

ENVIRONMETRICS

*Environmetrics*, **10**, 753–767 (1999)

## CONSTRAINED MIXTURE MODELS IN COMPETING RISKS PROBLEMS

S. K. NG,<sup>1</sup> G. J. McLACHLAN,<sup>1\*</sup> D. C. McGIFFIN<sup>2</sup> AND M. F. O'BRIEN<sup>3</sup><sup>1</sup>*Department of Mathematics, The University of Queensland, St. Lucia, Queensland 4072, Australia*<sup>2</sup>*Department of Surgery, The University of Alabama, Birmingham, AL 35294-0007, USA*<sup>3</sup>*Department of Cardiac Surgery, The Prince Charles Hospital, Brisbane, Australia*

### SUMMARY

We consider the problem of modelling the failure-time distribution, where failure is due to two distinct causes. One approach is to adopt a two-component mixture model where the components correspond to the two different causes of failure. However, routine application of this approach with typical parametric forms for the component densities proves to be inadequate in modelling the time to a re-replacement operation or death after the initial replacement of the aortic valve in the heart by a prosthesis, such as a xenograft valve. Hence we consider modifications to the usual mixture model approach to handle situations where there exists a strong dependency between the failure times of the distinct causes. With these modifications, a suitable model is able to be provided for the distribution of the time to a re-replacement operation conditional on the age of the patient at the time of the initial replacement operation. The estimate so obtained by the probability that a patient of a given age will undergo a re-replacement operation provides a useful guide to heart surgeons on the type of valve to be used in view of the patient's age. Copyright © 1999 John Wiley & Sons, Ltd.

KEY WORDS competing risks; constrained mixture models; EM algorithm; proportional hazards model

### 1. INTRODUCTION

In the analysis of failure time data, it is often necessary to consider different types of failure. Suppose that there are  $g = 2$  distinct causes of failure and that on each entity or individual under study, the aim is to record  $\mathbf{y} = (T, \mathbf{x}^T, \delta)^T$ , where  $T$  denotes the failure time,  $\mathbf{x}$  is a vector of covariates, and  $\delta = i$  if failure is due to the  $i$ th cause ( $i = 1, 2$ ). The superscript  $T$  denotes vector transpose. In the case where the study terminates before failure occurs,  $T$  is the censoring time and  $\delta$  is set equal to zero to indicate that the failure time is right-censored.

The traditional approach to the modelling of the distribution of failure time in the case of competing risks is to postulate the existence of so-called latent failure times corresponding to the two causes and to proceed on the basis that the two causes are independent of each other (see David and Moeschberger 1978; Kalbfleisch and Prentice 1980).

\* Correspondence to: G. J. McLachlan, Department of Mathematics, The University of Queensland, St. Lucia, Queensland 4072, Australia.

An alternative approach is to adopt a two-component mixture model, whereby the survival function of  $T$  is modelled as

$$S(t; \mathbf{x}) = \pi_1(\mathbf{x})S_1(t; \mathbf{x}) + \pi_2(\mathbf{x})S_2(t; \mathbf{x}), \quad (1)$$

where the  $i$ th component survival function  $S_i(t; \mathbf{x})$  denotes the conditional survival function given failure is due to the  $i$ th cause, and  $\pi_i(\mathbf{x})$  is the probability of failure from the  $i$ th cause ( $i = 1, 2$ ). It is common to assume that the mixing proportions  $\pi_i(\mathbf{x})$  have the logistic form.

$$\pi_1(\mathbf{x}; \boldsymbol{\beta}) = 1 - \pi_2(\mathbf{x}; \boldsymbol{\beta}) = \exp(a + \mathbf{b}^T \mathbf{T}) / (1 + \exp(a + \mathbf{b}^T \mathbf{x})), \quad (2)$$

where  $\boldsymbol{\beta} = (a, \mathbf{b}^T)^T$  is the vector of logistic regression coefficients (see Farewell 1977, 1982, 1986). This was the approach adopted by Larson and Dinse (1985) who assumed also that the component survival functions follow a proportional hazards model (Cox 1972) and that the baseline hazard functions  $h_{0i}(t)$  ( $i = 1, 2$ ) are piecewise constant for simplicity. That is,

$$\begin{aligned} h_i(t; \mathbf{x}) &= e^{\boldsymbol{\gamma}_i^T \mathbf{x} h_{0i}(t)} \\ &= e^{\boldsymbol{\gamma}_i^T \mathbf{x} + \alpha_{im}}, \text{ if } t \in J_m, \end{aligned} \quad (3)$$

where  $\boldsymbol{\gamma}_i$  is a vector of parameters,  $J_1, \dots, J_M$  are  $M$  prespecified disjoint intervals that totally exhaust the non-negative real line, and  $\alpha_{im}$  is the parameter representing the log of the  $i$ th component baseline hazard on the  $m$ th interval  $J_m$  ( $i = 1, 2; m = 1, \dots, M$ ). As an alternative to this specification of  $h_{0i}(t)$ , we can adopt some parametric form for the  $i$ th component baseline hazard function  $h_{0i}(t)$  ( $i = 1, 2$ ). For example, Gordon (1990a) adopted this Gompertz distribution to specify the conditional survival functions in the context of estimating the 'cure' rate of breast cancer after a treatment therapy. She also examined the applicability of the mixture model in fitting competing risks data through a simulation study (Gordon 1990b). Kuk (1992) on the other hand considered a semiparametric version of (3) by treating the baseline hazard functions  $h_{0i}(t)$  as nuisance parameters to be eliminated during the analysis.

In this paper, we consider the application of the two-component mixture model (1) to estimate the probability that a patient aged  $x$  years will undergo a re-replacement operation after having his/her native aortic valve replaced by a xenograft prosthesis. At the time of the initial replacement operation, the surgeon has the choice of using either a mechanical valve or a biologic valve such as a xenograft (made from porcine valve tissue) or an allograft (human donor valve). Modern day mechanical valves are very reliable, but a patient must take blood-thinning drugs for the rest of his/her life to avoid thromboembolic events. On the other hand, biologic valves have a finite working life, and so have to be replaced if the patient were to live for a sufficiently long enough time after the initial replacement operation.

With respect to this problem, we now let  $T$  denote the time to either a re-replacement operation or to death without a re-replacement operation. For brevity, we shall henceforth refer to the event of a re-replacement operation as a reoperation. Then for a patient aged  $x$  years at the time of the initial replacement operation, we can model the survival function of  $T$  by (1), where  $S_1(t; x)$  denotes the conditional survival function given the patient undergoes a re-replacement and  $S_2(t; x)$  denotes the conditional survival function given the patient dies without needing a reoperation. The mixing proportion  $\pi_1(x)$  then represents the probability of interest, namely that a patient

aged  $x$  years at the time of the initial replacement operation will undergo a reoperation in his/her lifetime.

In this application, specification of a typically used parametric form for the component survival functions in the mixture model (1) gave unsatisfactory results. Hence we had to consider a modified version of (1) in which constraints are effectively imposed on the component survival functions. It will be seen that this mixture model with constrained components gives results that are supported by the traditional latent failure-time approach based on the assumption of independent competing risks. The results of this latter approach for presented in the next section, while the constrained mixture approach is considered in Sections 4 and 5. Before we proceed with the implementation of these two approaches, we give a brief account in Section 3 on why there is a need to constrain the components of the mixture model (1) for this application.

## 2. LATENT FAILURE-TIME APPROACH

With the traditional approach to the handling of competing risks, consideration is given to the hypothetical latent failure times corresponding to each cause in the absence of the other (Moeschberger and David 1971). Accordingly, we let  $T_R$  be the latent failure time for reoperation in the absence of death and, likewise, we let  $T_D$  be the latent failure time for death in the absence of the risk due to reoperation. We let  $f_R(t; x)$  and  $S_R(t; x)$  denote the density and survival functions of  $T_R$ , and  $f_D(t; x)$  and  $S_D(t; x)$  the density and survival functions of  $T_D$ . With this approach, it is common to assume that the competing risks are independent. Under this assumption, the survival function  $S(t; x)$  for the observable  $T = \min(T_R, T_D)$ , is given by

$$S(t; x) = \int_t^\infty f_R(u; x)S_D(u; x) du + \int_t^\infty f_D(u; x)S_R(u; x) du. \quad (4)$$

This was the approach used by Grunkemeier *et al.* (1994) and McGiffin *et al.* (1997) in their statistical analyses of data on the time to reoperation after an initial valve-replacement operation.

To illustrate why there is a need to have the components of the two-component mixture model (1) inter-related or some constraints imposed on them, we can write (4) as the two-component mixture model,

$$S(t; \mathbf{x}) = \pi_1(x)S_1(t; x) + \pi_2(x)S_2(t; x), \quad (5)$$

where the mixing proportions  $\pi_i(x)$  are given by

$$\pi_1(x) = 1 - \pi_2(x) = \int_0^\infty f_R(u; x)S(u; x) du, \quad (6)$$

and the component survival functions  $S_i(t; x)$  by

$$S_1(t; x) = \left\{ \int_t^\infty f_R(u; x)S_D(u; x) du \right\} / \pi_1(x) \quad (7)$$

and

$$S_2(t; x) = \left\{ \int_t^\infty f_D(u; x)S_R(u; x) du \right\} / \pi_2(x). \quad (8)$$

It can be seen from the forms (7) and (8) for the component survival functions that if the latent failure due to one risk is much greater than the other, then this will affect the component survival function corresponding to the other risk through the effect of truncation in the integrand of the intervals on the right-hand sides of (7) and (8). For instance, in the present application, a reoperation is eventually needed if a patient were to live for a sufficiently long enough period following the initial replacement operation. Hence the presence of the latent survival function for reoperation,  $S_R(u; x)$ , in the integrand of the integral on the right-hand side of (8) can have a marked effect on the consequent form for the component distribution  $S_2(t; x)$  corresponding to death without a reoperation. This gives some idea why the usual Gompertz model for death is not appropriate for  $S_2(t; x)$ . We shall consider this further in Section 4, but firstly, we shall analyse the valve-replacement data by using the two-component mixture model (1) with (unconstrained) Gompertz component baseline survival functions and proportional hazards assumed for the effect of the age  $x$  of the patient at the time of the initial replacement operation. A guide to the applicability of the logistic model (2) for the mixing proportions  $\pi_i(x)$  can be obtained by plotting the log of  $\pi_1(x)/\pi_2(x)$  versus  $x$ , where  $\pi_1(x)$  is assessed using estimates of the latent failure-time density and survival functions in the right-hand side of (6).

### 3. MIXTURE MODELS WITH UNCONSTRAINED COMPONENTS

The reoperation data set considered in this paper was drawn from a more extensive data set involving 2100 patients undergoing 2366 aortic valve replacements with a variety of allograft, xenograft, and mechanical valves (McGiffin *et al.* 1993). The present data set consisted of  $n = 950$  cases of aortic valve replacements that were performed with xenograft prostheses. There were 62 subsequent valve-replacement reoperations due to either xenograft degeneration or some other reason, while 198 patients died without a reoperation. The remaining 690 survival times were all censored. The proportion of censoring observation is as large as 73%. Some preliminary results for a similar data set were reported in McGiffin *et al.* (1997), using the latent-failure time approach.

Under the logistic model (2),

$$\pi_1(x; \boldsymbol{\beta}) = 1 - \pi_2(x; \boldsymbol{\beta}) = \exp(a + bx)/(1 + \exp(a + bx)), \quad (9)$$

where  $\boldsymbol{\beta} = (a, b)^T$  is the vector of logistic regression, was adopted for the mixing proportions  $\pi_1(x)$  and  $\pi_2(x)$ , corresponding to reoperation and death without reoperation, respectively, as a function of the covariate  $x$ , being the age of the patient at the time of the initial replacement operation. Parametric forms for the component survival functions are specified by assuming proportional hazard models and taking the baseline hazard functions  $h_{0i}(t)$  ( $i = 1, 2$ ) to have the Gompertz form. That is, the  $i$ th component hazard function is specified as

$$h_i(t; \boldsymbol{\theta}_i, x) = \exp(\gamma_i x) h_{0i}(t), \quad (i = 1, 2), \quad (10)$$

where

$$h_{0i}(t) = \exp(\lambda_i + \xi_i t),$$

with  $\xi_i > 0$ , and  $\boldsymbol{\theta}_i = (\lambda_i, \xi_i, \gamma_i)^T$  ( $i = 1, 2$ ). We let

$$\boldsymbol{\Psi} = (\boldsymbol{\beta}^T, \boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$$

be the vector containing all the unknown parameters. Hence the component survival functions  $S_i(t; \boldsymbol{\theta}_i, x)$  are specified as

$$S_i(t; \boldsymbol{\theta}_i, x) = \exp\{-e^{\lambda_i + \gamma_i x} (e^{\xi_i t} - 1) / \xi_i\}, \quad (i = 1, 2). \quad (11)$$

The assumption of the Gompertz model for death in the absence of the competing risk of reoperation (that is, for the latent survival function  $S_D(t; x)$  for death) would appear to be reasonable, as this distribution is often used in the literature to model death in the absence of competing risks. However, it does not follow that this model will be applicable for death without a reoperation. Indeed, it will be seen that assumption (10) of the Gompertz model for the baseline hazard function and proportional hazards for the effect of age is inadequate for the second component corresponding to death without a reoperation. However, it does appear to be adequate for the first component corresponding to the event reoperation before death. As xenograft valves tend to degenerate more rapidly in younger patients (McGiffin *et al.* 1997), there is a need to have the first component hazard function (corresponding to reoperation) depending on the age of the patient, as well as the second component hazard function (corresponding to death without a reoperation), which clearly depends on age.

The observed data are of the form

$$\mathbf{y}_1 = (t_1, x_1, \delta_1)^T, \dots, \mathbf{y}_n = (t_n, x_n, \delta_n)^T,$$

where  $\delta_j = 1$  if the  $j$ th patient undergoes a reoperation,  $\delta_j = 2$  if the  $j$ th patient dies without a reoperation, and  $\delta_j = 0$  if the  $j$ th patient is still alive without having undergone a reoperation by the end of the study (that is, the failure time is then right-hand censored at time  $t_j$ ). The identifiability of mixtures of Gompertz distributions has been established by Gordon (1990a) in the case of mixing proportions that do not depend on any covariates. The extension to the case of mixing proportions specified by the logistic model (9) is straightforward. It follows that a sufficient condition for identifiability of the Gompertz mixture model is that the matrix  $(\mathbf{x}_1^+, \dots, \mathbf{x}_n^+)^T$  be of full rank, where

$$\mathbf{x}_j^+ = (1, x_j)^T.$$

The log likelihood function is given by

$$\begin{aligned} \log L(\boldsymbol{\Psi}) = & \sum_{j=1}^n [I(\delta_j = 1) \log \{\pi_1(x_j; \boldsymbol{\beta}) f_1(t_j; \boldsymbol{\theta}_1, x_j)\} \\ & + I(\delta_j = 2) \log \{\pi_2(x_j; \boldsymbol{\beta}) f_2(t_j; \boldsymbol{\theta}_2, x_j)\} \\ & + I(\delta_j = 0) \log S(t_j; \boldsymbol{\Psi}, x_j)], \end{aligned} \quad (12)$$

Table I. Maximum likelihood estimates (with standard errors) for a mixture of two Gompertz distributions with proportional hazards

Logistic model		Components					
$a$	$b$	Reoperation			Death without a reoperation		
		$\lambda_1$	$\xi_1$	$\gamma_1$	$\lambda_2$	$\xi_2$	$\gamma_2$
2.052 (0.57)	-0.0726 (0.013)	-3.530 (0.74)	0.507 (0.098)	-0.0296 (0.015)	-5.164 (0.55)	0.109 (0.025)	0.0245 (0.008)

where  $I(A)$  is the indicator function for the set  $A$ . The maximum likelihood estimate of  $\Psi$  can be obtained via the EM algorithm of Dempster *et al.* (1977). Further discussion of the EM algorithm in its application to mixture models in a general context may be found in McLachlan and Basford (1988) and McLachlan and Krishnan (1997). McLachlan *et al.* (1997) have developed a FORTRAN program for the fitting of a mixture of two Gompertz distributions to censored survival data, using a generalized version of the EM algorithm.

The estimates so obtained are presented in Table I, where the standard errors of the estimates are given in parentheses. The latter were computed by applying the nonparametric bootstrap approach of Efron (1979, 1982) with the resampling scheme slightly modified for the competing risks problems. Let  $N_i$  be the number of cause  $i$  failures ( $i = 1, 2$ ), and let  $N_3$  be the number of censored observations. The bootstrap data are obtained by sampling separately from each of the three sets, corresponding to cause  $i$  failures ( $i = 1, 2$ ) and the censored observations, with the sizes of these bootstrap subsamples taken equal to  $N_1$ ,  $N_2$ , and  $N_3$ , respectively. A similar resampling scheme has been used by Golbeck (1992) in bootstrapping life-table estimators. In this application,  $K = 100$  bootstrap samples were generated and the variances of the maximum likelihood estimates are the sample variances of the corresponding estimates based on the bootstrap samples.

Using the estimates  $\hat{a}$  and  $\hat{b}$ , the estimated probability  $\pi_1(x; \hat{\beta})$  of reoperation is plotted versus the age of the patient in Figure 1. The results in Figure 1 appear to be counterintuitive. For example, it is estimated that only 65% of 20 year old patients will need to undergo a reoperation. However, it is observed from the original data that 15 out of 18 non-censored patients of ages between 10 and 30 years (83%) had to undergo a reoperation, and the censored observations in this group would be expected to have very high chances of undergoing a reoperation. That is, the probability  $\pi_1(x; \hat{\beta})$  at  $x = 20$  is underestimated by the mixture model with components (11) using proportional hazards in conjunction with Gompertz baseline hazard functions. This is because during the estimation process, censored observations are given too high a probability of belonging to the second component, so that  $\pi_1(x; \hat{\beta})$  and  $S_1(t; \hat{\theta}_1, x)$  are underestimated. This bias is serious here, as there is a large proportion of censored observations. In this situation, there is inadequate information in the observed data on the tail of  $S_1(t; \hat{\theta}_1, x)$ . This is typically the case in practice, as most clinical studies require a very long follow-up to provide adequate information on the tails of both the component survival functions.

As we surmised at the start of this section, we would not expect the proportional hazards model to be appropriate for the second component for the mixture model corresponding to death before reoperation. To investigate this further now, we applied the semiparametric method of Kuk (1992) to this data set. This method adopts a marginal likelihood approach, whereby the

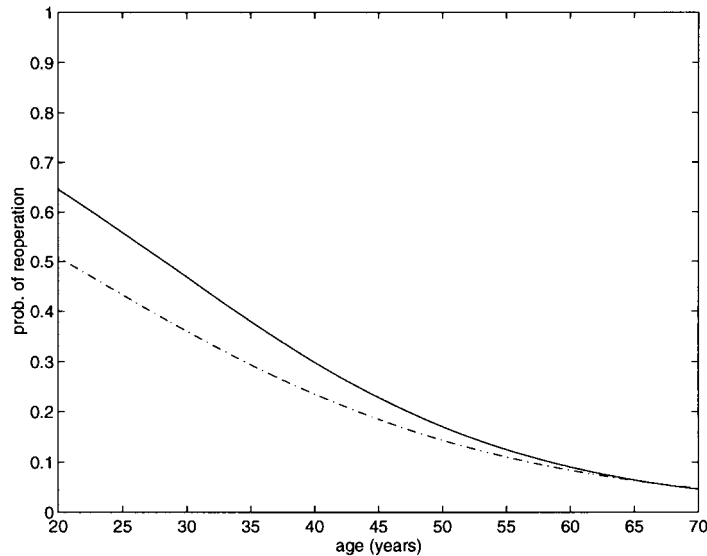


Figure 1. Estimated probability of reoperation at various age of patient: mixture model —; semi-parametric mixture model - - - - -

baseline hazard functions in (3) do not have to be specified, as they are treated as nuisance parameters and eliminated during the analyses. The estimate so obtained for the probability  $\pi_1(x; \beta)$  of a reoperation of a patient aged  $x$  years at the time of the initial replacement operation is displayed in Figure 1. It can be seen that it even gives a lower estimate of  $\pi_1(x; \beta)$  than the unsatisfactory assessment using the two-component mixture model with specified parametric forms (11) for the component baseline hazard functions. It would seem therefore that the problem lies with the assumption of proportional hazards for the effect of the age  $x$ .

#### 4. CONSTRAINED MIXTURE MODELS

We now proceed to consider a model for the second component hazard function, corresponding to death without a reoperation, that does not assume proportional hazards. As a xenograft prosthesis will eventually need replacement if the patient were to live long enough after the initial replacement operation, it means that the second component hazard function  $h_2(t; x)$  for death without a reoperation will be greater than the first component hazard function  $h_1(t, x)$  corresponding to reoperation before death; that is, for a given  $x$ ,

$$h_2(t; x) > h_1(t; x) \quad \forall t. \quad (13)$$

One way in which this constraint is automatically satisfied by the components of a two-component mixture model with component survival functions  $S_1(t; x)$  and  $S_2(t; x)$  is to specify the second component function  $S_2(t; x)$  as

$$S_2(t; x) = S_1(t; x)S_a(t; x), \quad (14)$$

where  $S_a(t; x)$  denotes some additional survival function. Then for (14), we have that

$$S(t; x) = \pi_1(x)S_1(t; x) + \pi_2(x)S_1(t; x)S_a(t; x). \quad (15)$$

From (14), it follows that

$$h_2(t; x) = h_1(t; x) + h_a(t; x) \quad (16)$$

$$> h_1(t; x) \quad \forall t, \quad (17)$$

which implies that the inequality (13) holds. In (16),  $h_a(t; x)$  denotes the hazard function corresponding to the survival function  $S_a(t; x)$  introduced in the model (14) for  $S_2(t; x)$ .

From (17), (15) can be considered as a constrained mixture model. We shall adopt it here with the component survival function  $S_1(t; x)$  for reoperation modelled as before by (11) and with the same parametric form for  $S_a(t; x)$ ; that is, it assumes a proportional hazards function with the Gompertz distribution used to model the baseline hazard function, namely

$$S_a(t; \boldsymbol{\theta}_a, x) = \exp\{-e^{\lambda_a + \gamma_a x}(e^{\xi_a t} - 1)/\xi_a\},$$

where  $\xi_a > 0$ , and  $\boldsymbol{\theta}_a = (\lambda_a, \xi_a, \gamma_a)^T$ . However, the component hazard function for death without a reoperation now no longer has the proportional hazards form.

The model (14) for the component survival function  $S_2(t; x)$  for death without a reoperation has an easy to understand interpretation. Given that a patient will die without a reoperation, his/her conditional hazard function can be viewed as being equal to the conditional hazard function for reoperation plus some an additional hazard denoted here by  $h_a(T; x)$ . A similar model to (14), but without any covariate, was used by Berkson and Gage (1952) to estimate the cure rate of breast cancer patients. The interpretation of their model was that 'cured' patients are subjected only to the 'normal' hazard for death, while uncured patients are subjected to an additional hazard from the risk of cancer. However, their formulation of the likelihood function was different from ours. Also, Berkson and Gate (1952) estimated the component survival function corresponding to normal death from standard life tables rather than from the observed data, which considerably simplified the estimation process. Previously, Boag (1949) had given a formulation of this problem in which the kernel of the log likelihood reduced to the same form as for the so-called long-term survival mixture model (see, for example, Farewell 1977; Maller and Zhou 1996).

The modified mixture model (15), which we shall call the constrained mixture model, can be fitted by maximum likelihood via the EM algorithm. To facilitate its implementation on the M-step, we used the ECM algorithm, where the M-step is replaced by computationally simpler CM-steps (Meng and Rubin 1993). The ECM algorithm shares all the appealing convergence properties of EM and always increases the likelihood after each iteration (Meng 1994). For the present problem, we divided the vector  $\boldsymbol{\Psi}$  of unknown parameters into three CM-steps corresponding to  $\boldsymbol{\beta}$ ,  $\boldsymbol{\theta}_1$ , and  $\boldsymbol{\theta}_a$ . The FORTRAN program of McLachlan *et al.* (1997) for the fitting of mixtures of Gompertz components with proportional hazards was modified to handle the constrained model (15). The resulting estimates are reported in Table II, along with their standard errors in parentheses.

In Figure 2, we have plotted the estimate of the probability  $\pi_1(x; \boldsymbol{\beta})$  that a patient aged  $x$  years at the initial replacement operation will undergo a reoperation, as obtained by the constrained mixture model. It can be seen that it is good agreement with the corresponding estimate as given



Table II. Maximum likelihood estimates (with standard errors) for constrained mixture model

Logistic model		Components					
$a$	$b$	Reoperation			Death without a reoperation		
		$\hat{\lambda}_1$	$\hat{\xi}_1$	$\hat{\gamma}_1$	$\hat{\lambda}_a$	$\hat{\xi}_a$	$\hat{\gamma}_a$
3.671 (0.82)	-0.0757 (0.016)	-2.783 (0.37)	0.275 (0.048)	-0.0441 (0.0092)	-5.454 (1.07)	0.0672 (0.026)	0.0308 (0.014)

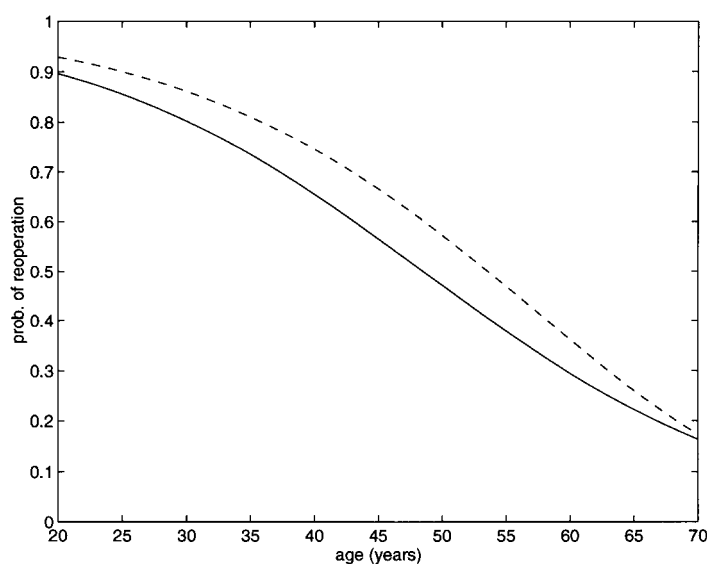


Figure 2. Estimated probability of reoperation at a given age of patient: Constrained mixture model (—); Latent failure-time approach (---)

by the latent failure-time approach applied under the assumption of independent competing risks. As this assumption would not appear to be a strong one for the application under study, it thus appears that the constrained mixture model provides a satisfactory model for the time to either reoperation or death without reoperation after the replacement of the native aortic valve by a xenograft prosthesis.

In Figure 3, we have plotted the fitted component survival function  $S_1(t; \hat{\theta}_1, x)$  along with the corresponding estimate as obtained with the latent failure-time approach assuming independent competing risks. It can be seen that there is good agreement between this latter method and the constrained mixture model approach. Although in Figure 3 we have plotted  $s_1(T; \hat{\theta}_1, x)$  for  $t$  up to 15 years for even a patient aged  $x = 70$  years at the time of the initial replacement operation, this estimate for large  $x$  should only be used for values of  $t$  in the practical range of interest. This is because for large values of  $t$ , the proportional hazards model (10) for the effect of age  $x$  provides only a crude approximation to reality for age  $x$ .

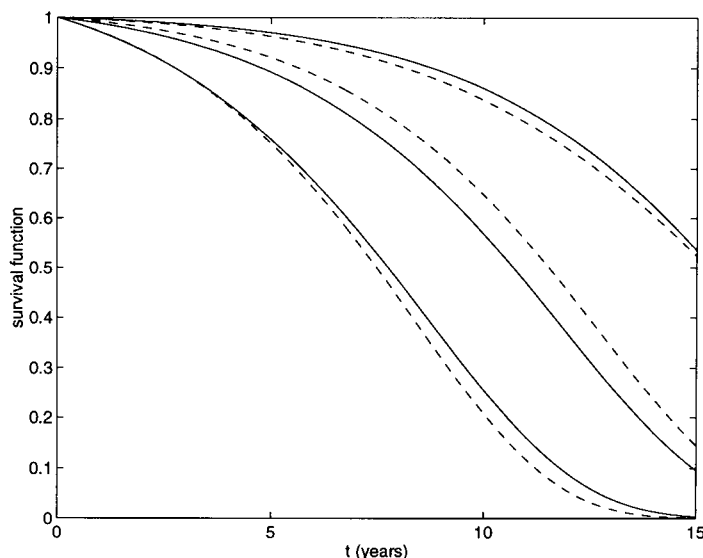


Figure 3. Conditional survival function of reoperation for specified age of patients. Constrained mixture model (—); latent failure-time approach (---). With each model:  $x = 20$  years (bottom line);  $x = 40$  years (middle line);  $x = 70$  years (top line)

### 5. CONDITIONAL PROBABILITY OF A REOPERATION

Up to now we have focused on the probability that a patient aged  $x$  years at the time of the initial replacement operation of the aortic valve by a xenograft prosthesis will have to undergo a re-replacement operation, which we refer to as reoperation. As a patient can avoid a reoperation by dying first, it is relevant to consider the conditional probability of a reoperation within a specified time  $t$  after the initial operation given that the patient does not die without a reoperation during this period. We shall denote this conditional probability by  $CP_R(t; x)$ .

In the latent failure-time framework, this probability is given by

$$\begin{aligned} CP_R(t; x) &= \text{pr}\{\text{reoperation by } t \mid \text{no death without reoperation by } t\} \\ &= \text{pr}\{\text{reoperation by } t\} / \{1 - \text{pr}(\text{death without reoperation by } t)\} \\ &= \frac{\int_0^t f_R(u; x) S_D(u; x) du}{1 - \int_0^t f_D(u; x) S_R(u; x) du}. \end{aligned} \quad (18)$$

It can be expressed in terms of the components of the mixture model (1) as

$$CP_R(t; x) = \pi_1(x) \{1 - S_1(t; x)\} / [\pi_1(x) + \pi_2(x) S_2(t; x)] \quad (19)$$

In terms of the constrained mixture model (15),  $CP_R(t; x)$  is given by

$$CP_R(t; \Psi, x) = \frac{\pi_1(x; \boldsymbol{\beta}) \{1 - S_1(t; \boldsymbol{\theta}_1, x)\}}{\pi_1(x; \boldsymbol{\beta}) + \pi_2(x; \boldsymbol{\beta}) S_1(t; \boldsymbol{\theta}_1, x) S_a(t; \boldsymbol{\theta}_a, x)}. \quad (20)$$

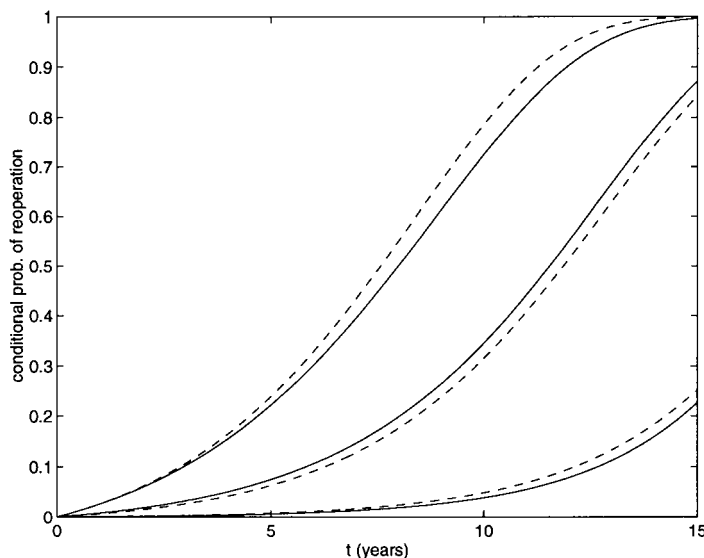


Figure 4. Conditional probability of reoperation for specified age of patients. Constrained mixture model (—); latent failure-time approach (---). With each model,  $x = 20$  years (top line),  $x = 40$  years (middle line),  $x = 70$  years (bottom line)

In Figure 4, we have plotted the estimated conditional probability of reoperation,  $CP_R(t; \hat{\Psi}, x)$ , versus  $t$  for various levels of the age  $x$  of the patient, along with the corresponding plots given by the latent failure-time approach applied under the assumption of independent competing risks. It can be seen that there is good agreement between this latter method and the constrained mixture model approach.

To illustrate further the close agreement between the estimates obtained with these two approaches, we computed the average value of the estimate of  $\pi_1(x)$  over the data set, namely,

$$\hat{\pi}_1 = \frac{1}{n} \sum_{j=1}^n \hat{\pi}_1(x_j).$$

It gave average estimated values of  $\hat{\pi}_1 = 0.323$  and  $0.368$ , for the constrained mixture approach and for the latent failure-time approach, respectively. In Figure 5, we plotted versus time  $t$  the average value of the estimate of  $CP_R(t; x)$  over the data set, given by

$$\overline{CP}_R(t) = \frac{1}{n} \sum_{j=1}^n \widehat{CP}_R(t; x_j),$$

for both the constrained mixture and latent failure-time approaches. It can be seen that there is very close agreement between the constrained and latent approaches over the whole range of  $t$ .

To investigate further the validity of the constrained mixture model (15), we calculated the residuals  $r_j$  defined as

$$r_j = H(t_j; \hat{\Psi}, x_j), (j = 1, \dots, n)$$

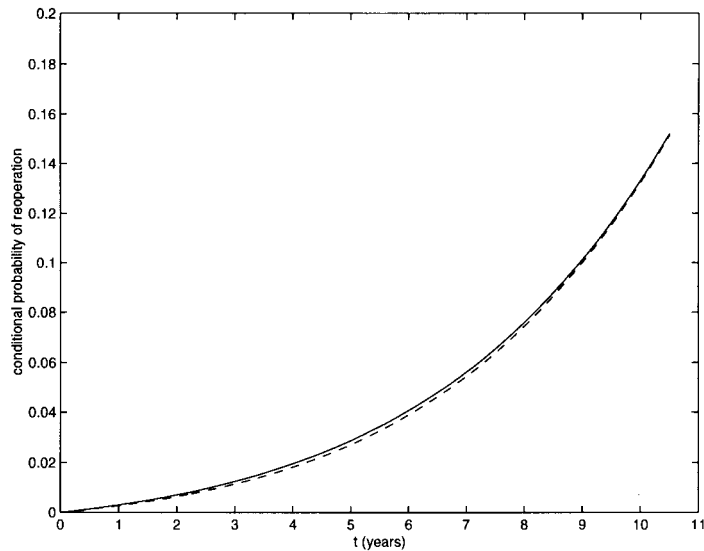


Figure 5. Estimated conditional probability of reoperation averaged over the data. Constrained mixture model (—); latent failure-time approach (---)

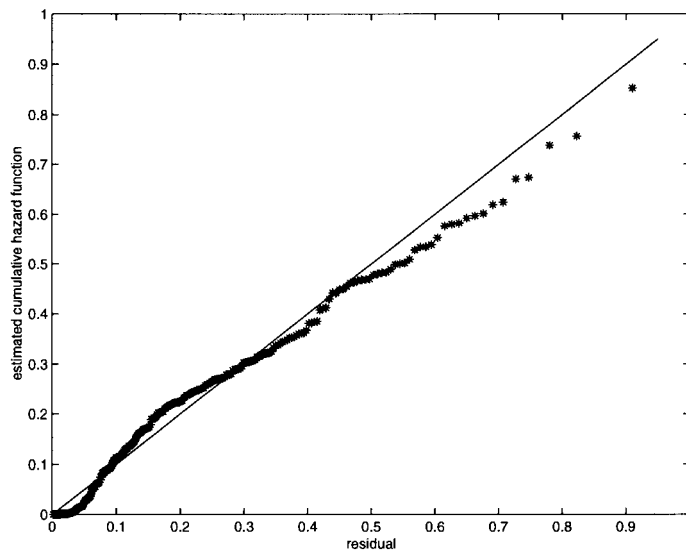


Figure 6. Plot of cumulative hazard function estimated from the residuals versus cumulative hazard function for the exponential distribution

If the model is correct and  $\hat{\Psi}$  is close to  $\Psi$ , then the residuals  $r_j$ s should behave like a censored sample from a unit exponential distribution. In Figure 6, we have plotted the Nelson-Aalen estimate of the cumulative hazard function formed from the residuals  $r_j$  versus  $r_1, \dots, r_n$ , along with the 45° line corresponding to the cumulative hazard function of the unit exponential distribution. It can be seen from this figure that the cumulative hazard function formed from the

residuals is quite close to the 45° line, except in the tail where the variability in the estimate of the cumulative hazard function is large. This suggests that the constrained mixture model is adequate.

## 6. DISCUSSION

We have seen that a mixture model with typically used parametric forms for its component survival functions is unable to provide a satisfactory model of the distribution of time to reoperation or to death without reoperation after the replacement of the aortic valve by a xenograft prosthesis. As a xenograft prosthesis will always need replacement if the patient were to live for a sufficiently long time after the initial operation, the two mixture model components corresponding to reoperation and death without a reoperation need to be inter-related in their parametric specification. We have handled this situation by constraining the components so that the component hazard function for death without a reoperation is always greater than the hazard for reoperation. This is effected implicitly by setting the component survival function corresponding to death without a reoperation to be equal to the component survival function for reoperation multiplied by some additional survival function. This constrained mixture model gives results similar to those obtained by the traditional latent failure-time approach applied under the commonly used assumption of independent competing risks. As this assumption would appear not to be a strong one in this application, it supports the validity of the constrained mixture model in this context. This constrained mixture model allows the component corresponding to the competing risk with the greater hazard to have a nonproportional hazard function, which is a more appropriate in the applications of the type here.

An attractive feature of the mixture model approach is that it does not have to make assumptions about the independence of the competing risks, as with the latent failure-time approach. As pointed by Lagakos (1979), assumption of independence of risks seems unreasonable and questionable in most real life situations. But any model that allows for dependence is restricted to have simple parametric form for the joint distribution of the latent times, for example, a bivariate normal mode, as proposed by Nádas (1971). Furthermore, the dependence of the competing risks is nontestable. Cox (1959, 1962, p. 112) and Tsiatis (1975) showed that for any joint distribution for the latent-failure times, there exists a joint distribution with independent latent-failure times that gives the same distribution of the observable failure times. Thus from the observable failure times and causes of failure alone, it is impossible to distinguish between an independent competing-risks model and an infinitude of dependent models. The indentifiability of the latent failure-time approach with covariates has been considered by Heckman and Honoré (1989) and Slud (1992). Klein and Moeschberger (1987) determined the effects of incorrect assumption and independence and concluded that the departures from independence were of great consequence on the estimations of the overall and marginal survival functions. In contrast to the latent failure-time approach, the postulated component survival functions and mixing proportions of the mixture model are able to be estimated directly from the observable data.

## REFERENCES

- Berkson, J. and Gage, R. P. (1952). 'Survival curve for cancer patients following treatment'. *Journal of the American Statistical Association* **47**, 501–515.

- Boag, J. W. (1949). 'Maximum likelihood estimates of the proportion of patients cured by cancer therapy' (with discussion). *Journal of the Royal Statistical Society B* **11**, 15–53.
- Cox, D. R. (1959). 'The analysis of exponentially distributed life-times with two types of failure'. *Journal of the Royal Statistical Society B* **21**, 411–421.
- Cox, D. R. (1962). *Renewal Theory*. London: Methuen.
- Cox, D. R. (1972). 'Regression models and life-tables' (with discussion). *Journal of the Royal Statistical Society B* **34**, 187–220.
- David, H. A. and Moeschberger, M. L. (1978). *The Theory of Competing Risks*. London: Griffin.
- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). 'Maximum likelihood from incomplete data via the EM algorithm'. *Journal of the Royal Statistical Society B* **39**, 1–38.
- Efron, B. (1979). 'Bootstrap methods: another look at the jackknife'. *Annals of Statistics* **7**, 1–26.
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: SIAM.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.
- Farewell, V. T. (1977). 'A model for a binary variable with time-censored observations'. *Biometrika* **64**, 43–46.
- Farewell, V. T. (1982). 'The use of mixture models for the analysis of survival data with long-term survivors'. *Biometrics* **38**, 1041–1046.
- Farewell, V. T. (1986). 'Mixture models in survival analysis: are they worth the risk?' *Canadian Journal of Statistics* **14**, 257–262.
- Golbeck, A. L. (1992). 'Bootstrapping current life table estimators'. In *Bootstrapping and Related Techniques*, eds. K. H. Jöckel, G. Rothe and W. Sandler Heidelberg: Springer, 197–201.
- Gordon, N. H. (1990a). 'Application of the theory of finite mixtures for the estimation of "cure" rates of treated cancer patients'. *Statistics in Medicine* **9**, 397–407.
- Gordon, N. H. (1990b). 'Maximum likelihood estimation for mixture of two Gompertz distributions when censoring occurs'. *Communications in Statistics – Simulation and Computation* **19**, 737–747.
- Grunkemeier, G. L., Jamieson, W. R. E., Miller, D. C. and Starr, A. (1994). 'Actuarial versus actual risk of porcine structural valve deterioration'. *Journal of Thoracic and Cardiovascular Surgery* **108**, 709–718.
- Heckman, J. J. and Honoré, B. E. (1989). 'The identifiability of the competing risks model'. *Biometrika* **76**, 325–330.
- Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Klein, J. P. and Moeschberger, M. L. (1987). 'Independent or dependent competing risks: Does it make a difference?' *Communications in Statistics – Simulation and Computation* **16**, 507–533.
- Kuk, A. Y. C. (1992). 'A semiparametric mixture model for the analysis of competing risks data'. *Australian Journal of Statistics* **34**, 169–180.
- Lagakos, S. W. (1979). 'General right censoring and its impact on the analysis of survival data'. *Biometrics* **35**, 139–156.
- Larson, M. G. and Dinse, G. E. (1985). 'A mixture model for the regression analysis of competing risks data'. *Applied Statistics* **34**, 201–211.
- Maller, R. A. and Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. New York: Wiley.
- McGiffin, D. C., O'Brien, M. F., Galbraith, A. J., McLachlan, G. J., Stafford, E. G., Gardiner, M. A. H., Pohlner, P. G., Early, L. and Kear, L. (1993). 'An analysis of risk factors for death and mode-specific death following aortic valve replacement using allograft, xenograft and mechanical valves'. *Journal of Thoracic and Cardiovascular Surgery* **106**, 895–911.
- McGiffin, D. C., Galbraith, A. J., O'Brien, M. F., McLachlan, G. J., Naftel, D. C., Adams, P., Reddy, S. and Early, L. (1997). 'An analysis of valve re-replacement after aortic valve replacement with biologic devices'. *Journal of Thoracic and Cardiovascular Surgery* **113**, 311–318.
- McLachlan, G. J. and Basford, K. E. (1988). *Mixture Models: Inference and Applications of Clustering*. New York: Marcel Dekker.
- McLachlan, G. J. and Krishnan, T. (1997). *The EM Algorithm and Extensions*. New York: Wiley.
- McLachlan, G. J., Ng, S. K., Adams, P., McGiffin, D. C. and Galbraith, A. J. (1997). 'An algorithm for fitting mixtures of Gompertz distributions to censored survival data'. *Journal of Statistical Software* **2**.
- Meng, X. L. (1994). 'On the rate of convergence of the ECM algorithm'. *Annals of Statistics* **22**, 326–339.
- Meng, X. L. and Rubin, D. R. (1993). 'Maximum likelihood estimation via the ECM algorithm: a general framework'. *Biometrika* **80**, 267–278.

- Moeschberger, M. L. and David, H. A. (1971). 'Life tests under competing causes of failure and the theory of competing risks'. *Biometrics* **27**, 909–933.
- Nádas, A. (1971). 'The distribution of the identified minimum of a normal pair determines the distribution of the pair'. *Technometrics* **13**, 201–202.
- Slud, E. (1992). 'Nonparametric identifiability of marginal survival distributions in the presence of dependent competing risks and a prognostic covariate'. In *Survival Analysis: State of the Art*, eds. J. P. Klein and P. K. Goel Dordrecht: Kluwer, 355–368.
- Tsatis, A. (1975). 'A non-identifiability aspect of the problem of competing risks'. *Proceedings of the National Academy of Sciences, USA* **72**, 20–22.