STATISTICS IN MEDICINE
Statist. Med. 2000; **19**:563–574

Choice of conditional models in bivariate survival

Robin Henderson[1,*] and Helen Prince[2]

[1] *Department of Mathematics and Statistics, University of Lancaster, Lancaster LA1 4YF, U.K.*
[2] *Department of Statistics, Newcastle University, Newcastle-upon-Tyne NE1 7RU, U.K.*

SUMMARY

We consider bivariate survival problems in which interest is in the conditional distribution of one survival variable given an uncensored observation of the other. The work is motivated by an analysis of time to cancer diagnosis then subsequent survival amongst a group of organ transplant recipients. The effect of conditioning is illustrated for five standard bivariate models. The consequences of adopting a misspecified marginal approach in which the conditioning variable is considered to be a fixed covariate are investigated. Copyright © 2000 John Wiley & Sons, Ltd.

1. INTRODUCTION

Incidence rates of various types of cancer, particularly lymphoma and Kaposi's sarcoma, can be very high amongst solid organ transplant recipients [1, 2]. In this paper we explore the relationship between two time measurements made on such patients: first the time from transplantation to diagnosis of cancer, and second the subsequent time from diagnosis to death. Most recipients are never diagnosed as having cancer, as although the relative incidence rates are very high the absolute rates are still low. However, for those who *are* so diagnosed it would be useful to know whether knowledge of the time to diagnosis holds any prognostic information for subsequent survival. Thus we have a bivariate survival problem with interest mainly in the conditional distribution of the second survival time T_2 given the value of the first, T_1.

Day *et al.* [3] consider a similar problem: the use of binary (present/absent) biological markers for prediction. Times to marker and event form bivariate survival, and interest is in the conditional distribution of time to event given the appearance of marker. This is close to our problem but an important difference is that they assume times are not sequential, that is, the event may occur before the marker. In contrast, in our application the events are serial, with T_2 by definition beginning when T_1 is observed. In addition we are interested only in the conditional distribution of T_2 given the exact value of T_1, not in a condition of the form $T_1 > t$.

Perhaps the most obvious and natural approach is to obtain a suitable bivariate survival model for (T_1, T_2) and then derive the associated conditionals. A simpler alternative might be to take a

*Correspondence to: Robin Henderson, Department of Mathematics and Statistics, University of Lancaster, Lancaster LA1 4YF, U.K.

CCC 0277–6715/2000/040563–12$17.50
Copyright © 2000 John Wiley & Sons, Ltd.

marginal approach, modelling T_2 directly, with T_1 considered as if it were a fixed covariate, as in for instance an approach of Gail *et al.* [4] to the analysis of multiple tumour times. In this work we compare these methods, beginning in Section 2 with a brief summary of some common methods of modelling bivariate survival data. The associated conditional distributions are then illustrated in Section 3 and the consequences of adopting a misspecified marginal approach are discussed in Section 4. In Section 5 we apply the techniques to data on a group of 634 patients who developed lymphoma or Kaposi's sarcoma following organ transplantation. The paper is completed by some comments and suggestions for future work in Section 6.

2. BIVARIATE SURVIVAL

Clearly almost any bivariate distribution can be adapted to survival through a suitable transformation of T_1 and T_2, though rather few seem to be used in practice. We will concentrate on the five different models summarized in Table I and described below.

2.1. Model 1: bivariate log-normal

The bivariate log-normal distribution (with $\log(T_1), \log(T_2)$ bivariate normal) may seem a natural and simple first choice but seems to have had little use in practice recently, in biostatistical applications at least. We will denote the parameters on the log scale as usual as $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ. Covariates are assumed to affect survival through μ_1 and μ_2, usually linearly.

2.2. Models 2 and 3: frailty

A more common method of modelling bivariate or multivariate survival is to adopt a frailty model [5–10]. We assume that T_1 and T_2 are conditionally independent given the value of some shared unobservable random effect or frailty Z, which acts multiplicatively on the hazard functions. If Z is large then both hazards are high and T_1 and T_2 both tend to be low. If Z is small then both hazards are low and T_1 and T_2 both tend to be large. Thus the frailty induces a positive association between responses.

Semi-parametric frailty models are in widespread use but here we prefer a parametric approach, assuming Weibull survival conditional on Z. Thus, for $i = 1, 2$ we assume

$$S_i(t \mid z) = \exp\{-z\lambda_i t^{\delta_i}\}$$

with covariate effects modelled through λ_1 and λ_2, usually log-linearly.

There are various choices available for the form of frailty distribution. Aalen [8] describes a useful general class, perhaps the two most important of which, or at least most commonly used, are gamma and positive stable.

Table I. Five bivariate survival models.

Model 1	Bivariate log-normal
Model 2	Weibull conditional distributions given gamma frailty
Model 3	Weibull conditional distributions given positive stable frailty
Model 4	Oakes' model with $c < 1$ and Weibull marginals
Model 5	Oakes' model with $c < 1$ and Burr marginals

Copyright © 2000 John Wiley & Sons, Ltd.

For model 2 we assume gamma frailty with (for identifiability) mean one, that is, $Z \sim \Gamma(\xi, \xi)$. An advantage of this model is that explicit expressions for the joint and marginal survivor functions can be obtained when the conditioning on the frailty is removed. The resulting distributions are of Burr form, for example the joint survivor function is

$$S(t_1, t_2) = \left(\frac{\xi}{\xi + \lambda_1 t_1^{\delta_1} + \lambda_2 t_2^{\delta_2}} \right)^{\xi}$$

and the marginal survivor functions for $i = 1, 2$ are

$$S_i(t) = \left(\frac{\xi}{\xi + \lambda_i t^{\delta_i}} \right)^{\xi}. \tag{1}$$

A disadvantage of the gamma choice is that the marginal distributions are no longer of proportional hazards form. This property is retained in the marginals if a positive stable distribution is assumed for the frailty terms. Hence for model 3 we retain the Weibull assumption given the frailty Z, but assume that Z has a positive stable distribution with parameter α $(0 < \alpha \leqslant 1)$, most easily defined through its Laplace transform $E[\exp\{-uZ\}] = \exp\{-u^{\alpha}\}$ [11]. In that case the joint survivor function becomes

$$S(t_1, t_2) = \exp\{-(\lambda_1 t_1^{\delta_1} + \lambda_2 t_2^{\delta_2})^{\alpha}\}$$

with marginals

$$S_i(t) = \exp\{-(\lambda_i t^{\delta_i})^{\alpha}\}$$

which are of Weibull (and hence proportional hazards) form.

The exact choice of frailty may not be important; Pickles and Crouchley [9] found that fitted mixture distributions were robust to choice of mixture (frailty) distribution.

2.3. Models 4 and 5: negative association

Usually a single frailty term is used to describe the association between responses in each cluster of cases. Crouchley and Pickles [12] and Xue and Brookmeyer [13], amongst others, prefer multiple component frailty models, where separate responses within the same cluster may have distinct, though not independent, associated frailties. Such a technique has the advantage that it can be used to model *negative* association between T_1 and T_2, by allowing their associated frailty terms to be negatively correlated. Negative correlation can also be accommodated within a bivariate model described by Oakes [14, 15], following earlier work by Clayton [16]. The model assumes that the marginal survivor functions S_1 and S_2 are specified and the degree of association is controlled through a parameter c, with independence at $c = 1$. If we restrict attention to $0 < c < 1$ then there is negative association and the joint survivor function is

$$S(t_1, t_2) = (\max[\{S_1(t_1)\}^{1-c} + \{S_2(t_2)\}^{1-c} - 1, 0])^{1/(1-c)}.$$

Model 4 is defined by taking this joint survivor function with both survival times assumed to have Burr marginal distributions, as in (1). Here ξ is now a free parameter to be estimated, rather than a property of an underlying frailty distribution. Model 5 has marginal distributions of Weibull

Copyright © 2000 John Wiley & Sons, Ltd.

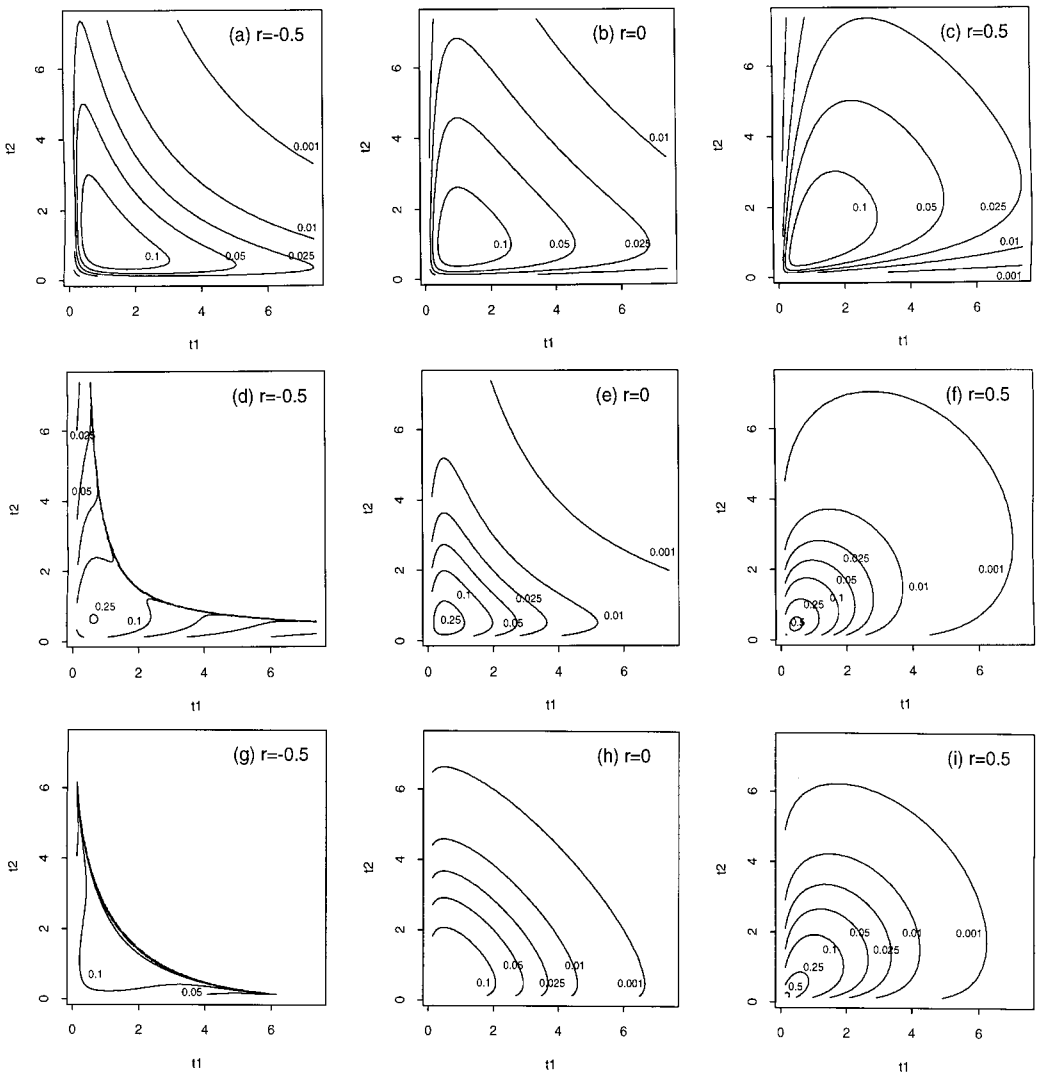

Figure 1. Bivariate densities: (a)–(c) log-normal marginals; (d)–(f) Burr marginals; (g)–(i) Weibull marginals. Correlations refer to log scale.

form. In both cases we model covariate effects through the scale parameters λ_i in the same way as models 2 and 3.

2.4. Illustration

Figure 1 illustrates the similarities and differences between the bivariate densities for the five models, in the absence of covariates. Correlations between $\log(T_1)$ and $\log(T_2)$ are -0.5 in the left-hand column, zero in the central column, and $+0.5$ in the right-hand column. Across the top

Copyright © 2000 John Wiley & Sons, Ltd.

row all marginals are log-normal, in the second row they are Burr, and in the third row Weibull. Plots (a)–(c) refer to model 1, (d) to model 4, (f) to model 2, (g) to model 5 and (i) to model 3, with the independence cases (e) and (h) included for reference. We scale so that in each plot the marginal distributions of T_1 and T_2 are the same, standardized to zero mean and unit variance on the log scale. Together with the required correlations this standardization uniquely defines the log-normal and Weibull parameters, but not Burr in plots (d) and (e). In those we select $\xi = 1$ for comparability with (f) (where $\xi = 1$ gives the required correlation of 0.5). Note that under the Oakes models (4 and 5) the bivariate density becomes concentrated along the curve $S_1 + S_2 = 1$ as the degree of negative correlation increases, and also that observations are always confined to the region $S_1 + S_2 \geqslant 1$. The concentration is noticeably more severe when the marginals are Weibull rather than Burr, whereas when there is positive correlation (plots (f) and (i)) the two bivariate contour plots are more similar. Both show more concentration near the origin than the bivariate log-normal distribution with the same correlation between $\log(T_1)$ and $\log(T_2)$.

3. CONDITIONAL DISTRIBUTIONS

We now turn to the conditional distribution of T_2 given an uncensored observation of T_1. The problem of censored T_1 will be considered later. Conditional survival distributions are given in Table II. The form for model 3 (positive stable frailty) was given by Hougaard [11] and the others are either standard or easily derived.

Figure 2 illustrates these using the same distributions and parameter values as in the outer columns of Figure 1. Each plot shows the marginal survival function of T_2 and the conditional survival functions given both $\log(T_1) = -2$ and $\log(T_1) = +2$. Given the scaling to unit variance for $\log(T_1)$ these represent fairly extreme values.

When there is negative correlation (left-hand column) we note the small variance in the conditional distribution of T_2 given large T_1 under the Oakes models (4 and 5), resulting from the restriction to the region $S_1 + S_2 \geqslant 1$. When there is positive correlation (right-hand column) the large variability in T_2 given large T_1 under model 2, gamma frailty, is worth noting, as is the high probability of T_2 being very small given small T_1 under model 3, positive stable frailty.

Table II. Conditional survival functions.

Model	$S(t_2 \mid T_1 = t_1)$
1	$1 - \Phi((t_2 - \mu_2 - \rho(\sigma_2/\sigma_1)(\log(t_1) - \mu_1))/\sigma_2\sqrt{\{1-\rho^2\}})$
2	$\left(\dfrac{\xi + \lambda_1 t_1^{\delta_1}}{\xi + \lambda_1 t_1^{\delta_1} + \lambda_2 t_2^{\delta_2}}\right)^{\xi}$
3	$\left(\dfrac{\lambda_1 t_1^{\delta_1}}{\lambda_1 t_1^{\delta_1} + \lambda_2 t_2^{\delta_2}}\right) \exp\{(\lambda_1 t_1^{\delta_1})^{\alpha} - (\lambda_1 t_1^{\delta_1} + \lambda_2 t_2^{\delta_2})^{\alpha}\}$
4, 5	$\{S(t_1, t_2)/S_1(t_1)\}^c$

Copyright © 2000 John Wiley & Sons, Ltd.

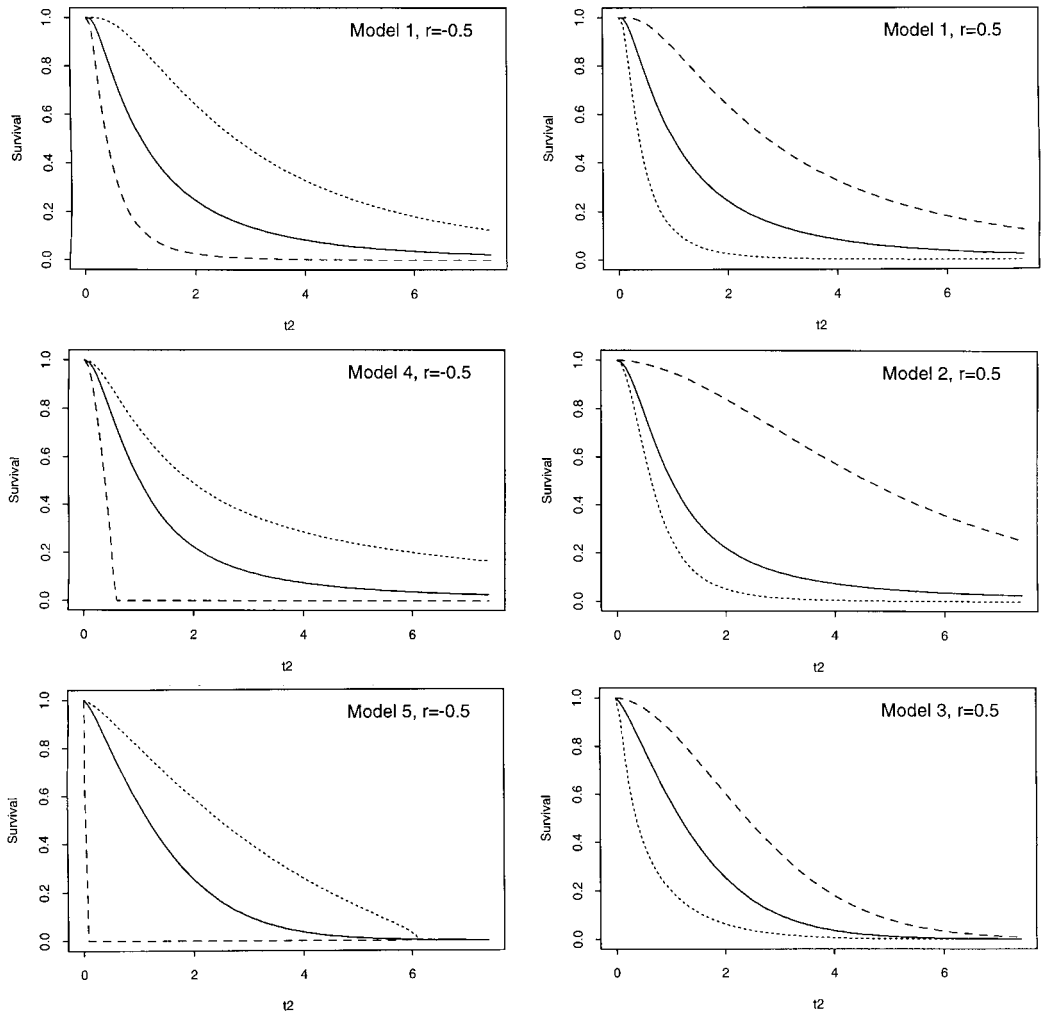

Figure 2. Conditional survival curves for T_2 given $\log(T_1) = -2$ (dotted line), $\log(T_1) = 2$ (dashed line) and unconditional (solid line).

4. MISSPECIFICATION

In principle, in practical applications the most appropriate bivariate model can be fitted and then conditional distributions can be derived as in the previous section. In practice however this can be quite cumbersome. If interest is solely in the conditional distribution of T_2 given T_1 then a much simpler approach is to model that conditional distribution directly, with T_1 (or more likely $\log(T_1)$) treated as a fixed covariate rather than a given value of a correlated response. If the conditional distribution is correctly specified then the two appoaches will lead to the same conclusions. The true conditional distribution can be a rather complicated function of T_1, however, and if simpler, more standard survival functions can provide good approximations then their use could be recommended.

Copyright © 2000 John Wiley & Sons, Ltd.

4.1. Model 1

As seen in Table II, for this model t_1 appears in the conditional distribution of $T_2 \mid T_1 = t_1$ only through a linear term in $\log(t_1)$ in the mean log survival time. Since the standard method of incorporating covariates in a log-normal analysis is to assume a linear predictor for the mean log lifetime, inclusion of $\log(t_1)$ as a covariate in a standard analysis will lead to the correct model and inferences, and there is no need for a bivariate analysis.

4.2. Models 2 and 3

Both of these models assume frailty effects acting multiplicatively on conditionally independent Weibull hazards, that is, if covariates are included

$$S_i(t \mid \boldsymbol{x}, z) = \exp\{-z\mathrm{e}^{\beta_i \boldsymbol{x}} t^{\delta_i}\}$$

for $i = 1, 2$, and assuming $\boldsymbol{x}$ includes an intercept term. It is of interest to investigate the consequences of simply fitting a Weibull distribution to T_2 with $\log(T_1)$ included as a covariate. Thus we might fit to T_2 a misspecified model of the form

$$S^{\mathrm{M}}(t \mid \boldsymbol{x}, t_1) = \exp\{-\mathrm{e}^{\beta \boldsymbol{x} + \gamma \log(t_1)} t^{\delta}\}.$$

Henderson [17] considers asymptotic properties of the resulting estimators $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\delta}$ for the gamma/Weibull model (model 2). His results can in fact be generalized to any frailty mixture of Weibulls. Assuming that, intercept apart, covariates are scaled to zero mean, it can be shown that for uncensored T_2 as sample size increases the estimators converge to β^*, γ^* and δ^* satisfying

$$\frac{E[Z^{(1-2\delta^*/\delta_2)} \log Z]}{E[Z^{(1-2\delta^*/\delta_2)}]} - E[\log Z] = C - \psi(\delta^*/\delta_2)$$

$$\gamma^*/\delta_1 = \delta^*/\delta_2 - 1$$

and

$$\beta^* = (\gamma^*/\delta_1)\beta_1 + (\delta^*/\delta_2)\beta_2 + \log\{E[Z^{1-2\delta^*/\delta_2}]\Gamma(\gamma^*/\delta_1 + 1)\Gamma(\delta^*/\delta_2 + 1)\}\boldsymbol{j}$$

where C is Euler's constant, $\psi(.)$ is the digamma function, and $\boldsymbol{j} = (1, 0, \ldots, 0)'$. When there is no frailty ($Z \equiv 1$) there is no misspecification and so the asymptotic fitted shape parameter δ^* is equal to the true shape parameter δ_2, the asymptotic fitted coefficient γ^* of $\log(t_1)$ is zero since the conditioning has no effect, and the asymptotic fitted regression coefficients β^* are equal to the true coefficients β_2. As the amount of frailty increases δ^* converges towards $0.5\delta_2$ and γ^* converges towards $-0.5\delta_1$, whilst the regression coefficients β^* move away from β_2 and towards $0.5(\beta_2 - \beta_1)$. It is interesting to note that the asymptotic regression coefficient β^* will be close to zero when frailty effects are large and $\beta_1 = \beta_2$.

4.3. Models 4 and 5

Interest is again in the consequences of fitting a misspecified Weibull distribution with $\log(t_1)$ treated as a covariate. For these models the asymptotic values δ^*, γ^* and β^* cannot be obtained algebraically and hence we have relied on numerical methods in the illustration below.

Copyright © 2000 John Wiley & Sons, Ltd.

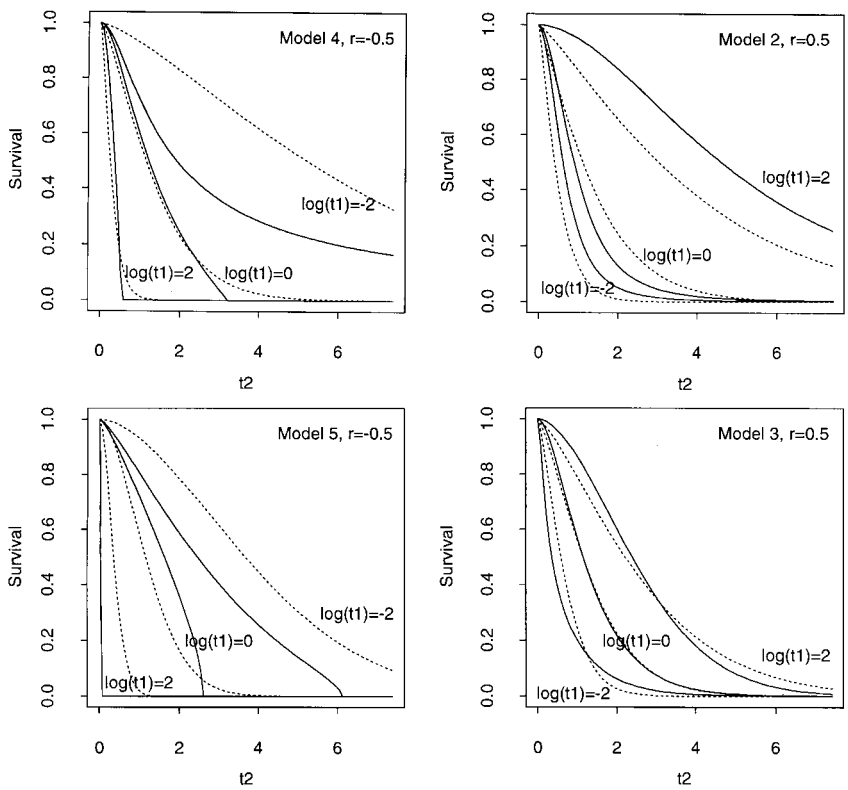

Figure 3. True (solid line) and misspecified (dotted line) conditional survival curves.

4.4. Illustration

Figure 3 compares true conditional survival curves for models 2–5 with those which would be obtained from a misspecified Weibull fit to uncensored T_2 with $\log(T_1)$ treated as a covariate. Scaling is the same as for the previous figures. Each plot shows three pairs of curves, corresponding to true and misspecified fits at $\log(T_1) = -2, 0$ and $+2$. Under model 2, the absolute difference between the true and misspecified fits is small when the conditioning is on small $\log(T_1)$, but there is more apparent difference when $\log(T_1)$ is large and the survival rate is higher. This is a consequence of attempting to fit a thin-tailed distribution (Weibull) to data arising from a heavy-tailed distribution (Burr). The relative error arising from the misspecification is less dependent on the conditioning value, however. Under model 3 both true and misspecified models are thin tailed and the misspecified fit is always close to the true for all three illustrations given. Recall that this model assumes positive stable frailty and has Weibull marginal distributions, but note that the true conditional distribution of T_2 given T_1 is *not* Weibull (Table II), explaining the small differences. Model 4, with negative association, leads to reasonably close true and misspecified conditional survival curves when T_1 is not low, but a very poor fit for the misspecified model when T_1 *is* low, with the misspecified fit considerably overestimating early survival rate. Model 5 leads to rather poor fits for all T_1, with the misspecified univariate fit tending to overestimate survival when we condition on an extreme value of T_1.

Copyright © 2000 John Wiley & Sons, Ltd.

In the above we assumed that T_2 is uncensored. Simulation results (not shown) with both random and type I censorship indicate that if censoring is allowed the misspecified distributions are only relatively slightly affected unless the conditioning has the effect of increasing survival rate, that is, high T_1 for models 2 and 3, low T_1 for models 4 and 5. In that case the fitted distributions can vary considerably according to the amount of censoring in T_2. There is no consistent pattern across all models however; censoring can lead to either increased or decreased fitted conditional survival rates depending on the model in use. For reference, positive stable simulations were obtained using a method given by Chambers *et al.* [18].

5. APPLICATION

We now turn to the application which motivated this work, the relationship between time T_1 from transplant to cancer diagnosis then time T_2 from diagnosis to death for a group of organ transplant patients. Data were provided by the Collaborative Transplant Study, based at Heidelberg, which collates details of transplant operations and outcomes from over 300 centres worldwide. Only those centres which provided written confirmation that their malignancy reports were complete were included in this study.

We consider a subset of the data consisting of results for 634 patients who developed lymphoma or Kaposi's sarcoma and who received heart or kidney transplants at centres in Western Europe or North America. Some 45 per cent of these patients died within the follow-up period, all due to cancer, leaving 55 per cent with censored T_2 values. Since the subset is confined to patients who developed cancer, T_1 is always known, never censored. Had follow-up been longer some additional patients may have developed the diseases of interest and we have to take this into account in the analyses. Fortunately this is straightforward since for each of the 634 patients the *potential* follow-up period f between transplant and study date is available. Thus we can obtain valid inference by conditioning on T_1 occurring within the follow-up period. With δ_2 indicating whether T_2 is observed ($\delta_2 = 1$) or censored ($\delta_2 = 0$) the likelihood contributions are thus of the form

$$\left(\frac{-\partial S(t_1, t_2)}{\partial t_1}\right)^{1-\delta_2} \left(\frac{-\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2}\right)^{\delta_2} (1 - S_1(f))^{-1}.$$

Five covariates were considered: organ transplanted (heart or kidney); age; sex; use of cyclosporin-A for immunosuppression, and ethnic group. These were allowed for in all the analyses summarized below, though results are not presented, since in the present work we are interested mainly in the relationship between T_1 and T_2. We note however that conclusions with respect to these covariates were reasonably consistent across all the models, though statistical significance tended to be less strong under the bivariate log-normal model than the others.

Table III shows the maximized log-likelihoods under the five bivariate models, together with that for a univariate Weibull fit to T_2 with $\log(T_1)$ included as a covariate. Also shown are the corresponding maximized log-likelihoods under an independence assumption for the bivariate models, and for the last row a univariate fit to T_2 without T_1. There is highly significant improvement in maximized log-likelihoods for models 4 and 5 against the independence alternative, and in the univariate analysis when $\log(T_1)$ is included. The coefficient of the latter is estimated to be 0.191 ($\mathrm{SE} = 0.056$) implying increased hazard at high T_1 and hence negative association. This is also the conclusion from the rejection of independence under models 4 and 5, which are appropriate

Copyright © 2000 John Wiley & Sons, Ltd.

Table III. Maximized log-likelihoods.

Model	Bivariate	Independent
1 Log-normal	−1603.48	−1605.39
2 Gamma frailty	−1642.11	−1642.11
3 Positive stable frailty	−1642.11	−1642.11
4 Oakes/Burr	−1630.90	−1639.36
5 Oakes/Weibull	−1633.73	−1642.11
	With $\log(T_1)$	Without $\log(T_1)$
Univariate Weibull	− 640.09	− 646.07

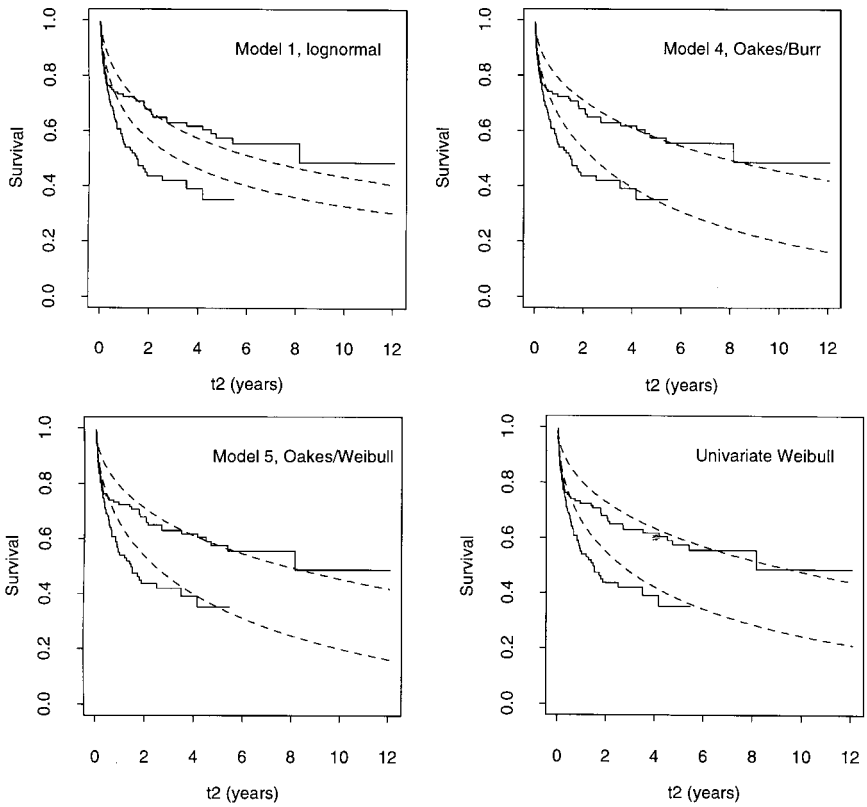

Figure 4. Kaplan–Meier survival plots for patients with 20 per cent highest and 20 per cent lowest T_1 values, together with fitted conditional survival curves assuming mean covariates and mean T_1 for each group.

for negatively associated responses. Estimates of the parameter c are 0.638 and 0.650, respectively. There is less convincing evidence of non-independence under model 1, log-normal, with estimated correlation $\hat{\rho} = -0.129$ (SE $= 0.077$). The log-likelihoods for models 2 and 3 are maximized at the independence boundaries, as expected if there is negative association since they are both appropriate for positive association only, and so neither are considered further.

It is not possible to illustrate the fit of the conditional distributions at any single choice of T_1 since few patients have common T_1. Instead, in each plot of Figure 4 we show observed survival

Copyright © 2000 John Wiley & Sons, Ltd.

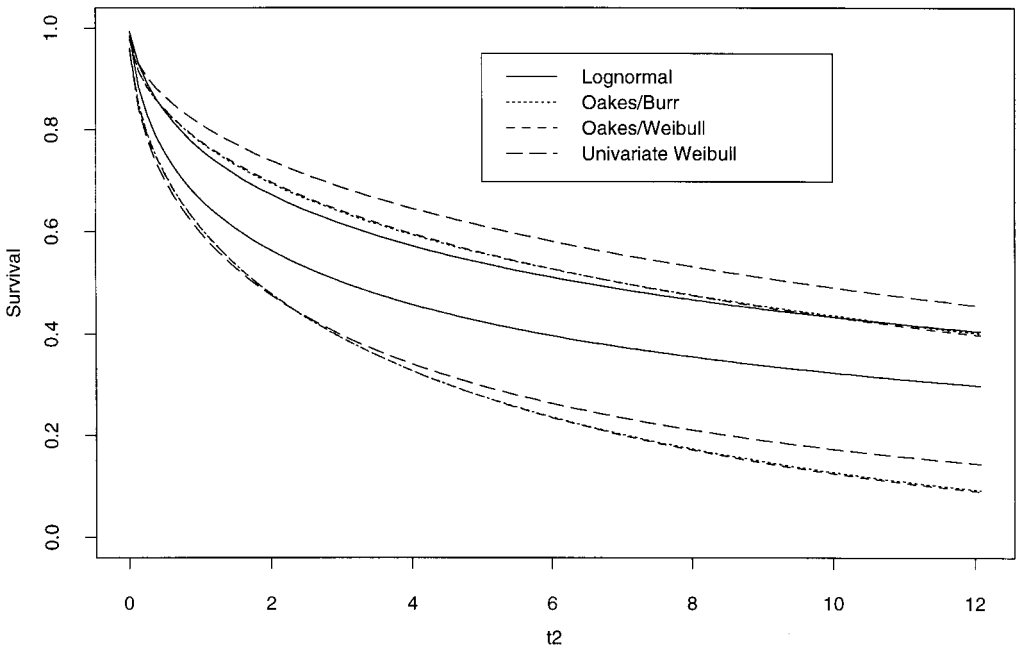

Figure 5. Estimated conditional survival curves. Two curves for each model: upper at the 5 per cent point of the T_1 distribution, lower at the 95 per cent point.

rates for the group of patients with the lowest 20 per cent of T_1 values, and for the group with 20 per cent highest T_1 values. This confirms the negative association; the lower fitted survival curve corresponds to the group with higher T_1. Each plot also includes fitted curves at the mean covariate value for each group, and conditional on T_1 taking the mean observed value for the group. The two versions of Oakes' model seem to fit the data well, but the log-normal fit underestimates the separation between the groups. The univariate fit seems reasonable, though early survival rate may be overestimated slightly more than under the Oakes' models.

Figure 5 illustrates the fitted conditional distributions at more extreme T_1 values, with predicted survival curves at the 5 per cent and 95 per cent quantiles of the observed distribution of T_1, which are 0.15 and 6.67 years, respectively. Conclusions are consistent with Figure 4; under the log-normal model T_1 has less effect than under the Oakes' models, which are indistinguishable in this plot. The univariate fit with T_1 treated as a covariate leads to higher conditional survival estimates over this time range, especially when T_1 is low so conditional survival is high. Note that this is consistent with the misspecification results of Section 4 for the negative correlation models (4 and 5).

6. DISCUSSION

There can be considerable differences in the shapes of conditional survival curves even when fitted marginal distributions and estimated correlations are similar. Hence it is, as always, very important to check the validity of a proposed model before any inferences are drawn. How best to do this

Copyright © 2000 John Wiley & Sons, Ltd.

for conditional survival remains an open question, complicated by the observation that in practice there are likely to be few patients with the same value of T_1, and also because a misspecified fit can be adequate for certain values of T_1 but not others, as seen in Figure 3.

We concentrated on parametric modelling in this work so that true conditional distributions could be compared. Clearly in practice semi-parametric models might be preferred, especially when treating T_1 as a covariate in a univariate analysis of T_2. Here a standard semi-parametric proportional hazards model might be selected rather than Weibull, but this has not been pursued in this work, partly because none of the true conditional distributions considered are themselves of proportional hazards form. We note however that there *are* bivariate distributions in which conditionals do have proportional hazards structure [19], whether the conditioning is on observed or censored T_1. In those cases there need be no misspecification in a univariate analysis, just as for the log-normal model considered above.

ACKNOWLEDGEMENTS

We thank Professor Gerhard Opelz, Director of the Collaborative Transplant Study, for providing the data analysed in Section 5, and Professor Mitchell Gail and two anonymous reviewers for helpful comments on an earlier version of this paper.

REFERENCES

1. Penn I. Cancer is a complication of severe immunosuppression. *Surgery in Gynecology and Obstetrics* 1986; **162**: 603–610.
2. Opelz G, Henderson R. Incidence of Non-Hodgkin's Lymphoma in kidney and heart transplant recipients. *Lancet* 1993; **342**: 1514–1516.
3. Day R, Bryant J, Lefkopoulou M. Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators. *Biometrika* 1997; **84**: 45–56.
4. Gail MH, Santner TJ, Brown CC. An analysis of comparative carcinogenesis experiments based on multiple times to tumour. *Biometrics* 1980; **36**: 255–266.
5. Hougaard P, Harvald B, Holm NV. Measuring the similarities between the lifetimes of adult Danish twins born between 1881–1930. *Journal of the American Statistical Association* 1992; **87**: 17–24.
6. Hougaard P. Frailty models for survival data. *Lifetime Data Analysis* 1995; **1**: 255–273.
7. Oakes D. Frailty models for multiple event times. In *Survival Analysis*: *State of the Art*, Klein JP, Goel PK (eds). Kluwer, 1992.
8. Aalen OO. Effects of frailty in survival analysis. *Statistical Methods in Medical Research* 1994; **3**: 227–243.
9. Pickles A, Crouchley R. A comparison of frailty models for multivariate survival data. *Statistics in Medicine* 1995; **14**: 1447–1461.
10. Aalen OO, Bjertness E, Sønju T. Analysis of dependent survival data applied to lifetimes of amalgam fillings. *Statistics in Medicine* 1995; **14**: 1819–1829.
11. Hougaard P. A class of multivariate failure time distributions. *Biometrika* 1986; **73**: 671–678.
12. Crouchley R, Pickles AR. Multivariate survival models for repeated and correlated events. *Journal of Statistical Planning and Inference* 1995; **47**: 95–110.
13. Xue X, Brookmeyer R. Bivariate frailty model for the analysis of multivariate survival time. *Lifetime Data Analysis* 1996; **2**: 277–289.
14. Oakes D. A model for association in bivariate survival. *Journal of Royal Statistical Society*, *Series B* 1982; **44**: 414–422.
15. Oakes D. Bivariate survival models induced by failures. *Journal of the American Statistical Association* 1989; **84**: 487–493.
16. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 1978; **65**: 141–151.
17. Henderson R. Modelling conditional distributions in bivariate survival. *Lifetime Data Analysis* 1996; **2**: 241–259.
18. Chambers JM, Mallows CL, Stuck BW. A method for simulating stable random variables. *Journal of the American Statistical Association* 1976; **71**: 340–344.
19. Arnold BC, Kim YH. Conditional proportional hazards models. In *Lifetime Data*: *Models in Reliability and Survival Analysis*, Jewell NP, Kimber AC, Lee, M-LT, Whittmore GA (eds). Kluwer: Dordrecht, 1996.

Copyright © 2000 John Wiley & Sons, Ltd.