostatistics and robust statistics. He teaches courses on statistics and probability theory at both the undergraduate and graduate school levels.