



City Research Online

City St George's, University of London

Citation: Kaishev, V. K., Dimitrova, D. S. & Haberman, S. (2007). Modelling the joint distribution of competing risks survival times using copula functions. Insurance: Mathematics and Economics, 41(3), pp. 339-361. doi: 10.1016/j.insmatheco.2006.11.006

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/4040/>

Link to published version: <https://doi.org/10.1016/j.insmatheco.2006.11.006>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Modelling the joint distribution of competing risks survival times using copula functions

by

Vladimir K. Kaishev*, Dimitrina S. Dimitrova and Steven Haberman

Cass Business School, City University, London

Abstract

The problem of modelling the joint distribution of survival times in a competing risks model, using copula functions is considered. In order to evaluate this joint distribution and the related overall survival function, a system of non-linear differential equations is solved, which relates the crude and net survival functions of the modelled competing risks, through the copula. A similar approach to modelling dependent multiple decrements was applied by Carriere (1994) who used a Gaussian copula applied to an incomplete double decrement model which makes it difficult to calculate any actuarial functions and draw relevant conclusions. Here, we extend this methodology by studying the effect of complete and partial elimination of up to four competing risks on the overall survival function, the life expectancy and life annuity values. We further investigate how different choices of the copula function affect the resulting joint distribution of survival times and in particular the actuarial functions which are of importance in pricing life insurance and annuity products. For illustrative purposes, we have used a real data set and used extrapolation to prepare a complete multiple decrement model up to age 120. Extensive numerical results illustrate the sensitivity of the model with respect to the choice of copula and its parameter(s).

Keywords: dependent competing risk model, disease elimination, failure times, overall survival function, copulas, spline function

*Corresponding author's address: Faculty of Actuarial Science and Statistics, Cass Business School, City University, London EC1Y 8TZ, UK. E-mail address: V.Kaishev@city.ac.uk, tel: +44(0)2070408453

1. Introduction

The competing risks model with independent failure time random variables has been considered by a number of authors in the (bio)statistical, econometric, medical, demographic and actuarial literature and the list of references, scattered throughout these areas, is extensive. We will mention here the textbooks due to David and Moeschberger (1978), Elandt-Johnson and Johnson (1980) and Bowers et al. (1997) and also the recent papers by Salinas-Torres et al. (2002) and Bryant and Dignam (2004), and also by Zheng and Klein (1994, 1995) where statistical methods for estimating related survival functions are considered.

The competing risk model, alternatively referred to as multiple decrement model, has also been considered under the assumption of dependence between the failure times in the early work of Elandt-Johnson (1976). Later, Yashin, Manton and Stallard (1986) considered conditional independence of the times to death, given an assumed stochastic covariate process. More recently, Carriere (1994) and Escarela and Carriere (2003) modelled dependence between two failure times by a two dimensional copula. Carriere (1994) has used a bivariate Gaussian copula to model the effect of complete elimination of one of two competing causes of death on human mortality. However, the mortality data used by Carriere (1994) was not complete with respect to older ages and therefore, it was not possible to calculate such important survival characteristics as expected lifetimes and life annuities and draw relevant conclusions. In Escarela and Carriere (2003), the bivariate Frank copula was fitted to a prostate cancer data set. The issues of identifiability of marginal survival functions in a copula based competing risk model have been considered by Tsiatis (1975), Prentice et al. (1978), Heckman and Honore (1989) and later by Carriere (1994).

In this paper, we will consider further the copula-based competing risk model, studied by Carriere (1994). We will investigate its sensitivity with respect to alternative choices of bivariate copula and its parameter(s). For this purpose, we have closed the survival model by applying a method of spline extrapolation up to a limiting age 120 and have explored the Gaussian copula, the Student t -copula, the Frank copula and the Plackett copula as alternatives. As discussed in Section 3, these copulas allow for modelling the dependence between failure times within the entire age range, from perfectly negative, to perfectly positive dependence. They belong to different families with different properties, and hence are appropriate for studying the sensitivity of the model. We develop this methodology, so as to model the effect of both partial and complete elimination of a cause of death on human mortality. The construction of multiple decrement tables, derived from the multivariate competing risk model is also addressed.

Since most real life applications are truly multivariate, i.e., there are more than two mutually dependent competing causes of decrement, our further goal here will be to extend and explore the applicability of the model to the multi dimensional case. This is in general a difficult task, since the bivariate copula theory does not extend to the multivariate case in a direct way. Although some fundamental results (e.g. Sklar's theorem) hold, constructing multivariate copula is related to some open problems, e.g., there is no unique multivariate dependence measure which extends the (bivariate) definitions of Kendall's τ and Spearman's ρ_s and the computational complexity increases. This makes multivariate copula applications less appealing. Here we have explored the applicability of the four dimensional Gaussian, t - and Frank copula to model the joint distribution of four competing risks, heart diseases, cancer, respiratory diseases and other causes of death, grouped together. The effect of simultaneously removing one, two or three of them on the overall survival, on the life expectancy at birth and at age 65, and on the value of a life annuity, which are important in pricing life insurance products, is also studied.

In the next section, we introduce the dependent multiple decrement model and the related crude, net and overall survival functions. Section 3 is devoted to copulas and their properties, and provides background material on the Gaussian, the Student t -, the Frank and the Plackett copula, which we have used to imple-

ment the proposed methodology. In Section 4, we address the problem of selecting an appropriate copula function and estimating its parameter(s). Then, in Section 5, we describe how, given some estimates of the crude survival functions and an appropriately selected copula, one can evaluate the net survival functions for the corresponding competing risks, by solving a system of nonlinear differential equations. In Section 6, we show how, by introducing an appropriate function, one can modify the net survival functions, obtained as solutions of the system of differential equations, so as to model not only complete but also partial elimination of any of the causes of death in the model. Finally, in Section 7 the proposed methodology is applied to the general US population, using cause specific mortality data set, provided by the National Center for Health Statistics(NCSH) (1999). Extensive numerical results and graphs illustrate the effect on survival of complete (partial) elimination of cancer, as a cause of death, in a two dimensional decrement model, and the elimination of any combination of heart diseases, cancer and respiratory diseases in a four variate model. Details of how the raw mortality data was used to obtain the crude survival functions and the method of smoothing and extrapolating the latter up to age 120 are provided in the Appendix.

2. The dependent multiple decrement model

We consider a group of lives, exposed to m competing causes of death, i.e., to m causes of withdrawal from the group. It is assumed that each individual may die from any single one of the m causes. To make the problem more formally (mathematically) tractable it is assumed that, at birth, each individual is assigned a vector of times $T_1, \dots, T_m$, $0 \leq T_j < \infty$, $j = 1, \dots, m$, representing his/her potential lifetime, if he/she were to die from each one of the m causes. Obviously, the actual lifetime span is the minimum of all the $T_1, \dots, T_m$. Thus, it is clear that under this model the lifetimes $T_1, \dots, T_m$ are unobservable, and we can only observe the $\min(T_1, \dots, T_m)$. In the classical multiple decrement theory the random variables $T_1, \dots, T_m$ are assumed independent, whereas here we will be interested in their joint distribution

$$F(t_1, \dots, t_m) = \Pr(T_1 \leq t_1, \dots, T_m \leq t_m) \quad (1)$$

and more precisely in the multivariate, joint survival function

$$S(t_1, \dots, t_m) = \Pr(T_1 > t_1, \dots, T_m > t_m) \quad (2)$$

which is considered absolutely continuous and where $t_j \geq 0$, for $j = 1, \dots, m$.

As pointed out by a number of authors (see e.g. Hooker and Longley-Cook (1957), Carriere (1994), Valdez (2001)), decrements in many real life actuarial applications tend to be dependent and hence, the random variables $T_1, \dots, T_m$ will be considered stochastically dependent and also non-defective, i.e., $\Pr(T_j < \infty) = 1$.

The overall survival of an individual, under our model assumptions, is defined by the random variable $\min(T_1, T_2, \dots, T_m)$, and so we will be interested in the overall survival function, which we define as $\mathcal{S}(t) \equiv S(t, \dots, t) = \Pr(T_1 > t, \dots, T_m > t) = \Pr(\min(T_1, \dots, T_m) > t)$, where $t \geq 0$. In particular, we are interested in the effect of elimination of a cause of death, say indexed j , on the overall survival function. We follow the approach to this elimination problem, taken by Elandt-Johnson (1976) and also by Carriere (1994), and assume that the removal of the j -th cause is equivalent to considering the marginal distribution

$$F(t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_m) = \Pr(T_1 \leq t_1, \dots, T_{j-1} \leq t_{j-1}, T_{j+1} \leq t_{j+1}, \dots, T_m \leq t_m).$$

The overall survival function then becomes

$$\Pr(T_1 > t, \dots, T_{j-1} > t, T_{j+1} > t, \dots, T_m > t) = \Pr(\min(T_1, \dots, T_{j-1}, T_{j+1}, \dots, T_m) > t), \quad (3)$$

i.e., we are considering the overall survival function with the j -th cause of death removed, and we will simply denote this by $\mathcal{S}^{(-j)}(t) \equiv S(t, \dots, t, 0, t, \dots, t)$, where $t = 0$ appears on the j -th position. Alterna-

tively, as pointed out by Elandt-Johnson (1976), eliminating a cause of death, may be interpreted as considering the limiting distribution (conditional on T_j) of surviving from all other causes, given $T_j \rightarrow \infty$. However, this approach is confronted with the difficulty that this limiting conditional distribution may not exist and even if it does exist, the evaluation of the overall survival function is more complicated. It has to be mentioned that in the case when $T_1, \dots, T_m$ are assumed independent the two approaches are equivalent (see Elandt-Johnson 1976). Thus, our major goal in this investigation will be to find a representation of the survival function $S^{(-j)}(t)$ which will allow us to estimate it and thereby measure the effect of removing a cause of death. As we will see in the following sections, one such representation is through the use of a suitable copula function.

Let us now introduce the notion of the *crude* survival function. The crude survival function $S^{(j)}(t)$ is defined as the survival function with respect to the j -th cause of death, due to which death actually occurs, i.e.,

$$S^{(j)}(t) = \Pr(\min(T_1, \dots, T_m) > t, \min(T_1, \dots, T_m) = T_j) \quad (4)$$

The survival function $S^{(j)}(t)$ is called *crude*, since it reflects the observed mortality of an individual and hence, may be estimated, from the observed mortality data of a population, as will be illustrated in Section 3. In the biostatistics literature the crude survival function $S^{(j)}(t)$ is sometimes called the *subsurvival function* or the *cumulative incidence function* (see e.g. Bryant and Dignam, 2004).

It is not difficult to see that

$$S(t, \dots, t) = S^{(1)}(t) + \dots + S^{(m)}(t), \quad (5)$$

since the events $\min(T_1, \dots, T_m) = T_j, j = 1, \dots, m$ are mutually exclusive.

This obviously suggests that the distributions in (4) are degenerate and that the crude survival functions are such that $S^{(j)}(0) < 1, j = 1, \dots, m$.

Let us now define the *net* survival function $S'^{(j)}(t)$ as $S'^{(j)}(t) = \Pr(T_j > t)$. Note that $S'^{(j)}(t)$ is the marginal survival function, due to cause j alone, associated with the joint multivariate survival function (2). Thus, we can view (1) as a multivariate distribution with marginal distribution functions $F'^{(j)}(t) = 1 - S'^{(j)}(t), j = 1, \dots, m$. As we will see, $S'^{(j)}(t)$ are the target quantities in our study since, if we know them we can identify and calculate the joint survival function $S(t_1, t_2, \dots, t_m)$ and hence, evaluate the overall survival function $S(t, \dots, t)$, under some appropriate assumptions. The classical model of independence of the r.v.s $T_1, T_2, \dots, T_m$ implies that

$$S(t_1, \dots, t_m) = S'^{(1)}(t_1) \times \dots \times S'^{(m)}(t_m).$$

When we consider the case of stochastic dependence between the causes of death, apart from knowing the marginals $F'^{(j)}(t)$ we will need to impose a certain dependence structure, in order to characterize the joint distribution of the r.v.s $T_1, T_2, \dots, T_m$. One way of doing so is to use copulas. Thus, to obtain the joint survival function $S(t_1, \dots, t_m)$, one would need to select a suitable copula, which mixes (couples) the net survival functions $S'^{(j)}(t), j = 1, 2, \dots, m$. However, in our model, $S'^{(j)}(t)$ are not known and we can only estimate the crude survival functions, $S^{(j)}(t), j = 1, 2, \dots, m$, based on an appropriate m -dimensional, cause specific mortality table. As an illustration of the model, in Section 7, we have used such a table constructed from the US Decennial Life Tables for 1989-91 published by NCHS (1999). In order to obtain estimates of the net survival functions $S'^{(j)}(t)$ on the basis of the estimated crude survival functions, a system of m non-linear differential equations, which links the two sets of functions, through the partial copula derivatives, is solved.

Copulas have recently attracted considerable attention as a tool for modelling dependence in a wide range of applications in finance, insurance, and economics. In Section 3, we introduce the required background

material on copulas and recall the Gaussian copula and the t -copula from the Elliptical class of copulas, the Frank copula from the Archimedean family, and also the Plackett copula. We provide also some general comments on the choice of copula, appropriate for our modelling purposes. For further details on copulas, the interested reader is referred to Georges et al. (2001), Frees and Valdez (1998), and Embrechts et al. (2001), and the books by Nelsen (1999), Joe (1997) and Cherubini et al. (2004).

3. Copulas and their properties

Copulas provide a very convenient way to model and measure dependence between failure time random variables since they give the dependence structure which relates the known marginal distributions of the failure times to their multivariate joint distribution. In order to see this, we first provide a short introduction on copulas.

If we assume that $\mathbf{u} = (u_1, \dots, u_m)'$, $u_j \in [0, 1]$, an m copula $C(\mathbf{u})$ is conventionally defined as a multivariate cumulative distribution function with uniform margins. A probabilistic way to define the copula is provided by the theorem of Sklar (1959). Let $X_1, \dots, X_m$ be random variables with continuous distribution functions $F_1, \dots, F_m$, and survival functions $S_1 = 1 - F_1, \dots, S_m = 1 - F_m$ respectively, and joint distribution and survival functions $H(x_1, \dots, x_m)$ and $S(x_1, \dots, x_m)$. Sklar's theorem states that if H is an m -dimensional d.f. of the random vector $(X_1, \dots, X_m)$ with continuous marginals $F_1, \dots, F_m$, then there exist a unique m copula C , such that for all $\mathbf{x}$ in $\mathbb{R}^m$

$$H(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m)) \quad (6)$$

and conversely, if C is an m copula and $F_1, \dots, F_m$ are d.f.s, then H is an m -dimensional distribution function with marginals $F_1, \dots, F_m$. Hence, the copula of the random vector $(X_1, \dots, X_m) \sim H$ is the distribution function of the random vector $(F_1(X_1), \dots, F_m(X_m))$. Thus, following (6), one can construct a dependence structure, i.e., an m -dimensional d.f. H by appropriately choosing a set of marginals $F_1, \dots, F_m$ and a copula function C . In order to construct a copula function, a corollary of Sklar's theorem can be applied, according to which a copula can be represented as an m -dimensional distribution function with continuous marginals, evaluated at the inverse functions $F_1^{-1}(u_1), \dots, F_m^{-1}(u_m)$, defined as $F_i[F_i^{-1}(x_i)] = x_i$, i.e.,

$$C(u_1, \dots, u_m) = H(F_1^{-1}(u_1), \dots, F_m^{-1}(u_m)) . \quad (7)$$

By using the probability integral transformation, $X_j \mapsto F_j(X_j) = 1 - S_j(X_j)$ as a result of which $S_j(X_j)$ has a uniform distribution on $[0, 1]$, it is easily verified (see Sklar, 1996) that Sklar's theorem, given by (6) can be restated to express the multivariate survival function $S(x_1, \dots, x_m)$ via an appropriate copula $\bar{C}$ called the survival copula of $(X_1, \dots, X_m)$. Thus,

$$S(x_1, \dots, x_m) = \bar{C}(S_1(x_1), \dots, S_m(x_m)). \quad (8)$$

The survival copula, relates the marginal survival functions $S_1(x_1), \dots, S_m(x_m)$ to the multivariate joint survival function $S(t_1, \dots, t_m)$ in much the same way as the copula C relates the marginal distribution functions to the multivariate distribution function.

Let us note that the copula C and the survival copula $\bar{C}$ of a random vector $(X_1, \dots, X_m)$ do not in general coincide. In order to see how the survival copula is expressed through its corresponding copula, we refer to Nelsen (1999) and Georges et al. (2001). The survival copula, $\bar{C}$, in (8) can be expressed through its copula, C , derived on the basis of specific distributions $F_1, \dots, F_m$, using (7). However, since $\bar{C}(\mathbf{u})$ is a copula, one can directly use any copula $C(\mathbf{u})$ to link the joint survival function in (8) to its marginals $S_1(x_1), \dots, S_m(x_m)$. This will be the approach taken here in modelling the joint survival function of competing risk survival times.

We conclude this description of the general background on copulas by recalling the fundamental Fréchet-Hoeffding bounds inequality which holds with respect to every m -dimensional copula function. To capture its importance we will first give its bivariate version, i.e.,

$$\max(u_1 + u_2 - 1, 0) \leq C(u_1, u_2) \leq \min(u_1, u_2) . \quad (9)$$

where the maximum and the minimum in (9) are correspondingly the lower and upper Fréchet-Hoeffding bounds which themselves are copulas. Following (9), in order that the random variables X_1, X_2 with a copula function $C(u_1, u_2)$ be perfectly negatively (positively) dependent, $C(u_1, u_2)$ needs to coincide with the lower (upper) Fréchet-Hoeffding bound. The r.v.s X_1, X_2 are independent iff their copula is equal to the product copula, i.e., iff $C(u_1, u_2) = u_1 u_2$. It can then be shown that (9) generalizes to the multivariate case as

$$\max(u_1 + \dots + u_m - m + 1, 0) \leq C(u_1, \dots, u_m) \leq \min(u_1, \dots, u_m) .$$

Many popular families of copulas depend on a set of parameters, not related to the parameters of the marginal distributions. This is due to the fact that copulas are invariant under increasing transformations of their corresponding random variables, hence are "scale-invariant". In order to model the full range of dependence, from perfect negative dependence, through independence, to perfect positive dependence, the copula parameters should be set so that the copula $C(u_1, u_2)$ attains correspondingly, its lower Fréchet-Hoeffding bound, coincides with the product copula, and attains its upper Fréchet-Hoeffding bound. Copulas which allow this are called *comprehensive* copulas (see Deheuvels, 1978). Since it is meaningful to investigate the full range of dependence between survival time r.v.s, it is desirable to use comprehensive copulas in modelling the dependence of competing causes of death. So, this is one of the important criteria in selecting the appropriate copula, i.e., it should be able to reproduce dependence throughout the whole range.

Sklar's theorem and its corollary, given by (7), provide a convenient tool for constructing copulas. Examples of such copulas are the multivariate Gaussian and Student t - copulas which belong to the wider class of Elliptical copulas. However, there are other ways of constructing copulas. For example, the popular Archimedean copulas are constructed as

$$C^A(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)) . \quad (10)$$

where ϕ is a continuous, convex function called a generator, such that $\phi(1) = 0$ and $\phi(0) = +\infty$ (see e.g. Nelsen, 1999). In what follows, we will introduce and use the Frank copula as one of the only two known comprehensive, bivariate Archimedean copulas. In addition, we will also use the Plackett copula, which does not belong to the Elliptical or to the Archimedean family but is comprehensive and hence, suitable for use in the competing risk model. For these and other properties of Frank and Plackett copulas we refer to Section 3.4 and also to Nelsen (1999).

3.1 Measures of association

We will consider here the standard dependence measures, Kendall's τ and Spearman's ρ_S , and the tail dependency as measures of association between two random variables X and Y . These measure are related to the copula since the latter is an expression of the stochastic relationship between X and Y within the entire range of values which the variables can take. It is not difficult to show that

$$\rho_S(X, Y) = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3 \quad (11)$$

and that the

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 1 \quad (12)$$

(see e.g. Cherubini et al., 2004), i.e., knowing the copula $C(u_1, u_2)$ of a pair of r.v.s X and Y one can evaluate these two measures of rank correlation using (11) and (12). For further properties of ρ_S and τ we refer e.g., to Nelsen (2001).

Another, important measure of dependence, which can be described as a measure of concordance in the tails of two r.v.s X and Y , is given by the upper and lower tail dependence coefficients

$$\lambda_L(X, Y) = \lim_{p \rightarrow 0} \Pr(Y \leq y_p \mid X \leq x_p).$$

$$\lambda_U(X, Y) = \lim_{p \rightarrow 1} \Pr(Y > y_p \mid X > x_p).$$

where x_p and y_p denote the lower p -quantiles of X and Y , i.e., $\Pr(X \leq x_p) = p$ and $\Pr(Y \leq y_p) = p$.

3.2 The Gaussian copula

We will first introduce the Gaussian copula. Let the random variables $X_1, \dots, X_m$ be standard normally distributed, i.e., $X_i \sim N(0, 1)$, $i = 1, \dots, m$ and let also the random vector $(X_1, \dots, X_m)$ have a standard m -variate normal distribution with correlation matrix $\mathbf{R} = \{R_{ij}\}$, $i, j = 1, \dots, m$. Clearly, R_{ij} are the linear correlation coefficients of the corresponding bivariate normal distributions, i.e., $R_{ij} = \rho(X_i, X_j) = \text{Cov}(X_i, X_j) / (\sqrt{\text{Var}(X_i)} \sqrt{\text{Var}(X_j)}) = \text{Cov}(X_i, X_j)$ in this case.

For $\mathbf{x} = (x_1, \dots, x_m)' \in \mathbb{R}^m$, denote by $\phi_R(x_1, \dots, x_m)$ the joint density function of the random vector $(X_1, \dots, X_m)$, i.e.,

$$\phi_R(x_1, \dots, x_m) = (2\pi)^{-m/2} |\mathbf{R}|^{-1/2} \exp \left\{ \frac{-\mathbf{x}' \mathbf{R}^{-1} \mathbf{x}}{2} \right\}$$

and denote also by $\Phi_R(x_1, \dots, x_m)$ the joint distribution function of $(X_1, \dots, X_m)$. Then, following (6) and (7), the copula $C^{\text{Ga}}(u_1, \dots, u_m)$ of the m -variate random vector $(X_1, \dots, X_m) \sim \Phi_R(x_1, \dots, x_m)$ may be defined as the distribution function of the random vector $(\Phi(X_1), \dots, \Phi(X_m))$, and is given by

$$C^{\text{Ga}}(u_1, \dots, u_m) = \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m)) = \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_m)} \phi_R(x_1, \dots, x_m) dx_1 \dots dx_m.$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard univariate normal distribution function. It is easy to see, after direct differentiation of (8) with respect to $u_1, \dots, u_m$, that the density function of the Gaussian copula is

$$c^{\text{Ga}}(u_1, \dots, u_m) = \frac{\partial^m C^{\text{Ga}}(u_1, \dots, u_m)}{\partial u_1 \dots \partial u_m} = \frac{\phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m))}{\phi(\Phi^{-1}(u_1)) \times \dots \times \phi(\Phi^{-1}(u_m))}.$$

In the two dimensional case, i.e., when $m = 2$, the matrix $\mathbf{R}$ is a 2×2 symmetric matrix with diagonal elements equal to 1 and an off-diagonal entry R_{12} , which completely defines the Gaussian copula in (8). For the Gaussian copula $\lambda_L(X, Y) = \lambda_U(X, Y) = 0$, i.e., there is no tail dependence and regardless of how high the value of ρ is, asymptotically extreme events in X and Y tend to occur independently.

3.3 The multivariate Student's t -copula

The multivariate Student's t -copula is defined through the multivariate t -distribution as follows. For $\mathbf{x} = (x_1, \dots, x_m)' \in \mathbb{R}^m$, denote by $t_{R,\nu}(x_1, \dots, x_m)$ the standardized multivariate joint t -distribution function with ν degrees of freedom and correlation matrix $\mathbf{R}$, i.e.,

$$t_{R,\nu}(x_1, \dots, x_m) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} \frac{\Gamma(\frac{\nu+m}{2}) |\mathbf{R}|^{-1/2}}{\Gamma(\frac{\nu}{2}) (\nu\pi)^{\frac{m}{2}}} \left(1 + \frac{1}{\nu} \boldsymbol{\xi}' \mathbf{R}^{-1} \boldsymbol{\xi}\right)^{-\frac{\nu+m}{2}} d\xi_1 \cdots d\xi_m.$$

Then, the multivariate Student's t -copula is defined through the multivariate t -distribution as

$$C^T(u_1, \dots, u_m) = t_{R,\nu}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_m)) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_m)} \frac{\Gamma(\frac{\nu+m}{2}) |\mathbf{R}|^{-1/2}}{\Gamma(\frac{\nu}{2}) (\nu\pi)^{\frac{m}{2}}} \left(1 + \frac{1}{\nu} \boldsymbol{\xi}' \mathbf{R}^{-1} \boldsymbol{\xi}\right)^{-\frac{\nu+m}{2}} d\xi_1 \cdots d\xi_m$$

where $t_\nu^{-1}(u_i)$, $i = 1, \dots, m$ is the inverse of the distribution function of a univariate t -distribution with ν degrees of freedom. The density of the Student's t -copula is

$$c^T(u_1, \dots, u_m) = |\mathbf{R}|^{-1/2} \frac{\Gamma(\frac{\nu+m}{2})}{\Gamma(\frac{\nu}{2})} \left(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right)^m \frac{\left(1 + \frac{1}{\nu} \boldsymbol{\xi}' \mathbf{R}^{-1} \boldsymbol{\xi}\right)^{-\frac{\nu+m}{2}}}{\prod_{j=1}^m \left(1 + \frac{\xi_j^2}{\nu}\right)^{-\frac{\nu+1}{2}}}.$$

where $\boldsymbol{\xi} = (t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_m))'$. Similarly to the Gaussian copula, the t -copula is symmetric but it has upper and lower tail dependence which, in the bivariate case, is given by

$$\lambda_L(X, Y) = \lambda_U(X, Y) = 2 t_{\nu+1} \left(-\sqrt{\nu+1} \sqrt{\frac{1 - \rho(X, Y)}{1 + \rho(X, Y)}} \right).$$

For the t -copula $\lambda_L(X, Y) = \lambda_U(X, Y) > 0$, i.e., there is tail dependence and it becomes stronger with the decrease of the degrees of freedom, ν , and/or with the increase of the linear correlation ρ . The existence of tail dependence means that asymptotically extremely large or extremely small events in X and Y tend to occur simultaneously.

Note that the Gaussian and the t -copulas, defined in Sections 3.2 and 3.3, depend on the correlation matrix $\mathbf{R}$, i.e., they depend on the pairwise linear correlation coefficients R_{ij} , $i, j = 1, \dots, m$ which, in general, are unknown parameters. The latter may be expressed in terms of Kendall's τ in the following form

$$R_{ij} = \sin(\pi \tau(X_i, X_j) / 2), i, j = 1, \dots, m, i \neq j. \quad (13)$$

3.4 Frank and Plackett Copulas

The Frank copula is an Archimedean copula, defined by (10) with generator $\phi(t) = -\ln[(e^{-\theta t} - 1)/(e^{-\theta} - 1)]$, $\theta \in (-\infty, \infty) \setminus \{0\}$, i.e.,

$$C^F(u_1, u_2) = (-1/\theta) \ln \{[(1 - e^{-\theta}) - (1 - e^{-\theta u_1})(1 - e^{-\theta u_2})]/(1 - e^{-\theta})\},$$

and its density is given as

$$c^F(u_1, u_2) = [\theta (1 - e^{-\theta}) e^{-\theta(u_1+u_2)}] / [(1 - e^{-\theta}) - (1 - e^{-\theta u_1})(1 - e^{-\theta u_2})]^2.$$

Another copula, which is neither Archimedean nor Elliptical, since it is constructed from Plackett's family of distributions on applying Sklar's theorem, is the Plackett copula, defined as

$$C^P(u_1, u_2) = \left\{ [1 + (\theta - 1)(u_1 + u_2)] - \sqrt{[1 + (\theta - 1)(u_1 + u_2)]^2 - 4u_1 u_2 \theta(\theta - 1)} \right\} / [2(\theta - 1)],$$

where $\theta \geq 0$, $\theta \neq 1$. If $\theta = 1$ then $C^P(u_1, u_2) = u_1 u_2$. The density of the Plackett copula is given as

$$c^P(u_1, u_2) = \theta [1 + (\theta - 1)(u_1 + u_2 - 2u_1 u_2)] / \sqrt{([1 + (\theta - 1)(u_1 + u_2)]^2 - 4u_1 u_2 \theta(\theta - 1))^3}$$

if $\theta \geq 0$, $\theta \neq 1$ and $c^P(u_1, u_2) = 1$ if $\theta = 1$. Both of the Frank and Plackett copulas are comprehensive and hence are appropriate in our modelling framework. Neither of them has tail dependence.

We will also use the straightforward multivariate generalization of the bivariate Frank copula

$$C^F(u_1, \dots, u_n) = (-1/\theta) \ln \left\{ 1 + [(e^{-\theta u_1} - 1) \dots (e^{-\theta u_n} - 1)] / (e^{-\theta} - 1)^{n-1} \right\},$$

which still depends on one parameter $\theta > 0$, but does not allow for modelling negative dependence (see Nelsen, 1999).

4 Estimating the model parameters

As noted in Section 2, in order to introduce and evaluate the functions of interest, arising from the competing risk model, one needs to specify a suitable copula, define its parameters and provide estimates of the crude survival functions, based on an appropriate multiple, cause specific mortality table. This will be discussed in somewhat greater detail in this section.

4.1 Specifying the copula and its parameters

Let us consider the bivariate case. As mentioned earlier, the association between the survival times T_1, T_2 , related to the two competing causes of death may in general vary, from extreme positive to extreme negative dependence. In order to capture such variation, one needs copulas whose parameters can be varied so that Kendall's τ and Spearman's ρ_s , expressed through the copula, by (11) and (12), take the whole range of values from -1 to 1 . Thus, it is desirable to use comprehensive copulas in modelling dependence of competing causes of death. Applying this criterion, we have selected the four copulas, introduced in Section 3, whose basic properties are summarized in Table 1. Here C^- , $C_\perp$ and C^+ denote the lower Fréchet-Hoeffding bound, the product copula and the upper Fréchet-Hoeffding bound respectively.

Table 1. Selected bivariate comprehensive copulas: main properties.

Copula	Comprehensive	Tail Dep.	Param. range	C^-	$C_\perp$	C^+
Gaussian(ρ)	Yes	No	$\rho \in (-1, 1)$	$\rho = -1$	$\rho = 0$	$\rho = 1$
t-copula(ρ, ν)	Yes	Yes	$\rho \in (-1, 1), \nu > 2$	$\rho = -1$	$\nu = \infty$	$\rho = 1$
Frank(θ)	Yes	No	$\theta \in (-\infty, +\infty) \setminus \{0\}$	$\theta = -\infty$	$\theta = 0$	$\theta = +\infty$
Plackett(θ)	Yes	No	$\theta \in (0, \infty) \setminus \{1\}$	$\theta = 0$	$\theta = 1$	$\theta = +\infty$

As seen from Table 1, the selected copulas depend on one or two parameters, not related to the parameters of the marginal distributions. The copula parameters, need to be estimated, based on a set of N pairwise observations on the survival times T_1 and T_2 . However, there is only one observable survival time, i.e., the $\min(T_1, T_2)$ and it is impossible to estimate the copula parameters, and hence to identify the joint survival function $S(t_1, \dots, t_m)$, by simply knowing the net survival functions $S^{(j)}(t)$, $j = 1, \dots, m$. For further

discussion on the identifiability of competing risk models we refer to Yashin, Manton and Stallard (1986), and Carriere (1994), from where related references can be followed up.

The fact that estimation of the copula parameter(s) is not possible imposes the necessity of considering them as free parameters. They could be set according to a priori available (medical) information, about the degree of pairwise dependence between the two competing risks, expressed through Kendall's τ and/or Spearman's ρ_S . In what follows, we have set the copula parameters equal to fixed values, covering the whole range of possible values of τ and ρ_S . Here, we provide a more detailed discussion on how this is done for the example of the Gaussian copula case as an illustration.

As pointed out in Section 3.2, R_{12} can be expressed through the Kendall's $\tau(T_1, T_2)$, as given by (13), which in this case is a free parameter, and hence, the dependence structure will have to be studied for different, fixed values of $\tau(T_1, T_2)$ within the open interval $(-1, 1)$. Let us note that $\tau(T_1, T_2)$ can not take the boundary values -1 and 1 , since in this case, the copula in (8) is not differentiable. Obviously, by fixing $\tau(T_1, T_2)$ we assume a certain degree of association (correlation) of the two r.v.s, T_1 and T_2 , as given in terms of the Kendall's τ . This means that, when constructing the joint survival function of T_1 and T_2 , we assume a certain degree of dependence between the two causes of death, which we believe is realistic to admit. Thus, if $\tau(T_1, T_2) = 0.99$, we assume a strong positive correlation between the two causes of death. This would mean that, removal of one of the causes from our model will not significantly affect, i.e. improve the overall survival, since, due to their positive correlation the remaining cause will operate similarly as the removed one. If $\tau(T_1, T_2) = 0$ then we are in the usual independence assumption. In the other extreme, $\tau(T_1, T_2) = -0.99$ would mean that removal of one of the causes will significantly affect the overall survival, since due to the negative correlation, we assume that the remaining cause has an opposite effect on mortality to the removed one, and hence can not compensate for its removal.

All of the considerations made here for the two dimensional case carry over to the multivariate case, noting however, that in this case there is not a unique definition of an association measure, equivalent to Kendall's τ or Spearman's ρ_S . One has to bear in mind, of course, that assumptions have to be made about all pairwise coefficients $\tau(T_i, T_j)$, $i, j = 1, \dots, m$, $i \neq j$ which, for large values of m , will require careful consideration (for $m = 2$ one numerical value for τ would need to be specified, but for $m = 3, 4, 5$ and 6 one would need to specify correspondingly 3, 6, 10 and 15 entries for the corresponding correlation matrix $\mathbf{R}$). So, one can conclude that the best approach to defining the values of τ is to set them, based on medical or expert knowledge about the level of pairwise dependence between diseases. Obviously, some of the diseases may be expected to be mutually pairwise independent so, the corresponding τ values can be set equal to zero.

Let us note that, in order to model extreme positive and negative dependence, i.e., when τ takes values 1 and -1 , it is more convenient to use the Frank copula.

4.2 Estimating and extrapolating the crude survival functions

As noted in Section 2, it is possible to estimate the crude survival functions, based on an appropriate m dimensional set of cause specific, mortality data. If the data for each cause have already been smoothed they may then be directly interpolated, using spline interpolation or another alternative approximation method (see e.g., Dellaportas et. al, 2001). Here, we propose using the 'data averaging' optimal spline interpolation method described by De Boor (2001). Applying this method, m spline models may be fitted to the sets of observed values for the m crude survival functions and smooth estimates of the crude survival functions, $S^{(j)}(t)$ $j = 1, \dots, m$, could then be obtained.

An important point to make in this connection is that often the multiple cause of death mortality data available may not be "closed", i.e., the data source may not include complete data for the older ages, beyond, say, 100 years of age, so that the final age group is "100 years and older". This makes it impossible to

produce complete net and overall survival functions and to estimate life expectancy and other actuarial functions, under the dependent competing risk model, and to investigate the complete or partial elimination of one or more causes of death. There have been several approaches to extending life tables beyond an upper age limit, such as the old-age mortality standard, developed by Himes, Preston and Condran (1994), the old age part of the Heligman-Pollard mortality model (see Heligman-Pollard, 1980) and also the Coale-Kisker method of closure of mortality tables (see Coale and Kisker, 1990, developed from a previous paper by Coale and Guo, 1989, and a recent investigation by Renshaw and Haberman, 2003). For a more detailed account on these methods and their application we refer to Buettner (2004). Here, we have applied a different approach (described in the Appendix), based on an appropriate extrapolation of the spline interpolants, produced by the 'data averaging' method. As will be seen from our numerical results (see Section 7), and based on our experience of a number of examples, this method produces reasonable extrapolations of the mortality experience up to the oldest ages, in the range 100-120.

If the mortality data have not been smoothed, a (spline) smoothing method may be applied to the raw data. A good candidate for the purpose is the method of constructing 'geometrically designed' free-knot splines, called GeD splines, which is automatic and does not require any knowledge of the number of knots and their locations (see Kaishev, Dimitrova, Haberman and Verrall, 2004).

5. Evaluating the net and overall survival functions

Having fixed the copula function, $C(u_1, \dots, u_m)$, one may use (8) and evaluate the joint survival function

$$S(t_1, \dots, t_m) = C(S^{(1)}(t_1), \dots, S^{(m)}(t_m)) \quad (14)$$

if the net survival functions $S^{(j)}(t_j)$, $j = 1, \dots, m$ were known. In order to find them, we may use the relationship between $S^{(j)}(t)$ and the crude survival functions, $S^{(j)}(t)$, $j = 1, \dots, m$ given by Heckman and Honore (1989) and also by Carriere (1994). Thus, under the assumption of differentiability of $C(u_1, \dots, u_m)$ with respect to $u_j \in (0, 1)$ and of $S^{(j)}(t_j)$ with respect to $t_j > 0$, for $t > 0$, the following system of differential equations holds

$$\begin{aligned} \frac{d}{dt} S^{(1)}(t) &= C_1(S^{(1)}(t), \dots, S^{(m)}(t)) \times \frac{d}{dt} S^{(1)}(t) \\ \frac{d}{dt} S^{(2)}(t) &= C_2(S^{(1)}(t), \dots, S^{(m)}(t)) \times \frac{d}{dt} S^{(2)}(t) \\ &\vdots \\ \frac{d}{dt} S^{(m)}(t) &= C_m(S^{(1)}(t), \dots, S^{(m)}(t)) \times \frac{d}{dt} S^{(m)}(t), \end{aligned} \quad (15)$$

where $C_j(u_1, \dots, u_m) = \frac{\partial}{\partial u_j} C(u_1, \dots, u_m)$, $j = 1, \dots, m$.

It is important to note that (15) is a system of nonlinear, differential equations which may be solved with respect to the net survival functions $S^{(j)}(t)$, once we have specified the type of copula function, and given estimates of the crude survival functions $S^{(j)}(t)$, $j = 1, \dots, m$ in a suitable functional form, which can then be substituted into the left-hand side of (15). Then, in order to obtain the net survival functions $S^{(j)}(t)$, we need to find an efficient numerical method of solving the system (15). Once the net survival functions are obtained, we may substitute them into the copula on the right-hand side of (14) and evaluate the joint survival function, and in particular the overall survival function $\mathcal{S}(t) \equiv S(t, \dots, t)$, which is of major interest in our investigation.

The derivatives with respect to time of the crude and net survival functions in (15) are actually the crude and net probability density functions of the r.v.s $T_1, T_2, \dots, T_m$. We will denote these densities as $f^{(j)}(t)$ and $f'^{(j)}(t)$, $j = 1, \dots, m$, respectively.

In order to solve the system (15) numerically, one can rewrite it as a system of difference equations, assuming that the time variable $t \geq 0$ takes integer values, (i.e., integral age values) $k = 0, 1, 2, \dots$. This has been the approach taken by Carriere (1994). Its disadvantage is that the net survival functions $S'^{(j)}(t)$, $j = 1, \dots, m$ are obtained only for integral ages. Here, we have used the *Mathematica* built-in function *NDSolve*, which solves numerically systems of differential equations and produces solutions, $S'^{(j)}(t)$, $j = 1, \dots, m$, for any $t > 0$.

Let us also note that we have used equality (5) as a check on the solution of (15). For this purpose, we can apply (14) to express the overall survival function on the left-hand side of (5) as

$$C(S'^{(1)}(t), \dots, S'^{(m)}(t)) = S^{(1)}(t) + \dots + S^{(m)}(t), 0 \leq t \leq 120. \quad (16)$$

Equality (16) implies that if, for any fixed t , we substitute the values of the net survival functions $S'^{(1)}(t), \dots, S'^{(m)}(t)$, obtained as a solution of (15), in the copula function, its value must be equal to the sum of the m crude survival functions, evaluated at t .

6. Partial and complete disease elimination

In order to study the effect of partial and complete disease elimination, we have adopted the following approach. Let us recall that, in our model, we have assumed that $T_1, T_2, \dots, T_m$ are the future lifetime spans of a newborn individual, under the operation of m causes of death, i.e., all the survival functions, introduced up to now, refer to age zero. We will now need to adjust explicitly the adopted notation for the crude and net survival functions, by adding a 0 subscript, indicating age at birth, i.e., we will write $S_0^{(j)}(t)$, and $S_0'^{(j)}(t)$ instead of $S^{(j)}(t)$ and $S'^{(j)}(t)$, and also $S_x^{(j)}(t)$, and $S_x'^{(j)}(t)$ to denote the corresponding crude and net survival functions for a life aged x . In order to illustrate the concept of partial and complete disease elimination, we will assume here that t takes integral values, i.e., $t \equiv k = 1, 2, \dots, 120$. Thus, we can express the net survival functions $S_0'^{(j)}(t)$, $j = 1, \dots, m$, which we obtained as a solution of (15) as

$$S_0'^{(j)}(k) = S_0'^{(j)}(1) \times S_1'^{(j)}(1) \times \dots \times S_{k-1}'^{(j)}(1). \quad (17)$$

Applying the well known relation $S_x'^{(j)}(k) = S_0'^{(j)}(x+k) / S_0'^{(j)}(x)$ to the factors on the right-hand side of (17), we obtain the convenient identity

$$S_0'^{(j)}(k) = (S_0'^{(j)}(1) / S_0'^{(j)}(0)) \times (S_0'^{(j)}(2) / S_0'^{(j)}(1)) \times \dots \times (S_0'^{(j)}(k) / S_0'^{(j)}(k-1)) \quad (18)$$

Rewriting (18) in actuarial notation gives

$$S_0'^{(j)}(k) = p'_0 \times p'_1 \times \dots \times p'_{k-1} = (1 - q'_0) \times (1 - q'_1) \times \dots \times (1 - q'_{k-1}), \quad (19)$$

where, for simplicity, the dependence on the index j on the right-hand side has been suppressed.

We now introduce the piecewise linear function $\theta_l(a, b; c, d)$, with respect to $l = 0, 1, 2, \dots$, as

$$\begin{aligned} \theta_l(a, b; c, d) &= a, & \text{if } l \in [0, c] \\ &= a + \frac{b-a}{d-c} \times (l - c), & \text{if } l \in [c, d] \\ &= b, & \text{if } l \in [d, 120], \end{aligned} \quad (20)$$

where a, b, c, d , are parameters, such that $0 \leq a, b \leq 1$ and $0 < c < d < 120$.

We shall use $\theta_l(a, b; c, d)$ to modify the net survival function in (19), in order to be able to control the degree of elimination of the j -th disease. Thus, if we denote by $S_0^{(j)}(k)$ the modified net survival function and set $q''_l = (1 - \theta_l) q'_l$, $l = 0, 1, 2, \dots, k-1$, (19) can be rewritten as

$$S_0^{(j)}(k) = (1 - q''_0) \times (1 - q''_1) \times \dots \times (1 - q''_{k-1}).$$

By appropriately choosing particular values of the parameters a, b, c and d , for each age, the degree of elimination may be varied with the age l , with reasonable flexibility, from partial to complete elimination of the particular j -th cause of death. We may see how the modification of the net survival function (i.e., the degree of its elimination) will affect the overall survival function by substituting $S_0^{(j)}(k)$ back into the copula, which defines the joint survival function in (14), i.e., the overall survival function becomes

$$\mathcal{S}(k) \equiv \mathcal{S}(k, \dots, k) = C(S_0^{(1)}(k), \dots, S_0^{(j-1)}(k), S_0^{(j)}(k), S_0^{(j+1)}(k), \dots, S_0^{(m)}(k))$$

where $k = 0, 1, \dots, 120$. So, by varying the parameters a, b, c and d , one can adjust the degree of elimination of the j -th cause of death at different ages and study its effect on $\mathcal{S}(k)$ in the m -variate dependent multiple cause of death model. Let us note that the choice of $a = b = 1$ corresponds to $\theta_l \equiv 1$, for $l = 0, 1, 2, \dots, k-1$ which leads to $S_0^{(j)}(k) \equiv 1$, $k = 0, 1, 2, \dots$, which corresponds to complete elimination of the j -th disease, since obviously, the overall survival function with j -th disease removed is

$$\begin{aligned} \mathcal{S}^{(-j)}(k) &= \Pr(T_1 > k, \dots, T_{j-1} > k, T_{j+1} > k, \dots, T_m > k) = \\ &C(S_0^{(1)}(k), \dots, S_0^{(j-1)}(k), 1, S_0^{(j+1)}(k), \dots, S_0^{(m)}(k)). \end{aligned}$$

A numerical implementation of complete and partial elimination is presented in Section 7.

7. Numerical results

In this section, we apply the methodology, described earlier, to a real data set, related to the US female general population, in which the data are grouped by causes of death, using "Table 10. Number of life table deaths from specific causes during age interval for the female population: United States, 1989-91" of the U.S. Decennial Life Tables for 1989-91 (see NCHS, 1999). For ease of presentation, we consider the two dimensional and the multidimensional competing risk models separately.

7.1 Two causes of death

We consider here the simplest case of only two competing causes of death, one due to cancer, and a second one due to all other, non-cancer causes, pooled together. Thus, all the results of Sections 2-6 apply here for the case of $m = 2$. Denote by T_c and T_o the lifetime random variables for the cancer and non-cancer causes of death and by $S^{(c)}(t)$, $S^{(c)}(k)$, and $S^{(o)}(t)$, $S^{(o)}(k)$, the crude and net survival functions for cancer and non-cancer respectively. As noted in Section 4.2, it is possible to estimate crude survival functions based on an appropriate set of cause specific, mortality data. In order to estimate the crude survival functions for cancer and other (non-cancer) causes, we have used a two decrement data set, obtained on the basis of the multiple decrement "Table 10" from NCHS (1999). The data are presented in 5 year age intervals and cover the age range from 0 to 100+ years. It should be noted that the 'observed' values of the crude survival functions are, in fact, the values which have already been smoothed and published in NCHS (1999). For this reason, no further smoothing has been required and we have interpolated the 'observed' values of the crude survival functions for ages from 0 to 100, using the 'data averages' optimal spline interpolation method (see De Boor, 2001). In order to obtain a "closed" mortality model up to a limiting age of 120, we have extrapolated the fitted cubic spline functions $S^{(c)}(t)$ and $S^{(o)}(t)$, for the cancer and the other (non-cancer) causes, over the 100-120 age range, under the condition that $S^{(c)}(120) = S^{(o)}(120) = 0$. The extrapolation has been performed using female mortality data for the old ages 100-115 given in "Table 3" from NCHS (1997). The

actual numbers, as well as the method and formulae, used to obtain the 'observed' and extrapolated values of the crude survival functions are given in the Appendix.

The fitted cubic spline curves $S^{(c)}(t)$ and $S^{(o)}(t)$, $0 \leq t \leq 120$ and their derivatives are given in Figure 1. As can be seen, the spline models of the two crude survival functions and the corresponding densities are smooth and possess very good visual quality.

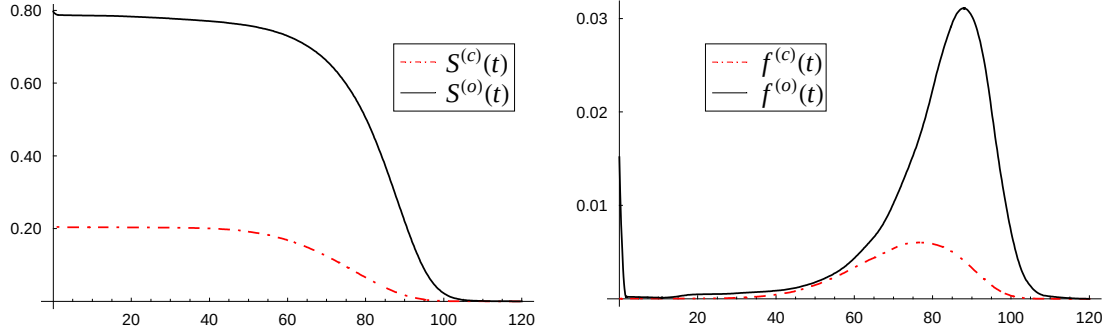


Fig. 1. Interpolated crude survival functions (left panel) and their densities (right panel) for 'cancer' and 'other' causes of death.

Having estimated the crude survival functions $S^{(c)}(t)$ and $S^{(o)}(t)$, $0 \leq t \leq 120$, we obtain the net survival functions $S'^{(c)}(t)$ and $S'^{(o)}(t)$, $0 \leq t \leq 120$, by solving the system (15), using the four copulas, specified in Table 1. The solutions $S'^{(c)}(t)$, $S'^{(o)}(t)$, $0 \leq t \leq 120$, obtained from (15) have been checked, using equation (16) for the case $m = 2$. Since the solutions $S'^{(o)}(t)$ and $S'^{(c)}(t)$ approach zero very closely in the age range $100 \leq t \leq 120$, solving (15) is much more difficult for $100 \leq t \leq 120$ and the solution there may not possess the built-in *Mathematica* precision. Another important point is that the numerical solution of (15), $S'^{(o)}(t)$ and $S'^{(c)}(t)$, is influenced by the extrapolated sections of the crude survival functions not only for $100 \leq t \leq 120$ but within the entire age range $0 \leq t \leq 120$. This in turn means that the results and conclusions with respect to survival under the dependent competing risk model, given later in this section, depend on the extrapolation that has been carried out.

7.1.1 The bivariate Gaussian copula case

The net survival functions, obtained as a solution of (15), using the Gaussian copula, $C^{\text{Ga}}(u_1, u_2)$, with values of ρ corresponding to five different values of Kendall's τ are plotted in Fig. 2 (so that $\tau = 0.91$ corresponds to $\rho = 0.99$, $\tau = 0.35$ corresponds to $\rho = 0.52$ and so on). As explained in Section 4.1, the linear correlation ρ is considered as a free parameter, by means of which different degrees of association, between the cancer and non-cancer modes of death, are preassigned. Thus, the system (15) has been solved for values of ρ equal to -0.99 , -0.52 , 0.00 , 0.52 , 0.99 and the obtained net survival functions $S'^{(o)}(t)$ and $S'^{(c)}(t)$, $0 \leq t \leq 120$, are represented by the curves, in the left and right panel in Fig. 2 (of course, other values of ρ could have been chosen). In the bivariate case, for fixed ρ , the net survival function, $S'^{(o)}(t)$, $0 \leq t \leq 120$, coincides with the overall survival function, $S^{(-c)}(t)$, $0 \leq t \leq 120$, when cancer has been removed. Obviously, if cancer is removed from the bivariate decrement model, the overall survival will entirely be determined by the only remaining cause of death, that of non-cancer, and vice-versa. Thus, in order to study the overall survival, when cancer is removed, we may directly study the non-cancer net survival function $S'^{(o)}(t)$.

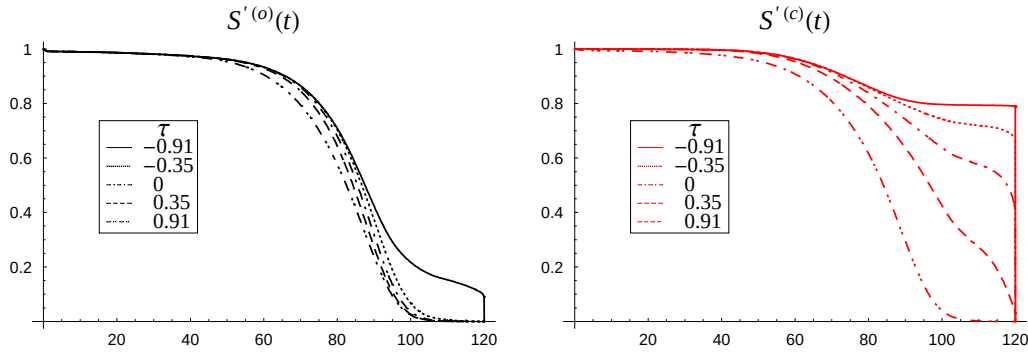


Fig. 2. The net survival functions $S'^{(o)}(t)$, $0 \leq t \leq 120$ for 'other' (non-cancer) cause (i.e. cancer removed) - left panel and for 'cancer' (i.e. 'other' cause removed) - right panel.

As can be seen from the left panel of Fig. 2, removing cancer affects survival most significantly when Kendall's $\tau = -0.91$ ($\rho_S = -0.99$), which corresponds to the case of extreme negative dependence. This effect of rectangularization of the net (overall) survival function is seen even more clearly on the right panel of Fig. 2, where the 'other' cause of death has been removed. In addition, we note that in the case of negative dependence or even independence between T_c and T_o , the trend of the overall survival curves suggests that the limiting age lies somewhere beyond 120 and it would not be natural to expect the old age survivors to die almost simultaneously at 120.

For the example of the bivariate Gaussian copula with Kendall's $\tau = -0.91$ ($\rho_S = -0.99$), we will illustrate how, not only a complete, but also a partial elimination of cancer, will affect the overall survival function $\mathcal{S}(t) \equiv S(t, t)$, $0 \leq t \leq 120$. As described in Section 6, we apply the function θ_t , given by (20) to modify the cancer net survival function $S'^{(c)}(t)$. The results are illustrated in the right panel of Fig. 3, where $S(t, t)$, $0 \leq t \leq 120$, is evaluated and plotted for five different choices of the function $\theta_t(a, b; c, d)$, plotted in the left panel of Fig. 3. As can be seen from Fig. 3, for any fixed age t , the probability of overall survival increases if we vary θ_t from $\theta_t \equiv 0$ ($a = b = 0$) - the solid curve, corresponding to no elimination, through $\theta_t \equiv 0.5$ ($a = b = 0.5$) - the dot-dashed curve, corresponding to half elimination equally applied for $0 < t < 120$, to $\theta_t \equiv 0.9$ ($a = b = 0.9$) - the double dot-dashed curve, corresponding to almost complete cancer elimination. $\theta_t(0.2, 0.8; 20, 65)$ represents respectively 20 % cancer elimination for ages $0 < t \leq 20$, linearly increasing elimination from 20 % up to 80 % for $20 < t \leq 65$ and 80 % cancer elimination for $65 < t < 120$. The remaining choice $\theta_t(0.8, 0.2; 20, 65)$ is similar but with the 20 % and 80 % parameters interchanged.

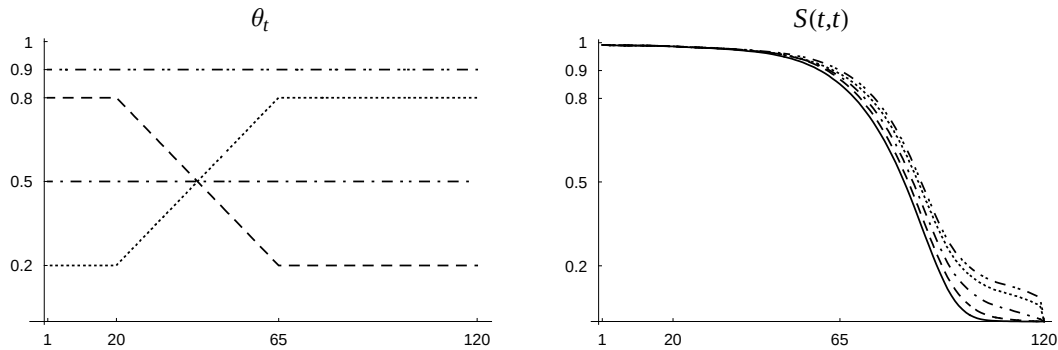


Fig. 3. The effect of different degree of partial elimination of cancer on the overall survival function $\mathcal{S}(t) \equiv S(t, t)$, $0 \leq t \leq 120$ - right panel for different choices of the elimination function θ_t - left panel.

7.1.2 Sensitivity of the results with respect to choice of copula

In this Section, we present a comparative study of the survival under the dependent competing risk model with respect to different choices of copula functions. For reasons stated in Section 4.1, we have performed the comparison by testing the proposed copula-based competing risk methodology with four different copulas, the Gaussian, the Student t -copula, the Frank and the Plackett copulas. For each copula we have used five different values of Kendall's τ equal to -0.91 , -0.35 , 0.00 , 0.35 , 0.91 . The results obtained show that the overall survival, given that cancer were eliminated as a cause of death, is most strongly affected by the choice of the copula in the case of extreme negative dependence, $\tau = -0.91$ and with the increase of τ this effect decreases. The curves of the density $f^{(o)}(t)$, $0 \leq t \leq 120$, for the four choices of copulas, are plotted in Fig. 4 in the case of $\tau = -0.35$ - left panel and $\tau = 0.35$ - right panel.

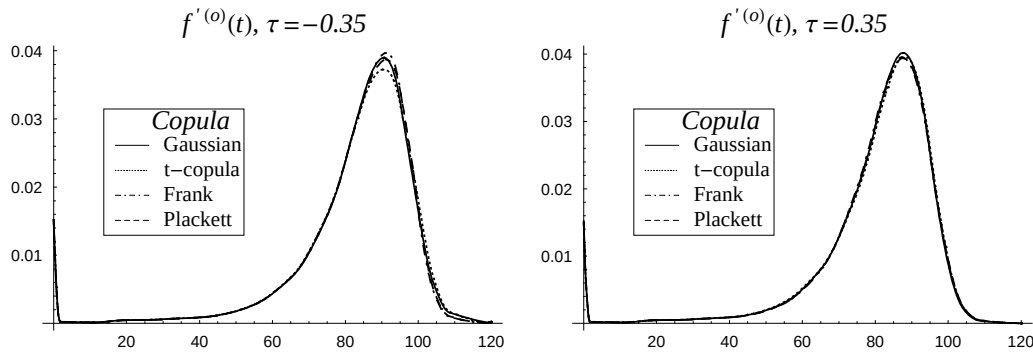


Fig. 4. The effect of the choice of copula on the overall survival density, $f^{(o)}(t)$, given cancer eliminated.

Although in the model presented here, it is more reasonable to look at the overall survival function $\mathcal{S}(t)$, $0 \leq t \leq 120$, the joint survival function of T_c and T_o , $S(t_1, t_2) = \Pr(T_c > t_1, T_o > t_2)$, $0 \leq t_j \leq 120$, $j = 1, 2$, is also of interest. However, since either one of the causes leads to death, and the other lifetime remains latent, probabilistic inference related to the joint distribution of T_c and T_o is somewhat artificial. Nevertheless, it is instructive and in Fig. 5-8 we have illustrated the joint density of T_c and T_o in case of the bivariate Gaussian, Student t -, Frank and Plackett copulas. For any bivariate copula, the joint density of T_c and T_o can be calculated from (14) as

$$\frac{\partial^2}{\partial t_1 \partial t_2} S(t_1, t_2) = c(S^{(c)}(t_1), S^{(o)}(t_2)) \times f^{(c)}(t_1) \times f^{(o)}(t_2). \quad (21)$$

In the upper panels of Fig. 5-8, negative dependence between T_c and T_o has been modeled with $\tau(T_1, T_2) = -0.35$, and in the lower panels, the modeled dependence is positive with $\tau(T_1, T_2) = 0.35$. The mass of the distribution is oriented in such a way that, under negative dependence, as seen from the upper panels of Fig. 5-8, higher values of T_o are more likely to occur jointly with smaller values of T_c . Under positive dependence, as seen from the lower panels, jointly increasing values of the lifetimes T_c and T_o are likely to occur. This is valid, regardless of what copula has been used, as can be seen from Fig. 5-8. There are, of course, some copula specific differences in the joint density functions, as is natural to expect in view of (21). For example, in the case of Student t -copula, there are probability masses in the corners inherited from the density of the copula itself (Fig. 6). In other cases, the joint density is significantly influenced by the (marginal) net survival functions, which cause probability masses to occur along the borders $T_c = 0$ and/or $T_o = 0$, as e.g., in the case of Frank copula (Fig. 7). Another, obvious characteristic of the joint density function is that for $\tau = -0.35$ it has one mode approximately around $T_o = 90$, $T_c = 80$ and for $\tau = 0.35$ it has a mode approximately around $T_o = 85$, $T_c = 95$.

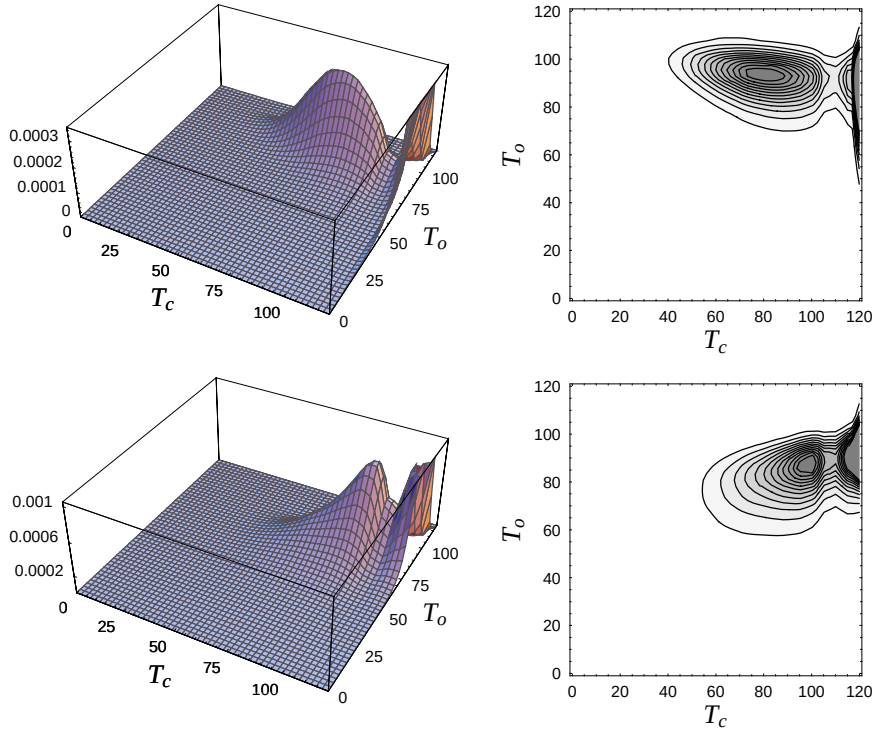


Fig. 5. A 3D plot and a contour plot of the joint density of T_c and T_o , expressed through the Gaussian copula for Kendall's $\tau = -0.35$ (upper panel) and $\tau = 0.35$ (lower panel).

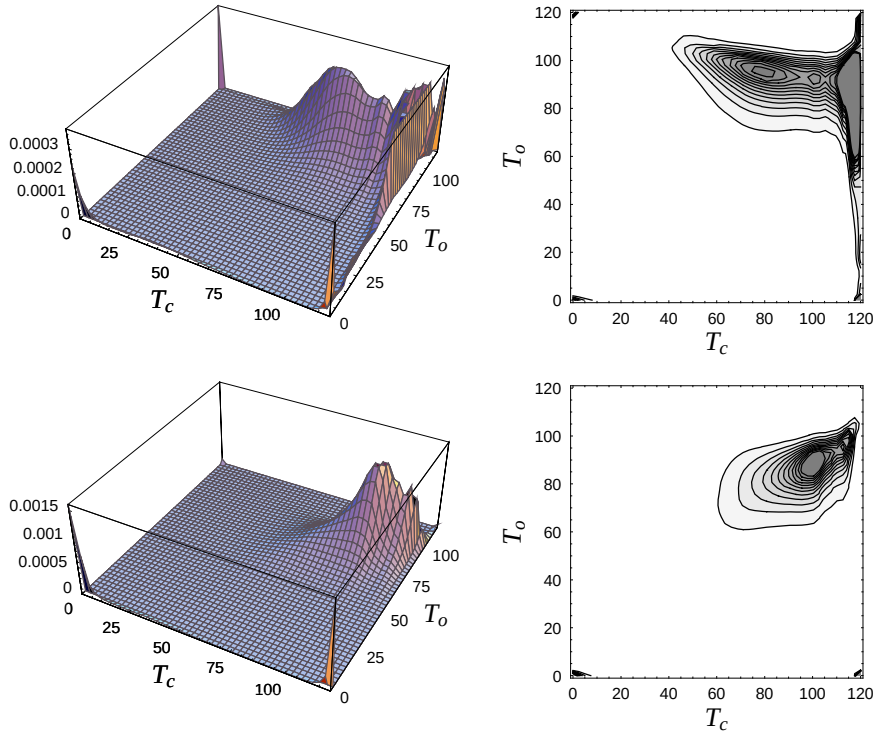


Fig. 6. A 3D plot and a contour plot of the joint density of T_c and T_o , expressed through the t -copula for Kendall's $\tau = -0.35$ (upper panel) and $\tau = 0.35$ (lower panel).

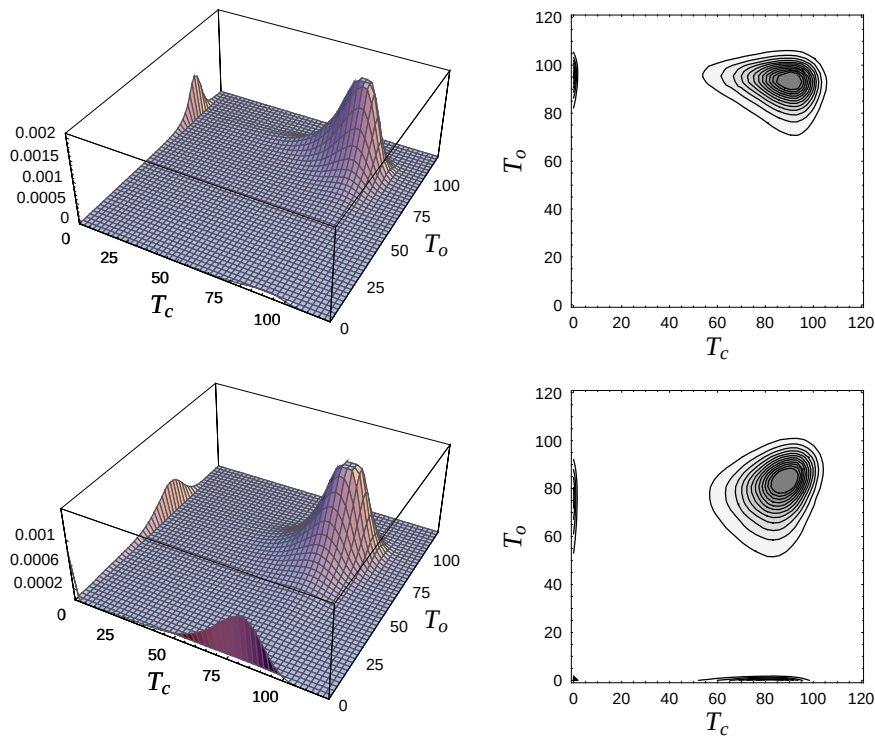


Fig. 7. A 3D plot and a contour plot of the joint density of T_c and T_o , expressed through the Frank copula for Kendall's $\tau = -0.35$ (upper panel) and $\tau = 0.35$ (lower panel).

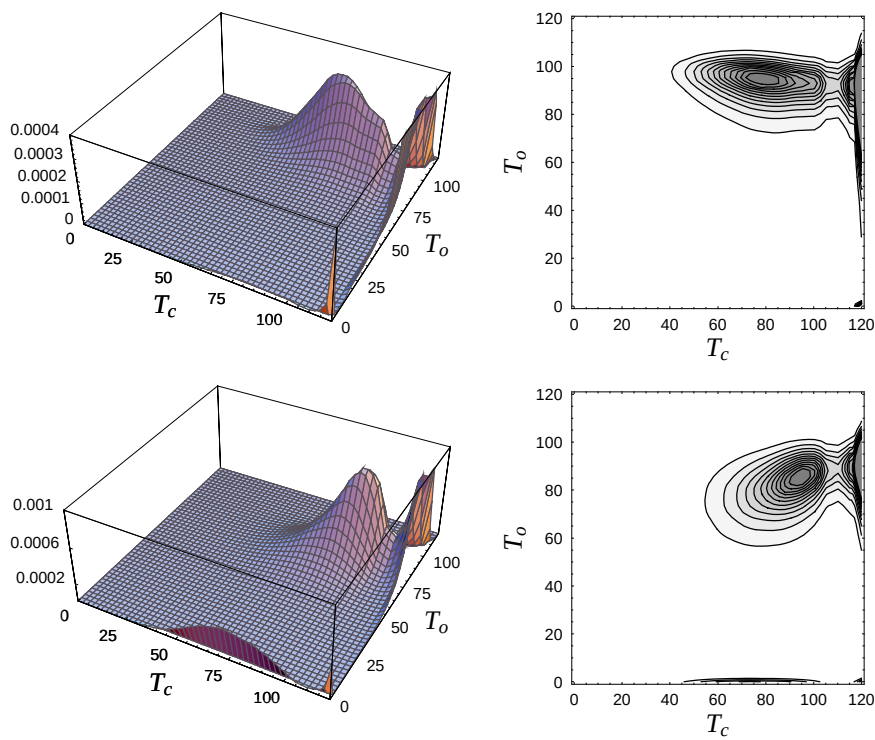


Fig. 8. A 3D plot and a contour plot of the joint density of T_c and T_o , expressed through the Plackett copula for Kendall's $\tau = -0.35$ (upper panel) and $\tau = 0.35$ (lower panel).

The sensitivity, with respect to the choice of copula, of the dependent competing risk model has been tested using actuarial functions, such as the life expectancy at birth, $e_0^{(-c)}$, at age 65, $e_{65}^{(-c)}$, and the whole of life annuity at age 65, $\bar{a}_{65}^{(-c)}$, calculated in case of complete cancer elimination. We are interested in assessing the gain due to cancer elimination, and so we have compared the latter actuarial functions, $e_0^{(-c)}$, $e_{65}^{(-c)}$ and $\bar{a}_{65}^{(-c)}$, with $e_0 = 78.83$, $e_{65} = 19.03$ and $\bar{a}_{65} = 12.63$, calculated in the case when none of the causes has been eliminated. The results of the sensitivity test are summarized in Tables 2-5 for the Gaussian, Student t -, Frank and Plackett copulas respectively.

Table 2. Gaussian copula results.

τ	$N(\rho)$	$e_0^{(-c)}$ [gain]	$e_{65}^{(-c)}$ [gain]	$\bar{a}_{65}^{(-c)}$ [gain]
-0.91	$\rho = -0.99$	87.43 [8.60]	26.75 [7.71]	15.08 [2.45]
-0.35	$\rho = -0.52$	83.37 [4.54]	22.26 [3.23]	14.13 [1.50]
0	$\rho = 0.00$	82.16 [3.34]	21.00 [1.97]	13.66 [1.03]
0.35	$\rho = 0.52$	81.05 [2.22]	20.00 [0.96]	13.21 [0.59]
0.91	$\rho = 0.99$	79.15 [0.32]	19.05 [0.02]	12.70 [0.07]

Table 3. Student's t -copula results.

τ	$T(\rho, \nu = 3)$	$e_0^{(-c)}$ [gain]	$e_{65}^{(-c)}$ [gain]	$\bar{a}_{65}^{(-c)}$ [gain]
-0.91	$\rho = -0.99$	87.00 [8.17]	26.27 [7.24]	15.06 [2.43]
-0.35	$\rho = -0.52$	83.43 [4.60]	22.36 [3.33]	14.14 [1.51]
0	$\rho = 0.00$	82.28 [3.46]	21.19 [2.16]	13.72 [1.09]
0.35	$\rho = 0.52$	81.18 [2.35]	20.20 [1.17]	13.29 [0.66]
0.91	$\rho = 0.99$	79.17 [0.34]	19.10 [0.07]	12.72 [0.09]

Table 4. Frank copula results.

τ	$F(\theta)$	$e_0^{(-c)}$ [gain]	$e_{65}^{(-c)}$ [gain]	$\bar{a}_{65}^{(-c)}$ [gain]
-0.91	$\theta = -44.88$	86.96 [8.13]	26.23 [7.20]	15.04 [2.41]
-0.35	$\theta = -3.46$	83.20 [4.37]	22.09 [3.05]	14.08 [1.45]
0	$\theta = 0.00$	82.16 [3.34]	21.00 [1.97]	13.66 [1.03]
0.35	$\theta = 3.46$	81.13 [2.30]	20.00 [0.97]	13.20 [0.57]
0.91	$\theta = 44.88$	79.27 [0.45]	19.00 [-0.04]	12.66 [0.03]

Table 5. Plackett copula results.

τ	$P(\theta)$	$e_0^{(-c)}$ [gain]	$e_{65}^{(-c)}$ [gain]	$\bar{a}_{65}^{(-c)}$ [gain]
-0.91	$\theta = \frac{1}{735.8}$	86.33 [7.50]	25.53 [6.50]	14.97 [2.34]
-0.35	$\theta = \frac{1}{5.022}$	83.24 [4.42]	22.14 [3.11]	14.09 [1.46]
0	$\theta = 0.00$	82.16 [3.34]	21.00 [1.97]	13.66 [1.03]
0.35	$\theta = 5.022$	81.17 [2.34]	20.07 [1.04]	13.23 [0.60]
0.91	$\theta = 735.8$	79.27 [0.44]	19.11 [0.08]	12.72 [0.09]

A general conclusion is that the gain in $e_0^{(-c)}$, $e_{65}^{(-c)}$ and $\bar{a}_{65}^{(-c)}$ is bounded by the gain in these actuarial functions, obtained with the Fréchet-Hoeffding copula bounds (9). As seen from Tables 2-5, the approximate upper bound for the gain in $e_0^{(-c)}$ is 8.60, for the gain in $e_{65}^{(-c)}$ is 7.71 and for the gain in $\bar{a}_{65}^{(-c)}$ is 2.45 years. So, the choice of copula can affect these actuarial functions only up to the corresponding upper bounds. As expected, each of $e_0^{(-c)}$, $e_{65}^{(-c)}$ and $\bar{a}_{65}^{(-c)}$ decreases as Kendall's τ increases, regardless of the choice of copula. It is important to note that the stated actuarial functions, computed for the four selected copulas, differ more significantly in the case of extreme negative dependence. In this case the maximum gain, in any of $e_0^{(-c)}$, $e_{65}^{(-c)}$ and $\bar{a}_{65}^{(-c)}$, is obtained for the Gaussian copula (Table 2) and the minimum for the

Plackett copula (Table 5). As can be seen from Tables 2-5, the differences between the selected actuarial functions become insignificant for τ equal to $-0.35, 0, 0.35, 0.91$ for the four copulas.

It is worth mentioning that our estimates of the gain in ${}^{\circ(-c)}e_0$ and in ${}^{\circ(-c)}e_{65}$, for the independent case, are 3.34 and 1.97, and these figures are consistent with the values 3.30 and 1.92 respectively, given in NCHS (1999). The small differences between these 2 sets of estimates may be due to the extrapolation up to the limiting age of 120, that has been performed in our study.

7.2 Multiple causes of death ($m = 4$)

We now illustrate the extension of the proposed methodology to the multivariate case by considering four competing causes, heart (h), cancer (c), respiratory diseases (r) and other causes (o). As in the bivariate case, we have constructed the four decrement table using "Table 10" from NCHS (1999). The interpolated crude survival functions $S^{(h)}(t), S^{(c)}(t), S^{(r)}(t), S^{(o)}(t), 0 \leq t \leq 120$ and their derivatives are given in Fig. 9.

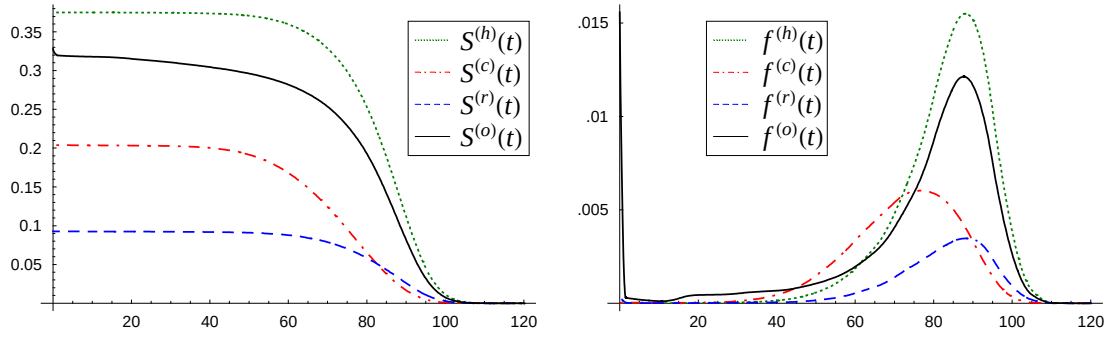


Fig. 9. The crude survival functions (left panel) and their densities (right panel).

As noted in Section 3.4, the multivariate Frank copula allows for modelling only positive dependence. So, here for illustrative purposes we have used the Frank copula to model purely positive dependence along with the Gaussian and t -copula which can capture a wide variety of pairwise dependencies between the lifetimes T_h, T_c, T_r and T_o given that the matrix $\mathbf{R}$ is positive definite. We have been successful in solving the system (15) in this four dimensional case for the selected three copulas. The four net survival functions and their densities for the Gaussian copula with (arbitrarily chosen) pairwise correlation coefficients $R_{hc} = -0.5, R_{hr} = -0.5, R_{ho} = 0.5, R_{cr} = 0.5, R_{co} = -0.5, R_{ro} = -0.5$, are presented in Fig. 10.

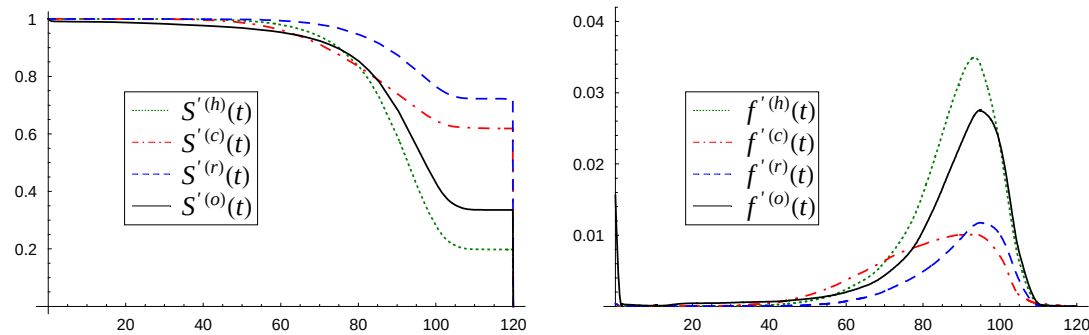


Fig. 10. The net survival functions (left panel) and their densities (right panel).

In the left panel of Fig. 11, we have given the obtained overall survival functions with no disease eliminated, $\mathcal{S}(t)$, and each one of the diseases eliminated, $\mathcal{S}^{(-h)}(t), \mathcal{S}^{(-c)}(t)$ and $\mathcal{S}^{(-r)}(t)$. As can be seen, the most significant improvement in survival for the age range $40 \leq t \leq 85$ is achieved if cancer is eliminated whereas

for $85 \leq t \leq 120$ the best improvement in survival is due to removal of heart diseases. As expected, improvement in survivorship due to removal of respiratory diseases is not as significant. Fig. 12 presents examples of the cases where 2 causes of death are simultaneously eliminated: the simultaneous elimination of cancer and heart diseases gives the best improvement in the overall survival over the entire range $40 \leq t \leq 120$, compared to the removal of any other combination of two causes. The densities given in the right panels of Fig. 11-12 combined with Table 6 provide information about the effect of removal of one disease and a pair of diseases on the life expectancy. As is natural to expect, the complete simultaneous elimination of heart, cancer and respiratory diseases provides the highest gain in the actuarial functions $e_0^{(-j)}$, $e_{65}^{(-j)}$ and $\bar{a}_{65}^{(-j)}$, $j \in \{h, c, r, hc, hr, cr, hcr\}$. The results obtained with the Student t -copula with identical correlation matrix and $\nu = 3$ are almost identical and therefore are omitted.

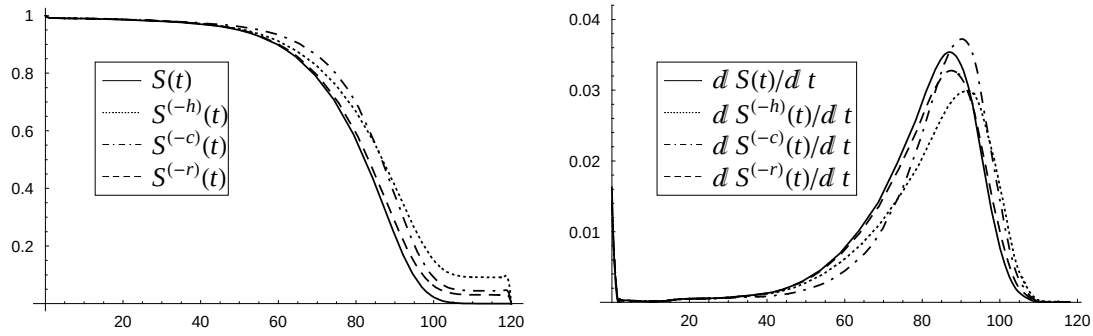


Fig. 11. The overall survival functions with no elimination and only one disease eliminated (left panel) and their densities (right panel).

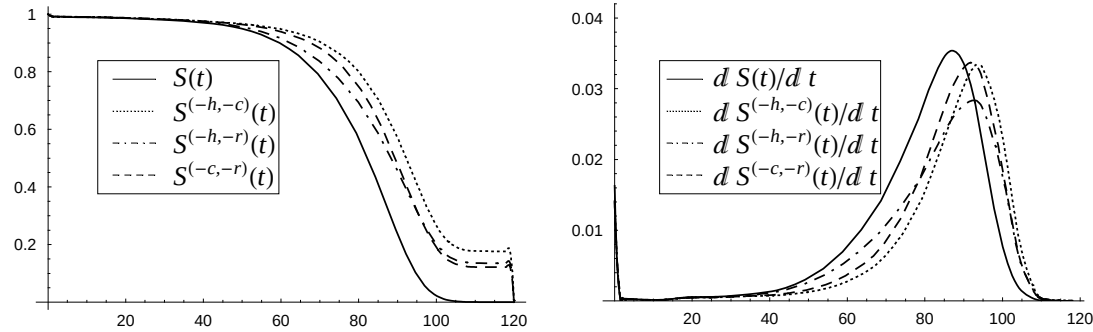


Fig. 12. The overall survival functions with no elimination and only two diseases eliminated (left panel) and their densities (right panel).

Table 6. Multivariate Gaussian copula results.

Disease, j	$j \equiv \{h\}$	$j \equiv \{c\}$	$j \equiv \{r\}$	$j \equiv \{hc\}$	$j \equiv \{hr\}$	$j \equiv \{cr\}$	$j \equiv \{hcr\}$
$e_0^{(-j)}$	84.77	84.16	80.54	91.51	86.86	88.17	97.51
[gain]	[5.96]	[5.35]	[1.72]	[12.69]	[8.05]	[9.35]	[18.70]
$e_{65}^{(-j)}$	25.06	23.18	20.82	30.20	27.13	27.14	36.09
[gain]	[6.04]	[4.16]	[1.80]	[11.18]	[8.12]	[8.12]	[17.08]
$\bar{a}_{65}^{(-j)}$	14.62	14.24	13.19	16.42	15.23	15.45	17.82
[gain]	[1.99]	[1.61]	[0.56]	[3.80]	[2.61]	[2.82]	[5.19]

Using the multivariate Frank copula, we have modelled positive dependence with parameter $\theta = 34.64$ which in the bivariate case corresponds to $\tau = 0.35$. The results are similar in nature but expressed to a lesser extent, as seen from Fig. 13 and Table 7, where the maximum gain achieved is just over half of that achieved with the Gaussian copula, where we have assumed positive as well as negative pairwise correlations.

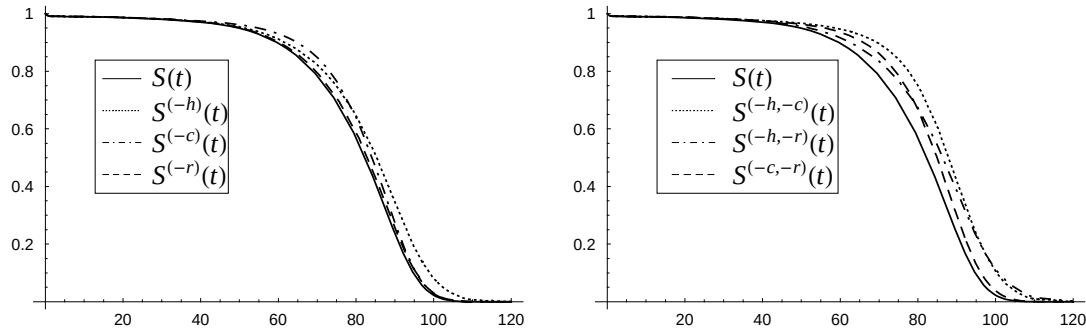


Fig. 13. The overall survival functions with no elimination and only one disease eliminated (left panel) and only two disease eliminated (right panel).

Table 7. Multivariate Frank copula results.

Disease, j	$j \equiv \{h\}$	$j \equiv \{c\}$	$j \equiv \{r\}$	$j \equiv \{hc\}$	$j \equiv \{hr\}$	$j \equiv \{cr\}$	$j \equiv \{hcr\}$
${}^{\circ(-j)}e_0$	81.95	81.11	79.48	85.28	83.24	81.96	87.13
[gain]	[3.14]	[2.30]	[0.67]	[6.46]	[4.43]	[3.15]	[8.31]
${}^{\circ(-j)}e_{65}$	21.83	19.98	19.53	23.59	23.00	20.62	25.22
[gain]	[2.82]	[0.96]	[0.52]	[4.57]	[3.99]	[1.60]	[6.20]
$\overline{a}_{65}^{(-j)}$	13.78	13.14	12.86	14.63	14.22	13.43	15.24
[gain]	[1.16]	[0.52]	[0.23]	[2.01]	[1.59]	[0.81]	[2.61]

8. Conclusions

The objective of this paper is to demonstrate how copulas may be used in the modelling of dependencies among causes of death for the purposes of analyzing the impact of the complete or partial elimination of causes of death on survival functions and related indices, expectations of life and annuity values.

The paper extends the earlier work of Carriere (1994) and Valdez (2001) to include more than two competing risks, to investigate the sensitivity of the model to the choice of copula and to provide a more rapid and universal computation process for solving the underlying system of non-linear differential equations, which yields a solution not just at integral ages (as in Carriere 1994) but at any (intermediate) age $0 \leq t \leq 120$.

The methodology and results may be applied to

- setting target levels for mortality rates that will assist with scenario testing and sensitivity analyses in the presence of dependence between causes of death
- population forecasting and planning
- life insurance business where the financial impact of mortality improvements on life insurance and annuities products may be investigated.

Possible extensions include allowing for the introduction of "new" causes of death (for example, as happened in the late 1980s with HIV/AIDS) and exploring the alternative approach to elimination as formulated in Section 2, following Elandt-Johnson (1976).

Acknowledgements

The authors would like to thank the anonymous referee for his valuable comments and suggestions which helped to improve the presentation of the paper.

References

- Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. (1997). *Actuarial Mathematics*. Itasca, III: Society of Actuaries.
- De Boor, C. (2001). *A practical guide to splines*. Revised Edition, New York: Springer.
- Bryant, J. and Dignam, J.J. (2004). Semiparametric Models for Cumulative Incidence Functions. *Biometrics*, 60, 182-190.
- Buettner, T. (2004). Approaches and experiences in projecting mortality patterns for oldest-old. *North American Actuarial Journal*, 6 (3), 14-29.
- Carriere, J. (1994). Dependent Decrement Theory. *Transactions, Society of Actuaries*, v. XLVI, 45-65.
- Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula methods in finance*. John Wiley and Sons Ltd.
- Coale, A., and Guo, G. (1989). Revised regional model life tables at very low levels of mortality. *Population Index* 55:613-643.
- Coale, A., and Kisker, E.E. (1990). Defects in data on old-age mortality in the United States: New Procedures for calculating schedules and life tables at the Higher Ages. *Asian and Pacific Population Forum* 4, 1-31.
- David, H.A. and Moeschberger, M.L. (1978). *The theory of competing risks*. Macmillan, New York.
- Deheuvels, P. (1978). Caractérisation complète des Lois extrêmes multivariées et de la convergence des types extrêmes, *Publ. l'Institut de Statist. L'université de Paris*, 23, 1-36.
- Dellaportas, P., Smith, A.F.M. and Stavropoulos, P. (2001). Bayesian analysis of mortality data, *Journal of the Royal Statistical Society, Series A*, 164, Part 2, 257-291.
- Elandt-Johnson, R.C. (1976). Conditional Failure Time Distributions under Competing Risk Theory with Dependent Failure Times and Proportional Hazard Rates. *Scandinavian Actuarial Journal*, 37-51.
- Elandt-Johnson, R.C. and Johnson, N.L. (1980). *Survival models and data analysis*. Wiley, New York.
- Embrechts, P., Lindskog, F. and McNeil, A. (2001). *Modelling Dependence with Copulas and Applications to Risk Management*. Department of Mathematics, ETHZ, manuscript.
- Escarela, G. and Carriere, J. (2003). Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research*, 12 (4), 333-349.

- Frees, E. and Valdez, E. (1998). Understanding Relationships Using Copulas. *North American Actuarial Journal*, v.2 No 1, 1-25.
- Georges, P., Lamy, A-G., Nicolas, E., Quibel, G. and Roncatti, T. (2001). Multivariate survival modelling: a unified approach with copulas. Groupe de Recherche Opérationnelle Crédit Lyonnais France, working paper.
- Heckman, J. J. and Honore B. E. (1989). The identifiability of the Competing Risk Model. *Biometrika*, v. 76, Issue 2, 325-330.
- Heligman, L. and Pollard, J. H. (1980). The age pattern of mortality. *J. Inst. Act.*, 107, 49-80.
- Himes, C.L., Preston, S. H. and Condran, G. (1994). A relational model of mortality at older ages in low mortality countries. *Population Studies*, 48, 269-291.
- Hooker, P.F., and Longley-Cook, L.H. (1957). Life and other Contingencies. Vol. II. Cambridge, Great Britain: University Press.
- Joe, H. (1997). Multivariate Models and Dependent Concepts. Chapman & Hall, London.
- Kaishev, V., Dimitrova, D., Haberman, S., Verrall, R. (2004). Automatic, computer-aided, geometric design of free-knot regression splines. Faculty of Act. Sci. and Stat., Cass Business School, Statistical Research Paper No 24, August 2004, IBSN 1 901615 81 2
- National Center for Health Statistics, (1997). U.S. decennial life tables for 1989–91, vol. 1, No 1. United States Life Tables. Hyattsville, Maryland.
- National Center for Health Statistics, (1999). U.S. decennial life tables for 1989–91, vol. 1, No 4. United States life tables eliminating certain causes of death. Hyattsville, Maryland.
- Nelsen, R. (1999). An Introduction to Copulas. Springer, New York.
- Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V., Flournoy, N., Farewell, V.T. and Breslow, N.E., (1978). The analysis of failure times in the presence of competing risks, *Biometrics*, 34, 541-554.
- Renshaw, A.E. and Haberman, S. (2003). Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Applied Statistics*, 52(1), 119-137.
- Salinas-Torres, V. H., de Brança Pereira, C. A., and Tiwardi, R.C. (2002). Bayesian Nonparametric Estimation in a Series or a Competing Risk Model. *Journal of Nonparametric Statistics*, 14, 449-458.
- Sklar, A. (1959). Functions de repartition a n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* 8, 229-231.
- Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences U.S.A.*, 72, 20-22.
- Valdez, E. (2001). Bivariate analysis of survivorship and persistency. *Insurance: Mathematics and Economics*, 29, 357-373.
- Yashin, A., Manton, K., and Stallard, E. (1986). Dependent competing risks: A stochastic process model. *Journal of Mathematical Biology*, 24, 119-140.
- Zheng, M., and Klein (1994). A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula. *Communications in Statistics - Theory and Methods*, 23 (8), 2299-2311.

Zheng, M., and Klein (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula, *Biometrika*, 82 (1), 127-138.

Appendix

We first describe how we have constructed the two decrement US female population data set (FP), using the multiple decrement "Table 10. Number of life table deaths from specific causes during age interval for the female population: United States, 1989-91" of NCHS (1999). This table contains the number of deaths by cause of death, relating to five year age groups. The first and the last age spans for which data are given in the table are correspondingly 0-1 and 100+. For the cancer column $d_x^{(c)}$ in our two decrement FP data set (see Table A1), we have used the "Malignant neoplasms" column of "Table 10" from NCHS (1999), in which numbers of deaths from all cancer causes, coded 140-208 with respect to ICD-9 (International Classification of Diseases, 9-th revision), are pooled together. In order to obtain the column of the other (non-cancer) deaths, $d_x^{(o)}$, in our FP data set, we subtracted the cancer deaths, $d_x^{(c)}$, from the difference $l_x^{(\tau)} - l_{x+5}^{(\tau)}$, where $l_x^{(\tau)}$ is the number of living at the beginning of age x . The values $l_x^{(\tau)}$ are listed in column "Number of living at beginning of age interval" of "Table 10" from NCHS (1999). Thus, we have formed two columns of data in our FP data set, number of deaths due to cancer, $d_x^{(c)}$, and number of deaths due to other causes, $d_x^{(o)}$. Based on these crude data, we obtain the observed values at ages $k = 1, 5, 10, \dots, 95, 100$ of the crude survival function $S_0^{(c)}(k)$ and $S_0^{(o)}(k)$, (see Table A2).

The extrapolation over the 100-120 age range has been performed using the US female mortality data for the old ages 100-115 given in "Table 3. Life table for females: United States, 1989-91" of NCHS (1997). Thus, we have obtained the number of female deaths from cancer for the range 100-120 as

$${}_5d_x^{(c)} = {}_\infty d_x^{(c)} \times ({}_5d_x / {}_\infty d_x), \quad x = 100, 105, 110, 115, 120,$$

where ${}_5d_x$ and ${}_\infty d_x$, are the numbers of female deaths from all causes, including cancer, given in "Table 3" from NCHS (1997).

Table A1: The extrapolated US female general population data set.

x	$d_x^{(c)}$	$d_x^{(o)}$	$l_x^{(\tau)}$
0–1	221	82579	10000000
1–4	1246	15354	9917200
5–9	1475	8025	9900600
10–14	1279	8421	9891100
15–19	1637	20063	9881400
20–24	2430	24770	9859700
25–29	3941	27259	9832500
30–34	8295	33405	9801300
35–39	16491	39809	9759600
40–44	31121	49979	9703300
45–49	56365	72635	9622200
50–54	93955	111145	9493200
55–59	140508	173392	9288100
60–64	194175	272525	8974200
65–69	244473	410827	8507500
70–74	288062	635438	7852200
75–79	299411	930689	6928700
80–84	276256	1310844	5698600
85–90	211549	1533351	4111500
90–95	117510	1314490	2366600
95–100	40591	668909	934600
100 – 105	7393	188407	225100
105 – 110	1050	26750	29300
110 – 115	45	1155	1500
115 – 120	11	289	300
120	0	0	0

x - age spans

$d_x^{(c)}$ - the number of deaths due to cancer during the age interval x

$d_x^{(o)}$ - the number of deaths due to non-cancer causes during the age interval x

$l_x^{(\tau)}$ - the number of living at the beginning of the age interval x

The following quantities are necessary in order to find the 'observed' values of the crude survival functions $S_0^{(c)}(x)$ and $S_0^{(o)}(x)$:

${}_{\infty}q_0^{(c)}$ - the multiple-decrement probability that a newborn will die from cancer

${}_{\infty}q_0^{(o)}$ - the multiple-decrement probability that a newborn will die from non-cancer causes

${}_{\infty}d_0^{(c)}$ - the total number of deaths from cancer for all ages from 0 to ∞

${}_{\infty}d_0^{(o)}$ - the total number of deaths from non-cancer causes for all ages from 0 to ∞

The following formulae were used to obtain the values of ${}_{\infty}q_0^{(c)} = 0.2039$, ${}_{\infty}q_0^{(o)} = 0.7961$, based on the values given in Table A1:

$${}_{\infty}q_0^{(c)} = \frac{\sum_x d_x^{(c)}}{l_0^{(\tau)}} = \frac{{}_{\infty}d_0^{(c)}}{l_0^{(\tau)}}; \quad {}_{\infty}q_0^{(o)} = \frac{\sum_x d_x^{(o)}}{l_0^{(\tau)}} = \frac{{}_{\infty}d_0^{(o)}}{l_0^{(\tau)}}.$$

Table A2: The FP data set.

k	${}_k d_0^{(c)}$	${}_k d_0^{(o)}$	$S_0^{(c)}(k)$	$S_0^{(o)}(k)$	$\mathcal{S}_0(k)$
0	—	—	0.203949	0.796051	1
1	221	82579	0.203927	0.787793	0.99172
5	1467	97933	0.203802	0.786258	0.99006
10	2942	105958	0.203655	0.785455	0.98911
15	4221	114379	0.203527	0.784613	0.98814
20	5858	134442	0.203363	0.782607	0.98597
25	8288	159212	0.203120	0.780130	0.98325
30	12229	186471	0.202726	0.777404	0.98013
35	20524	219876	0.201897	0.774063	0.97596
40	37015	259685	0.200248	0.770083	0.97033
45	68136	309664	0.197135	0.765085	0.96222
50	124501	382299	0.191499	0.757821	0.94932
55	218456	493444	0.182103	0.746707	0.92881
60	358964	666836	0.168053	0.729367	0.89742
65	553139	939361	0.148635	0.702115	0.85075
70	797612	1350188	0.124188	0.661032	0.78522
75	1085674	1985626	0.095382	0.597488	0.69287
80	1385085	2916315	0.065441	0.504420	0.56986
85	1661341	4227159	0.037815	0.373335	0.41115
90	1872890	5760510	0.016660	0.220000	0.23666
95	1990400	7075000	0.004909	0.088551	0.09346
100	2030991	7743909	0.000850	0.021660	0.02251
105	2038384	7932316	0.000111	0.002819	0.00293
110	2039433	7959067	0.000006	0.000144	0.00015
115	2039479	7960221	0.000001	0.000029	0.00003
120	2039490	7960510	0	0	0

k - exact age in years

${}_k d_0^{(c)}$ - the number of deaths due to cancer from age 0 to age k

${}_k d_0^{(o)}$ - the number of deaths due to other causes from age 0 to age k

$S_0^{(c)}(k)$ - 'observed' values at age k of the crude survival function for cancer

$S_0^{(o)}(k)$ - 'observed' values at age k of the crude survival function for other causes

$\mathcal{S}_0(k)$ - 'observed' values at age k of the overall survival function

The following formulae were used to calculate the values of ${}_k d_0^{(c)}$, ${}_k d_0^{(o)}$, $S_0^{(c)}(k)$, $S_0^{(o)}(k)$ and $\mathcal{S}_0(k)$ given in Table A2, based on the values given in Table A1:

$${}_k d_0^{(c)} = \sum_{x < k} d_x^{(c)}; \quad {}_k d_0^{(o)} = \sum_{x < k} d_x^{(o)};$$

$$S_0^{(c)}(k) = {}_\infty q_0^{(c)} - \frac{{}_k d_0^{(c)}}{l_0^{(\tau)}}; \quad S_0^{(o)}(k) = {}_\infty q_0^{(o)} - \frac{{}_k d_0^{(o)}}{l_0^{(\tau)}}; \quad \mathcal{S}_0(k) = S_0^{(c)}(k) + S_0^{(o)}(k) = \frac{l_k^{(\tau)}}{l_0^{(\tau)}};$$

$$S_0^{(c)}(0) = {}_\infty q_0^{(c)}; \quad S_0^{(o)}(0) = {}_\infty q_0^{(o)}; \quad \mathcal{S}_0(0) = S_0^{(c)}(0) + S_0^{(o)}(0) = \frac{l_0^{(\tau)}}{l_0^{(\tau)}} = 1.$$