Lifetime Data Analysis, 8, 177–203, 2002

Parametric Modeling for Survival with Competing Risks and Masked Failure Causes

BETTY J. FLEHINGER[1]
IBM Research Division, Mathematical Sciences Department, Thomas J. Watson Research Ctr., P.O. Box 218, Yorktown Heights, NY 10598, USA

BENJAMIN REISER
Department of Statistics, University of Haifa, Haifa, Israel

EMMANUEL YASHCHIN
IBM Research Division, Mathematical Sciences Department, Thomas J. Watson Research Ctr., P.O. Box 218, Yorktown Heights, NY 10598, USA

Received September 12, 2000; Revised February 5, 2001; Accepted April 10, 2001

Abstract. We consider a life testing situation in which systems are subject to failure from independent competing risks. Following a failure, immediate (stage-1) procedures are used in an attempt to reach a definitive diagnosis. If these procedures fail to result in a diagnosis, this phenomenon is called masking. Stage-2 procedures, such as failure analysis or autopsy, provide definitive diagnosis for a sample of the masked cases. We show how stage-1 and stage-2 information can be combined to provide statistical inference about (a) survival functions of the individual risks, (b) the proportions of failures associated with individual risks and (c) probability, for a specified masked case, that each of the masked competing risks is responsible for the failure. Our development is based on parametric distributional assumptions and the special case for which the failure times for the competing risks have a Weibull distribution is discussed in detail.

Keywords: life testing, masking, reliability, 2-stage experimentation, Weibull distribution

1. Introduction

System life data is often used to estimate component reliability. Consider a series system of k components operating independently. Typically the data consist of system lifelength (failure or censoring time) and identification of the failing component. Analysis follows standard competing risk methodology. Frequently, however, due to such factors as lack of proper diagnostic equipment or cost and time constraints, the failed component may be unknown or masked for some of the failed systems. Often the failure analysis will restrict the failure cause to a subset of components.

The same problem can arise in biomedical settings concerned with survival. In clinical trials and epidemiological studies it is not uncommon to have missing information on the cause of death. Andersen and Ryan (1998) discuss a study on colon cancer in which the cause of death was masked for 25% of the deaths. Lapidus et al. (1994) in a study of motorcycle fatalities found that 40% of death certificates are missing information. This type of problem can also occur in animal bioanalysis (Kodell and Chen, 1987).

Frequently, engineering and medical considerations require that some subset of masked failures be subject to a second stage of study to arrive at a definitive resolution of the failure cause. This occurs in failure analysis laboratories or on autopsy tables. This stage-2 data, appropriately collected, can be used to improve survival estimation. This article focuses on parametric survival analysis with independent competing risks and masked failures. We show how stage-1 and stage-2 information can be combined to produce statistical inference on

- survival (reliability) functions of the individual risks (components)
- the proportion of failures associated with individual components, and
- probability, for a specified masked case, that each of the masked competing risks is responsible for the failure.

There is a series of papers dealing with estimation for masked data under the simplifying assumption that components have exponentially distributed lifetimes. Miyakawa (1984) provides closed form expressions for the maximum likelihood estimators (MLE's) of the failure rates for a two-component series system. Usher and Hodgson (1988) and Guess et al. (1991) extend these results to larger systems. Doganaksoy (1991) discusses how confidence intervals can be obtained while Reiser et al. (1995) present a Bayesian approach. When dealing with more than two components and allowing for the possibility that any subset of components may be masked, the problem becomes very complex and closed form expressions for the MLE's are not in general available. Aboul-Seoud and Usher (1996) show that for modular systems, where no masking across modules is possible, closed form solutions can be obtained. Usher and Guess (1989) and Usher (1996) discuss numerical procedures assuming Weibull distributions while Albert and Baxter (1995) emphasize the usefulness of the EM algorithm in the analysis of parametric masked models. Bayesian analysis under the Weibull assumption is described by Mukhopadhyay and Basu (1997) and Basu et al. (1999).

All of the above papers make the strong symmetry assumption that the probability of masking does not depend on the true failure cause. The nonparametric approaches which attempt to generalize Kaplan-Meier estimates generally make this symmetry assumption (Andersen and Ryan, 1998; van der Laan and McKeague, 1998; Goetghebeur and Ryan, 1990, 1995; Miyakawa, 1984; Lo, 1991; Kodell and Chen, 1987), sometimes only implicitly. Dinse (1986) discusses masking both with and without this symmetry assumption. He develops a likelihood based procedure based on grouping data which leads to non-identifiability and breaks down if the hazard functions of the competing risks are proportional. Lin and Guess (1994) and Guttman et al. (1995) show how one must beware of the symmetry assumption. Flehinger et al. (1996) and Reiser et al. (1996) discuss masking in the go/no-go situation where failure times are assumed to be irrelevant. They recommend a 2-stage experimental procedure in which a sample of the masked cases obtained in stage-1 are subject to sufficient failure analysis to resolve the causes of each of the sampled masked cases. Flehinger et al. (1998) extend these results to lifetime data assuming that the hazard functions of the competing risks are proportional. Their methodology shows how to combine information from standard testing with results of

expensive stage-2 resolution procedures. This approach eliminates the need for arbitrary assumptions about symmetry or other structural properties. For go / no-go data (Reiser et al., 1997) and for lifetime data under the proportional hazards assumption (Flehinger et al., 1998) stage-1 data alone results in non-identifiability; hence, maximum likelihood estimation cannot be carried out. Typically the symmetry assumption is used to overcome the problem. Alternatively a Bayesian approach (Kuo and Yang, 2000) has been suggested.

In many practical situations the assumption of proportional hazards cannot be justified. One possible alternative is to assume a fully parametric model such as gamma, lognormal or Weibull for the lifetime distributions. In Section 2 we describe the likelihood function for 2-stage experimentation in a general parametric setting. In Section 3 we provide details for the Weibull case and describe statistical inference procedures. The Weibull is a useful choice due to its flexibility, and it is often recommended for lifetime data (e.g., see Tobias and Trinidade, 1986).

We illustrate the utility of this approach with several examples presented in Section 4 and compare Weibull modelling for masked data with Flehinger et al.'s (1998) proportional hazards approach. Section 5 concludes with a discussion.

2. Parametric Modelling

Consider a situation in which N systems individually are subject to failure due to independent competing causes. For each system a failure or censoring time is available. Each system failure is due to only one of the competing causes; however, our stage 1 test procedures may result in masking. A sample of cases corresponding to each masked subset is taken to further stage 2 diagnostic tests which provide a definite identification of the cause responsible for the failure. We assume that the probability of choosing the cases taken to stage 2 testing does not depend on the model parameters but only on the observed stage 1 data.

2.1. Notation and Assumptions

We require the following notation related to the observed data:

g indexes each masked group

$i \subset g$ means that cause i is contained in masked group g

n_i times to failure identified as being caused by cause i have lifetimes $t_j^{(i)}, j = 1, \ldots n_i$

n_g times to failure restricted to masking group g have lifetimes $t_j^{(g)}, j = 1, \ldots n_g$

n_c is the number of censored systems. The corresponding censoring (survival) times are lifetimes $t_j^{(c)}, j = 1, \ldots n_c$

$N - n_c = \Sigma_{i=1}^{k} n_i + \Sigma_g n_g$ times to failure have lifetimes $t_j, j = 1, \ldots, N - n_c$; these include all failures, identified or masked.

$n_{g,i}$ is the number of system failures restricted to masking group g in stage 1 and identified with cause i in stage 2.

$n_g^+ = \Sigma_{i=1}^{k} n_{g,i}$ is the total number of system failures restricted to masking group g in stage 1 that are identified in stage 2.

$\tilde{n}_g$ is the number of system failures restricted to masking group g in stage 1 that are not taken to stage 2 for further identification. Clearly, $n_g^+ + \tilde{n}_g = n_g$.

$n_i^* = n_i + \Sigma_{g \supset i} n_{g,i}$ is the total number of system failures identified with cause i in either stage 1 or 2.

$\bar{t}_j$ is the j-th event time that corresponds either to a failure or a censoring time; $j = 1, \ldots N$.

$t_j^{(i^*)}$ is the j-th failure time among all failure times that were identified with cause i in either stage 1 or 2.

$\tilde{t}_j^{(g)}$ is the j-th failure time among all unresolved failure times in masking group g.

The notation below relates to the model and its parameters:

β_i is a vector of unknown parameters associated with the lifetime distribution of cause i.

$f_i(t; \beta_i)$, $S_i(t; \beta_i)$ are the probability density functions and survival functions associated with the cause i.

$f(t; \beta_1, \ldots, \beta_k)$, $S(t; \beta_1, \ldots, \beta_k)$ are the probability density function and survival function associated with the system.

P_i, the ith identification probability, is the probability that a system failure induced by cause i is correctly identified in stage 1.

$P_{g|i}$, the masking probability, is the probability that a system failure induced by cause i is restricted to masking group g in stage 1.

We assume that neither P_i nor $P_{g|i}$ depends on time. Furthermore, under the assumption of independent causes we have that

$$S(t; \beta_1, \ldots, \beta_k) = \prod_{i=1}^{k} S_i(t; \beta_i) \tag{2.1}$$

and

$$f(t; \beta_1, \ldots, \beta_k) = -\frac{\partial}{\partial t} S(t; \beta_1, \ldots, \beta_k) = S(t; \beta_1, \ldots, \beta_k) \times \sum_{i=1}^{k} h_i(t; \beta_i) \tag{2.2}$$

where

$$h_i(t; \beta_i) = f_i(t; \beta_i)/S_i(t; \beta_i) \tag{2.3}$$

is the hazard function for cause i. We denote the cumulative hazard for cause i by

$$H_i(t; \beta_i) = \int_0^t h_i(u; \beta_i) du \tag{2.4}$$

We further assume that the probability density functions satisfy standard regularity conditions and are known up to the unknown parameters β_i.

Finally, denote by $\pi_{i|g}(t)$ the probability that a failure at time t that is restricted to a masking group g is due to component i. One can see from Bayes theorem that

$$\pi_{i|g}(t) = \frac{P_{g|i} f_i(t;\beta_i)}{\Sigma_{r\subset g} P_{g|r} f_r(t;\beta_r)} \tag{2.5}$$

In what follows we refer to $\pi_{i|g}(t)$, $i = 1, 2, \ldots, k$ as the vector of *diagnostic probabilities*.

For masked data in the parametric setting we are interested in estimating not only the β_i (in order to obtain estimates of the lifetime distributions) but also the diagnostic probabilities which can be used as aids to decision making when masking occurs. A more detailed discussion of diagnostic probabilities can be found in Flehinger et al. (1996).

2.2. The Likelihood Function

Based on the above notation and assumptions the likelihood function using both stage-1 and stage-2 data can be seen to be

$$\begin{aligned} L = {} & \Pi_{j=1}^{n_c} S\left(t_j^{(c)};\beta_1,\ldots,\beta_k\right) \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \left\{ P_i f_i(t_j^{(i)};\beta_i) \Pi_{l\neq i} S_l\left(t_j^{(i)};\beta_l\right) \right\} \\ & \times \Pi_g \Pi_{j=1}^{\tilde{n}_g} \left[\Sigma_{r\subset g} P_{g|r} f_r\left(\tilde{t}_j^{(g)};\beta_r\right) \Pi_{l\neq r} S_l\left(\tilde{t}_j^{(g)};\beta_l\right) \right] \\ & \times \Pi_{i=1}^{k} \Pi_{g\supset i} \Pi_{j=1}^{n_{g,i}} \left[P_{g|i} f_i\left(t_j^{(g,i)};\beta_i\right) \Pi_{l\neq i} S_l\left(t_j^{(g,i)};\beta_l\right) \right]. \end{aligned} \tag{2.6}$$

To provide motivation for the likelihood given above note that our N systems can be divided into 4 types: (i) censored, (ii) failure cause identified in stage 1, (iii) failure cause masked and never resolved, (iv) failure cause resolved in stage 2. The likelihood (2.6) is constructed by the multiplication of the contributions of each of these types. The first 2 lines in (2.6) express the contribution of types (i), (ii) and (iii). This is similar to the likelihood given for stage 1 data in Flehinger et al. (1996), differing in that only masked systems not taken to stage 2 are considered. The 3rd line in (2.6) gives the contribution of those systems resolved in stage 2.

After some algebra, the likelihood function can be written in the form:

$$\begin{aligned} L = {} & \Pi_{j=1}^{N} S(\bar{t}_j;\beta_1,\ldots,\beta_k) \Pi_{i=1}^{k} \left(P_i^{n_i} \Pi_{g\supset i} P_{g|i}^{n_{g,i}} \right) \\ & \times \Pi_{i=1}^{k} \Pi_{j=1}^{n_i^*} h_i(t_j^{(i^*)};\beta_i) \Pi_g \Pi_{j=1}^{\tilde{n}_g} \Sigma_{r\subset g} \left[P_{g|r} h_r(\tilde{t}_j^{(g)};\beta_r) \right] \end{aligned} \tag{2.7}$$

Through use of (2.1) and the relation

$$\log S_i(t;\beta_i) = -H_i(t;\beta_i) \tag{2.8}$$

the log-likelihood function can be written as

$$\begin{aligned} logL = &-\Sigma_{j=1}^{N}\Sigma_{i=1}^{k} H_i(\bar{t}_j;\beta_i) + \Sigma_{i=1}^{k}\left(n_i logP_i + \Sigma_{g\supset i} n_{g,i} logP_{g|i}\right) \\ &+ \Sigma_{i=1}^{k}\Sigma_{j=1}^{n_i^*} logh_i(t_j^{(i^*)};\beta_i) + \Sigma_g \Sigma_{j=1}^{\tilde{n}_g} log\left[\Sigma_{r\subset g} P_{g|r} h_r(\tilde{t}_j^{(g)};\beta_r)\right]. \end{aligned} \tag{2.9}$$

In the most general case where none of the parameters can be treated as known quantities, we are interested in maximizing (2.9) over the distributional parameters (β_1, β_2, ..., β_k), the identification probabilities P_1, P_2, ..., P_k, and the masking probabilities $P_{g|i}$, for all possible i and g, in order to obtain the desired MLE's. An iterative approach to this problem will be taken. First we present an algorithm for estimating the identification and masking probabilities for given values of the β_i. Then we discuss estimation of β_i given the identification and masking probabilities. To implement an iterative procedure for the complete set of parameters we cycle between these two procedures. Specifically, we start with some initial estimate of the identification and masking probabilities - in our examples we use the initial values corresponding to the non-parametric proportional hazards model, see Flehinger et al. (1998). Then we find the MLE's of $\{\beta_i\}$ under the assumption that the identification and masking probabilities are known and equal to the initial values. Assuming that the values of $\{\beta_i\}$ are known and equal to these estimates, we re-estimate the identification and masking probabilities, and so forth. The process stops when the results of the new iteration are sufficiently close to those of the previous one.

If only stage-1 data is available, setting all the $n_{g,i} = 0$ gives the appropriate log-likelihood function. The iterative estimation algorithm outlined in this paper can still be used in this situation. Previous approaches discussed in the literature (see introduction) for obtaining MLE's for stage-1 data for parametric models have made the symmetry assumption, which is not used under our approach. One has to take into account, however, that the model based on stage-1 data only is not always identifiable - for example, basic parameters cannot be estimated in the case of two causes of failure with proportional hazards (e.g., see Flehinger et al., 1998).

It is also worth mentioning that in parametric models one will generally be able to obtain parameter estimates related to cause i even if no failures attributable to this cause were observed, either directly or as a result of stage-2 resolution. For example, consider the case of two causes of failure with a single masked group (1, 2). Suppose that a cluster of failures of type 1 was observed at one point in time and a cluster of masked failures was observed at a much later time. Under the assumption that lifetimes for both causes have a Weibull distribution, the likelihood estimation process will still produce estimates for masking probabilities and cause 2 lifetime parameters; the corresponding diagnostic probabilities will tend to attribute the masked failures to cause 2.

2.3. Estimation of the Identification and Masking Probabilities

In order to maximize *logL* given by (2.9) for the identification and masking probabilities (given the distributional parameters (β_1, β_2, ..., β_k), we introduce the Lagrangian

$$G = \log L + \sum_{i=1}^{k} \lambda_i \times \left(1 - P_i - \sum_{g \supset i} P_{g|i}\right). \tag{2.10}$$

Conditions $0 \leq P_i \leq 1$ and $0 \leq P_{g|i} \leq 1$ are not explicitly represented in the Lagrangian since our solution satisfies them automatically, as can be seen from the iterative scheme presented below.

By setting the derivatives of G by P_i and $P_{g|i}$ to zero for every i we obtain the equations

$$\begin{cases} \frac{\partial G}{\partial P_i} = \frac{n_i}{P_i} - \lambda_i = 0 \\ \frac{\partial G}{\partial P_{g|i}} = \frac{n_{g,i}}{P_{g|i}} + \frac{1}{P_{g|i}} \Sigma_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)}) - \lambda_i = 0, \end{cases} \tag{2.11}$$

for every set g that contains i. It is not difficult to see that λ_i estimates the total number of failures that are due to cause i.

Under the assumption that $(\beta_1, \beta_2, \ldots, \beta_k)$ are known, the equations (2.11) can be solved by using the following iterative process.

Procedure. Start with an arbitrary set of initial identification and masking probabilities (for example, for every i and $g \supset i$ one can set $P_i = P_{g|i} = 1/1 + \Sigma_{g \supset i} 1$).

Step1. For every $i = 1, 2, \ldots, k$ compute the estimated values of (λ_i) corresponding to the current values of P_i and $P_{g|i}$ as follows:

$$\lambda_i = n_i^* + \sum_{g \supset i} \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)}). \tag{2.12}$$

Step2. Update the identification and masking probabilities,

$$\begin{cases} P_i^{(new)} = n_i / \lambda_i \\ P_{g|i}^{(new)} = (1/\lambda_i) \left[n_{g,i} + \Sigma_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)}) \right], \end{cases} \tag{2.13}$$

and return to Step 1. Declare a solution when the difference between the new values of the identification and masking probabilities, and the previous ones becomes sufficiently small.

Note that at the end of the iteration process the values $\pi_{i|g}(\tilde{t}_j^{(g)})$ provide the estimates of diagnostic probabilities for every unresolved masked case.

It is easy to show that the log-likelihood is a convex function, and, since the constraints are all linear, the maximum resulting from the above algorithm is unique. The convergence of the above procedure follows easily from the fact that it can be represented as a version of an EM algorithm (McLachlan and Krishnan, 1997).

2.4. Estimation of the Distributional Parameters

The equations for the β_i, obtained by setting the derivatives of (2.9) to zero are

$$\sum_{j=1}^{n_i^*} \frac{\partial}{\partial \beta_i} \log h_i(t_j^{(i^*)}; \beta_i) + \sum_{g \supset i} \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)}) \frac{\partial}{\partial \beta_i} \log h_i(\tilde{t}_j^{(g)}; \beta_i)$$
$$= \sum_{j=1}^{N} \frac{\partial}{\partial \beta_i} \log H_i(\bar{t}_j; \beta_i), \quad i = 1, 2, \ldots, k. \tag{2.14}$$

These equations can be solved by an iterative process, the details of which will depend on the specific parametric form of the hazard functions.

A likelihood ratio test based on L can be obtained for any particular parameter of interest. This test can be inverted (Meeker and Escobar, 1998) to provide approximate confidence bands for the parameter. This procedure involves constraining the parameter of interest to take a specified value and then maximizing L under the constraint. By varying the specified value a profile likelihood for the parameter is obtained, which serves as a basis for computing an approximate confidence interval.

2.5. Estimation of the Diagnostic Probabilities

The diagnostic probabilities $\pi_{i|g}(t)$, $i = 1, 2, \ldots, k$ corresponding to an observed masking group g are computed by substitution of the identification, masking and distributional parameter estimates, obtained as described above, into (2.5). Note that in the case of proportional hazards the diagnostic probabilities do not depend on time t at which the masked failure is observed.

3. The Weibull Case

In this case the survival function for the cause i is

$$S_i(t) = exp[-(t/\theta_i)^{\delta_i}], \quad i = 1, \ldots, k, \tag{3.1}$$

and the 2-stage log-likelihood function can be represented in the form

$$logL = -\Sigma_{i=1}^{k} \Sigma_{j=1}^{N} (\bar{t}_j/\theta_i)^{\delta_i} + \Sigma_{i=1}^{k} \left(n_i log P_i + \Sigma_{g \supset i} n_{g,i} log P_{g|i} \right)$$
$$+ \Sigma_{i=1}^{k} \left[n_i^* log(\delta_i/\theta_i) + (\delta_i - 1) \Sigma_{j=1}^{n_i^*} log(t_j^{(i^*)}/\theta_i) \right]$$
$$+ \Sigma_g \Sigma_{j=1}^{\tilde{n}_g} log \left[\Sigma_{r \subset g} P_{g|r} (\delta_r/\theta_r) (\tilde{t}_j^{(g)}/\theta_r)^{\delta_r - 1} \right] \tag{3.2}$$

3.1. Estimation and Inference

Given the (θ_i, δ_i), $i = 1, 2, \ldots, k$ the P_i and $P_{g|i}$ can be estimated as described in Section 2.3. Note that in this case

$$\pi_{i|g}(t) = \frac{P_{g|i}(\delta_i/\theta_i)\,(t/\theta_i)^{\delta_i - 1}}{\Sigma_{r \subset g} P_{g|r}(\delta_r/\theta_r)\,(t/\theta_r)^{\delta_r - 1}}. \tag{3.3}$$

Applying (2.14) just to the Weibull scale parameters results in the system of equations

$$\sum_{j=1}^{N} (\bar{t}_j/\theta_i)^{\delta_i} = \lambda_i, \quad i = 1, 2, \ldots, k, \tag{3.4}$$

where λ_i is defined by (2.12). Under the assumption that all the other parameters are known, these equations can be solved by using a simple and efficient iterative process,

$$\theta_i^{(new)} = \left[\frac{\sum_{j=1}^{N} \bar{t}_j^{\delta_i}}{\lambda_i} \right]^{1/\delta_i}. \tag{3.5}$$

The equations for $(\delta_1, \delta_2, \ldots, \delta_k)$ are somewhat more complex,

$$\delta_i \sum_{j=1}^{N} (\bar{t}_j/\theta_i)^{\delta_i} log(\bar{t}_j/\theta_i) = \lambda_i + \delta_i \tilde{\lambda}_i, \quad i = 1, 2, \ldots, k, \tag{3.6}$$

where $\tilde{\lambda}_i$ is defined by

$$\tilde{\lambda}_i = \sum_{j=1}^{n_i^*} log(t_j^{(i^*)}/\theta_i) + \sum_{g \supset i} \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)}) log(t_j^{(i^*)}/\theta_i). \tag{3.7}$$

These equations also can be solved by using an iterative process similar to that for the scale parameters; the details will be omitted.

Note that by using the processes described above one can obtain, in a relatively straightforward way, solutions for the equations to maximize the likelihood. Substitution of the appropriate MLE's into (3.1) results in $\hat{S}_i(t)$, the estimate of the survival function for failure cause i. We denote by $\hat{L}$ the maximal value of L. Furthermore, one can run the same iterative process when a selected subgroup of parameters assumes fixed values, leading to profile likelihood based inference for this subgroup.

In particular, for any specified model parameter of interest (i.e., any of the P_i, $P_{g|i}$, θ_i or δ_i) say γ, the above set of the equations are iteratively solved as described above for a fixed value of γ and ignoring the equation which involves the derivative with respect to γ (this equation can be used to obtain the Lagrange multiplier associated with γ). Let $L^*(\gamma)$ denote the maximal value of L obtained under the constraint that the value of γ is fixed. The log-likelihood ratio associated with this value is

$$\Psi(\gamma) = -2\{\log L^*(\gamma) - \log \hat{L}\}; \tag{3.8}$$

asymptotically, at the hypothesized value of γ, this value has a chi-squared distribution with one degree of freedom. Approximate confidence limits for γ can be obtained by finding the values of γ such that $\Psi(\gamma)$ is equal to the required percentile point of this distribution.

The profile likelihood analysis of the individual survival curves $S_i(t_0)$ or the system survival curve at a specified time t_0 is more complicated and is given in Appendix A.

3.2. Testing the Equality of Weibull Shape Parameters

Under the parametric Weibull assumption, testing the equality of the shape parameters of the lifetime distributions of the k causes is equivalent to testing the proportionality of the respective hazard functions.

In order to derive a likelihood ratio test of the hypothesis of equal shape parameters, i.e.,

$$H : \delta_1 = \delta_2 = \ldots = \delta_k, \tag{3.9}$$

it is necessary to obtain the maximal value of the likelihood function under this hypothesis; we denote the common (and unknown) value of the shape parameter by δ.

Under H, the survival function of cause i is $S_i(t) = exp\,[-(t\,/\,\theta_i)^{\delta}\,]$ and we can write the system survival function as

$$S(t) = exp[-(t/\theta)^{\delta}], \tag{3.10}$$

where the system scale parameter θ is defined by

$$\theta = \left(\sum_{i=1}^{k} \frac{1}{\theta_i^{\delta}}\right)^{-1/\delta} \tag{3.11}$$

and, correspondingly,

$$f(t) = \delta t^{\delta-1}\theta^{-\delta} exp[-(t/\theta)^{\delta}]. \tag{3.12}$$

The probability ϕ_i that a failure is due to cause i can in this situation readily be seen to be

$$\phi_i = (\theta/\theta_i)^{\delta}, \tag{3.13}$$

independently of the time at which the failure occurs.

After some algebra L_H, the likelihood under H can be written as

$$L_H = L_1 \times L_2 \tag{3.14}$$

where

$$\begin{aligned} L_1 &= \prod_{i=1}^{k} (P_i\phi_i)^{n_i} \prod_{g} \left(\sum_{i \subset g} P_{g|i}\phi_i \right)^{\tilde{n}_g} \prod_{i=1}^{k} \prod_{g \supset i} (P_{g|i}\phi_i)^{n_{g,i}} \\ L_2 &= \left\{ \prod_{j=1}^{n_c} S\left(t_j^{(c)}\right) \right\} \prod_{j=1}^{N-n_c} f(t_j). \end{aligned} \tag{3.15}$$

L_1 has been examined in Flehinger et al. (1998) where analytic formulae for its optimization are presented. L_2 is just the usual likelihood for censored Weibull data whose optimization is well known (e.g., see Meeker and Escobar, 1998).

Consequently, an asymptotic likelihood ratio test for H is obtained by treating $-2\log(\max L_H/\hat{L})$ as a χ^2_{k-1} variate.

4. Examples

In this section we consider two applications of the proposed methodology. The first application is in the field of reliability; it discusses a situation in which the Weibull model is likely to be much more useful than a nonparametric proportional hazards model. The second application is in the field of cancer research. We re-examine a study previously published and show that the Weibull model results in similar conclusions. A special purpose program was written by the third author in APL2 in order to carry out the computations described above for the following examples.

Example 1. Consider a scenario in which a company manufacturing hard drives for computers is trying to analyze causes of failures of a certain sub-assembly. Some of these causes, such as "defective head", are related to components, but others (e.g., "particle contamination") are not; in this example, we will not treat these separately and will refer to them simply as causes of failure. In the situation we consider there are three major causes of failures which, without going into details, we denote as causes 1, 2 and 3. We assume that these causes act independently and in series. Four years ago 10,000 drives were

Table 1. Estimates of the Weibull parameters (δ_i, θ_i) for failures due to Cause i.

Parameter	MLE	95%	Bounds
δ_1	.691	.511	.908
δ_2	1.006	.632	1.516
δ_3	2.151	1.695	2.686
θ_1	7568	1170	124000
θ_2	1566	202	59600
θ_3	34.9	22.6	62.9

Table 2. Estimates of masking probabilities (Weibull model).

Masking Group	Masking Probs		
	1	2	3
1,3	.412	–	.446
1,2,3	.310	.469	.436

manufactured and since then information about failures was collected in a database. The number of failures observed in this period was 172. Some of the failures were masked and a selected number of those were analyzed to complete resolution in the defect isolation laboratory. The only observed masked groups were (1, 2, 3) and (1, 3).

The resulting data is given in the first four columns of the table in Appendix B. The first two columns give the sequential number and failure time (in years), respectively. The third column (outcome) gives the cause of failure if it was either identified immediately or resolved in stage-2; the value -1 correspond to unresolved failures. Finally, the fourth column gives the masking information. Consider, for example, failure #6 for which the entry in the third column is 1 and the entries in the 4-th column are 2, 3. This failure originally corresponded to the masked group (1, 2, 3), but then was resolved as a failure due to cause 1. Failure #20 was immediately identified as one related to cause 2. In cases of unresolved failures column 4 contains the complete masking group; for example, it is still unknown whether failure #3 is due to cause 1 or 3.

Assuming that the failure causes have associated Weibull lifetimes we apply the methodology described in Section 3. The resulting MLE estimates of the shape parameters (dimensionless) and scale parameters (measured in years) are $(\delta_1, \delta_2, \delta_3) = (0.691, 1.006, 2.151)$ and $(\theta_1, \theta_2, \theta_3) = (7568, 1566, 34.9)$; the 95% confidence bounds for these parameters are given in Table 1. One can see that the upper confidence bounds for θ_1 and θ_2 are quite high, primarily because the data is censored at four years; however, as we will see below, the bounds for survival probabilities are still practically meaningful.

The point estimates of masking probabilities are given in Table 2. For each observed masked group the corresponding $P_{g|i}$ are presented in the rows of the Table. Since $P_i + \Sigma_{g \supset i} P_{g|i} = 1$, the P_i can be obtained by subtracting the sum of the columns in Table 2 from 1. For example, based on Table 2 we estimate that if a failure is due to cause 1 then

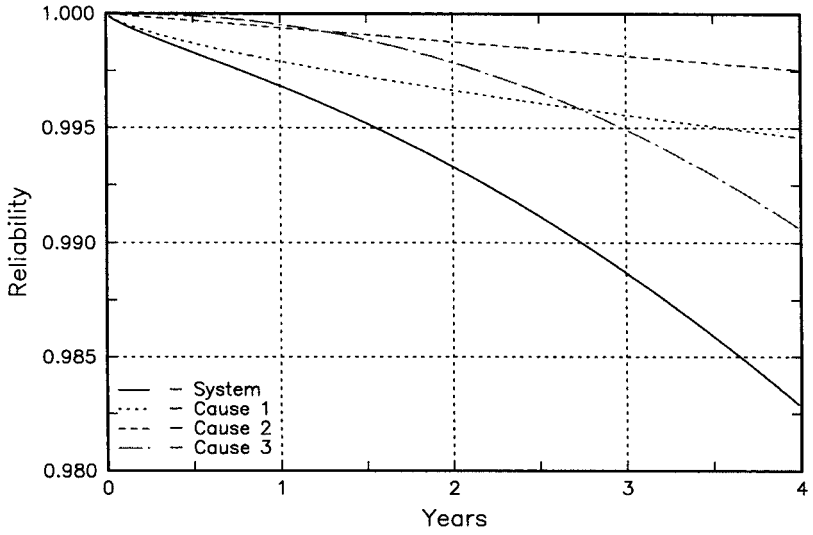

Figure 1. Estimates of reliability (Weibull model) for Example 1 data.

with probability 0.412 we will observe a group (1, 3), with probability .310 we will observe (1, 2, 3), and with probability 0.278 = 1-0.412-0.310 we will immediately identify cause 1 as culprit. The estimated diagnostic probabilities $\pi_{i|g}(\tilde{t}_j^{(g)})$ are shown in the last 3 columns of the data table in Appendix B. For example, in the unresolved (1 3) masked case corresponding to failure #3 the estimated probability that the failure is due to cause 1 is 0.992. In a similar masked case #166, however, this probability is only 0.149.

Figure 1 gives the estimated survival functions $\hat{S}_i(t)$ under the Weibull model, while Figure 2 provides, for comparison, estimates obtained under a nonparametric proportional hazards assumption (see Flehinger et al., 1998). Such estimates can be obtained for up to four years. The Weibull model enables one to get estimates for longer time periods. For example, survival function estimates for up to 5 years, along with lower 95% confidence bounds, are given in Table 3.

Finally, Figure 3 gives the estimated (Weibull) hazards for the system and the individual causes of failure.

The log-likelihood ratio statistic for the hypothesis H: $\delta_1 = \delta_2 = \delta_3$ is $-2\log(\max L_H/\max L) = 32.20$ (the estimated common value of the shape parameter is $\delta = 1.18$). This value is clearly incompatible with the χ_2^2 null distribution, indicating the unsuitability of the proportional hazards assumption. One could have expected this conclusion from the confidence bounds for the shape parameter given in Table 1, or on purely graphical grounds, from Figures 1 and 3.

For comparison, we performed an analysis of the case where none of the masked failures were resolved (i.e., analysis based on stage-1 data only). The likelihood analysis results in estimates for survival probabilities that are quite close to those obtained by bringing in the 2-nd stage - but the estimates of masking and diagnostic probabilities are different.

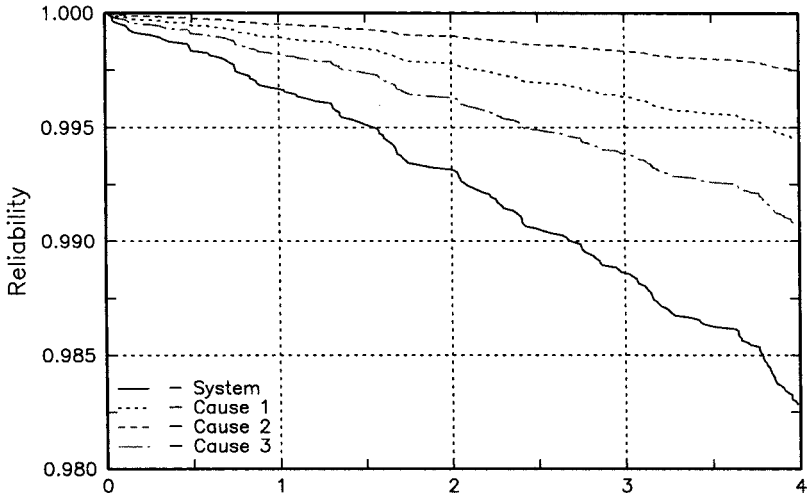

Figure 2. Estimates of reliability (proportional hazards model) for Example 1.

Specifically, the stage-1-only estimates are $(\delta_1, \delta_2, \delta_3) = (0.63, 0.84, 1.94)$ and $(\theta_1, \theta_2, \theta_3) = (34000, 3500, 41)$; The estimated identification probabilities are $(p_1, p_2, p_3) = (0.47, 0.38, 0.10)$ and the estimated masking probabilities are $p_{(123)|1,2,3} = (0\ 0.62\ 0.45)$ and $p_{(13)|1,3} = (0.53\ 0.45)$. As one can see based on the group (1 2 3), the maximum of the likelihood in stage-1-only case can be attained on the boundary of the feasible region. This does not prevent one from executing likelihood based inference on $p_{(123)|1}$ - however, we cannot rely on the standard asymptotic theory to do the job.

Example 2. We reconsider the cancer study examined by Flehinger et al. (1998). The data consists of survival times from cancer detection for 838 men with confirmed diagnoses of non−small cell lung cancer. The data are censored at six years and during this time 543 deaths were observed with 295 lifetimes censored. After a great deal of effort it was determined that 476 deaths were directly attributable to lung cancer (cause 1) and 67 to other causes (cause 2). Figure 4 presents both Kaplan-Meier and Weibull estimates of the

Table 3. Values of $\hat{S}_i(t)$ and corresponding 95% lower confidence bounds.

	Cause 1		Cause 2		Cause 3		System	
Years	Prob	Bound	Prob	Bound	Prob	Bound	Prob	Bound
1	.9979	.9970	.9994	.9988	.9996	.9991	.9968	.9958
2	.9966	.9954	.9988	.9980	.9979	.9968	.9933	.9918
3	.9956	.9940	.9882	.9971	.9949	.9934	.9987	.9967
4	.9946	.9927	.9975	.9962	.9906	.9883	.9828	.9801
5	.9937	.9915	.9969	.9952	.9848	.9808	.9756	.9713

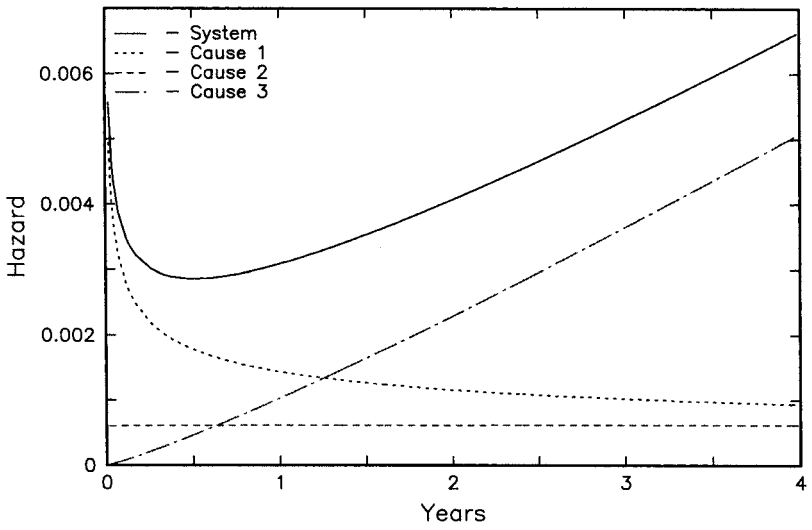

Figure 3. Estimates of the hazards (Weibull model) for Example 1.

survival probabilities for both causes based on this data set. Both the parametric (based on the Weibull model) and non-parametric approaches give essentially the same survival curves. The estimated scale and shape parameters for the Weibull model are $(\delta_1, \delta_2) = (0.859, 1.21)$ and $(\theta_1, \theta_2) = (5.52, 25.3)$ for causes 1 and 2, respectively. For this data both the parametric and nonparametric fits are quite similar.

We examine the properties of our procedures by introducing artificial masking for a fraction of the patient population, along with the correct assessment of the cause of death for a random fraction of the masked group. Following Flehinger et al. (1998), masking was introduced with probability 0.1 for cancer–caused deaths and with probability 0.5 for deaths from other causes. Twenty per cent of the masked cases were taken to stage 2 for complete diagnosis. After masking was introduced, stage 1 data showed 543 deaths of which 425 were due to cancer, 33 to other causes, and 85 masked. Of the masked cases, 17 were resolved, with 10 diagnosed as cancer deaths and 7 as due to other causes.

Assuming Weibull lifetimes the resulting MLE's of the shape and scale parameters are $(\delta_1, \delta_2) = (0.863, 1.16)$ and $(\theta_1, \theta_2) = (5.53\ 27.2)$. The estimated identification probabilities are $(p_1, p_2) = (0.89, 0.48)$. In spite of the masking, these estimates are close to those produced for the complete data set. Figure 5 shows the estimated survival curves based on (i) the Weibull model and (ii) the proportional hazards model of Flehinger et al. (1998). These can be compared with the unmasked data results in Figure 4. Both (i) and (ii) based on a 20% resolution of masked data well reproduced the original Kaplan-Meier estimates. Table 4 presents 95% confidence intervals of survival probabilities based on profile likelihoods for both approaches.

One can see that the results are quite similar; however, the test of the Section 3.2 does not support the hypothesis of equal shape parameters. The best Weibull proportional

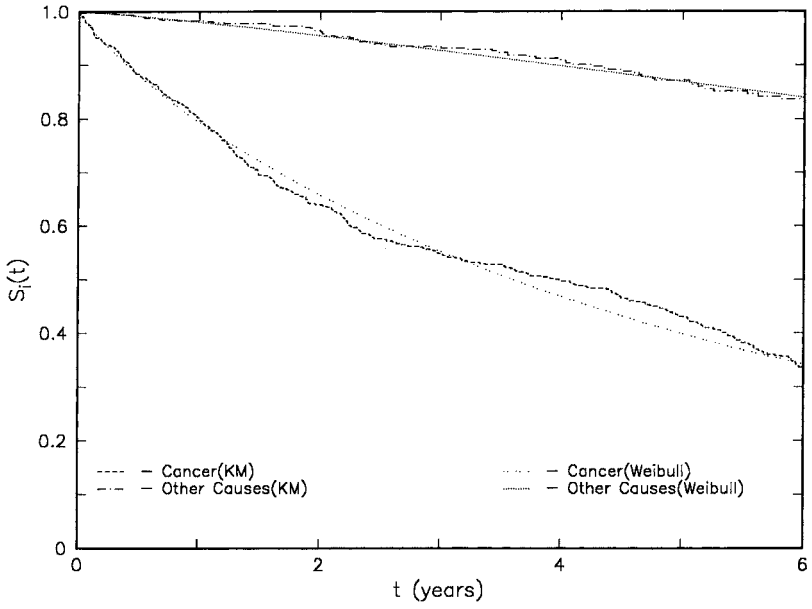

Figure 4. Kaplan-Meier (KM) and Weibull based estimates of survival functions for cancer and other causes, for the complete data set in Example 2.

hazards model corresponds to the parameters $(\delta_1, \delta_2) = (0.892, 0.892)$ and $(\theta_1, \theta_2) = (5.46\ 48.24)$. These parameters (together with the corresponding estimates of P_i and $P_{g|i}$) yield $-2\log(\max L_H/\hat{L}) = 5.18$. The p-value of the test for equality of shape parameters, computed on the basis of the χ_1^2 null distribution, is 0.023. This, examined together with Figure 5 and the shape parameter estimates given above indicate that the deviation from hazard proportionality, although statistically significant, is too small to be of practical importance; this lack of proportionality is far less prominent than that observed in Example 1.

For comparison, we also consider the case where none of the 85 failures are resolved in the second stage. Under Weibull assumption, the estimated shape and scale parameters are $(\delta_1, \delta_2) = (0.861, 1.45)$ and $(\theta_1, \theta_2) = (5.22\ 25.6)$. The estimated identification probabilities are $(p_1, p_2) = (0.85, 0.76)$. The estimated survival probabilities for 1, 2, 3 and 5 years are (0.79, 0.65, 0.54, 0.38) for cancer and (0.99, 0.98, 0.96, 0.91) for deaths from other causes. Given that the results obtained from Weibull and non-parametric proportional hazards are quite similar and that we know the true values, $(p_1, p_2) = (0.9, 0.5)$, of the identification probabilities, this example suggests that one could get substantially better estimates by resolving a small fraction of masked cases.

Example 3. We consider the same general situation as in Example 2, but in this case all 543 deaths are masked and 20% of the cases (109) are taken to stage-2 resolution; of these

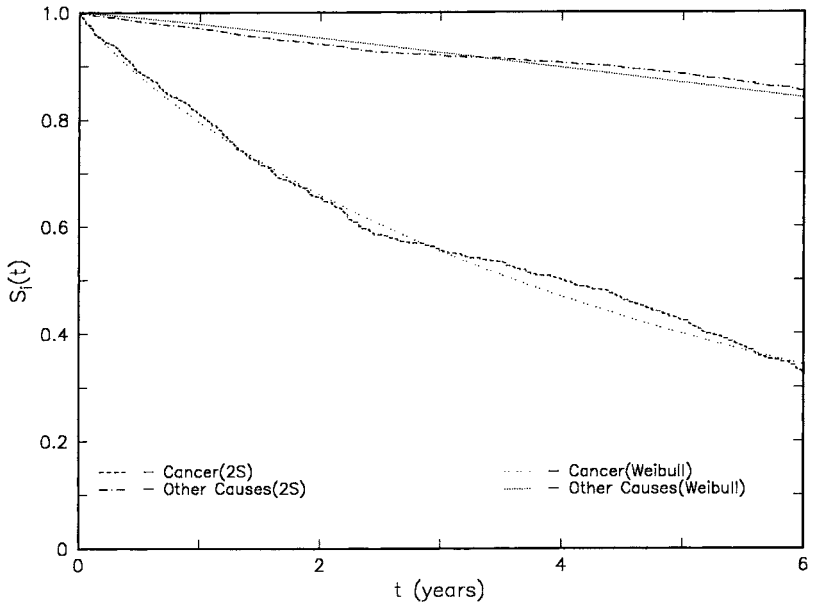

Figure 5. Estimates of survival functions for cancer and other causes, based on Example 2 data with masking, for the estimation model (2S) of Flehinger et al. (1998) and for the Weibull model.

92 are found to be due to cancer and 17 to other causes. The estimated Weibull parameters were $(\delta_1, \delta_2) = (0.847, 1.23)$ and $(\theta_1, \theta_2) = (5.80\ 20.4)$. Figure 6 shows the survival curves for both the Weibull and proportional hazard models. Again both approaches on the

Table 4. Estimated Survival Probabilities in Example 2. The top table corresponds to the nonparametric proportional hazards model, the bottom table—to the Weibull model.

	Cancer Death		Other Causes	
Years From Detection	$\hat{S}_1(t)$	95% Conf. Int.	$\hat{S}_2(t)$	95% Conf. Int.
1	0.81	(0.78, 0.84)	0.97	(0.96, 0.98)
2	0.66	(0.62, 0.69)	0.94	(0.92, 0.96)
3	0.56	(0.52, 0.59)	0.92	(0.89, 0.95)
5	0.43	(0.39, 0.46)	0.89	(0.84, 0.92)

	Cancer Death		Other Causes	
Years From Detection	$\hat{S}_1(t)$	95% Conf. Int.	$\hat{S}_2(t)$	95% Conf. Int.
1	0.80	(0.77, 0.82)	0.98	(0.96, 0.99)
2	0.66	(0.63, 0.69)	0.95	(0.93, 0.97)
3	0.55	(0.52, 0.59)	0.92	(0.89, 0.95)
5	0.40	(0.36, 0.44)	0.87	(0.82, 0.91)

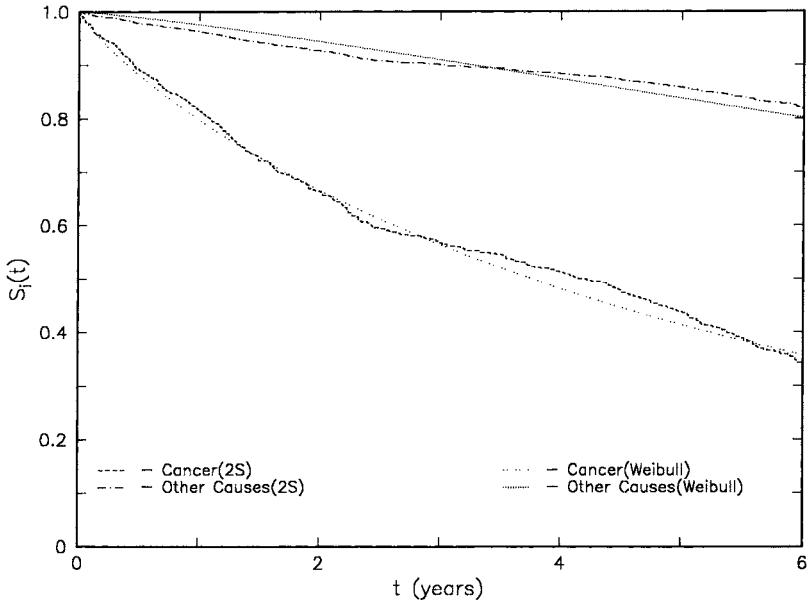

Figure 6. Estimates of survival functions for cancer and other causes for the Example 3 data.

masked data reproduce the Kaplan-Meier estimate on the original data even though only 20% of the deaths have an identified cause.

When the percentage of resolved cases was reduced we found that the Weibull continued to fit the cancer survival function quite well but did not do as well on the other causes. This is a consequence of the high survival rate for other causes resulting in very few observed "other causes" deaths when only a small number of masked cases is resolved. The proportional hazards modelling does better in this situation as it "borrows" observations from the cancer deaths. For example, consider the situation where all 543 deaths are masked and only 5% of the cases (28) are taken to complete resolution; of these 26 are due to cancer and 2 to other causes. In this case the estimated Weibull parameters are $(\delta_1, \delta_2) = (0.840, 9.52)$ and $(\theta_1, \theta_2) = (5.12\ 7.23)$, indicating that no reliable estimates can be expected for death rate from other causes.

5. Discussion

Our approach to lifetime masked data has the advantage of producing an estimable model without requiring strong assumptions about the masking probabilities, such as the symmetry assumption. In this paper we have emphasized the use of parametric modelling for the competing causes. This is more general than the proportional hazards approach of

Flehinger et al. (1998) in the sense that the competing risks are not assumed to have proportional hazards functions, while less general by forcing a particular parametric structure on the data. This restriction is mitigated by the flexibility of many classes of parametric models such as Gamma, Lognormal or Weibull. The examples we have studied indicate that in strongly non-proportional hazard cases parametric models tend to be clearly superior while if lack of proportionality is small, little is lost by making an adequate parametric assumption.

A natural question arises with regards to conditions under which the set of equations (2.14) has a unique solution that corresponds to the maximum of the log-likelihood (2.9). Unfortunately, we do not have such conditions, even in the Weibull case. In practice, we never saw the algorithms presented for this case fail to produce a local maximum (we suspect that this happens because for most data sets compatible with the Weibull model the underlying log-likelihood is pseudo-convex) - however, we do not have a proof that this is also a global maximum for all possible data sets. The situation here is quite similar to that arising in neural network models or image analysis.

Another important issue concerns the decisions related to Stage 2 sampling. In principle, one could run into situations where stage-1 data alone is sufficient to draw acceptably accurate conclusions (in this case, one could abolish stage-2 resolution altogether). The impact of masking fraction on the stage-1 estimates depends on a number of factors, such as number of masking groups, their size and composition, and the underlying life time distributions. For example, if there are only two possible causes of failure (i.e., only one masking group is possible) and the masking probabilities for the two components are fixed - the power of inference will tend to be higher in the case where the causes of failure correspond to opposite Failure Rate families (i.e., one of the causes corresponds to Increasing Failure Rate family and the other - to Decreasing Failure Rate family) than in the case where both causes belong to the same family. As noted earlier, in the extreme case of proportional hazards the model based on Stage-1 data only becomes non-identifiable.

Once it is decided to proceed with stage-2 resolution, one will need to establish the sample size and identify cases that are to be resolved. A number of strategies can be effective, depending on the concrete situation - study of such strategies is an interesting research topic in its own right. For example, in situations where masked cases can be resolved sequentially, one could, in principle, formulate criteria to determine which masked case is to be resolved next - one possibility would be to pick a case that would lead to most sizeable reduction in some measure of a confidence region related to parameters deemed most important at the present point in time. However, in practice these issues are typically handled much less formally - the question of which masked cases are resolved next generally depends on a large number of factors and constraints related to cost, availability of tools, time and personnel, management pressures related to getting information about particular components or development lab pressures. It is important to note that the likelihood based analysis presented here does not depend on the policy of choosing a set of cases for stage-2 analysis.

If the cause of failure is known for all systems, standard graphical methods are available for investigating the suitability of the chosen distributional model. How to carry this out under masking is an open research question which we are currently investigating. We

expect that in many applications the choice of distributional models will be strongly influenced by traditional preferences in the area of interest; for example, one working with reliability engineering community is likely to be asked to emphasize Weibull model even if other models happen to provide a better fit for the data set of interest.

Numerous generalizations of the presented model are possible. For example, estimation similar to that presented in Section 2 can also be carried out when the distributional parameters depend on covariates or when identification or masking probabilities depend on time. One can also generalize the analysis to the case when competing causes are dependent, but this type of analysis is quite complex as there are generally many forms of dependence that result in the same likelihood.

In summary, parametric analysis of masked failures is one of the cardinal reliability issues in today's manufacturing, given the increasing trend in modularity; we also believe that this issue is important in medical applications. We believe that the basic idea of consolidating data observed in the field and in the failure resolution labs into a single model is very promising and will eventually be accepted by a number of industries, in spite of the fact that it requires radical changes in the process of data management.

Appendix A. Inference on the survival functions

Performing profile likelihood based inference on $S_r(t_0)$ requires maximization of the likelihood function for a fixed value, say, $exp(-c)$ of this parameter. This approach has been used for survival functions in a number of nonparametric survival models (e.g., see Flehinger et al., 1998; Thomas and Grunkemeier, 1975). The corresponding Lagrangian can be written as

$$G_r = \log L + \sum_{i=1}^{k} \lambda_i \times \left(1 - P_i - \sum_{g \supset i} P_{g|i}\right) + \Delta_r \left[\log c - \delta_r \log(t_0/\theta_r)\right], \tag{A.1}$$

where Δ_r is the Lagrange multiplier. First, consider the case where the shape parameters are known. Then the equations for the scale parameters are

$$\begin{aligned} \left(\frac{\theta_i}{\delta_i}\right)\frac{\partial G_r}{\partial \theta_i} &= \Sigma_{j=1}^{N} (\bar{t}_j/\theta_i)^{\delta_i} - \lambda_i = 0, \quad i = 1, 2, \ldots, k, i \neq r \\ \left(\frac{\theta_r}{\delta_r}\right)\frac{\partial G_r}{\partial \theta_r} &= \Sigma_{j=1}^{N} (\bar{t}_j/\theta_r)^{\delta_r} - \lambda_r = -\Delta_r. \end{aligned} \tag{A.2}$$

The equations for P_i and $P_{g|i}$ are given by (2.11). For every Δ_r one can find a solution of (A.2), $\{\theta_1, \theta_2, \ldots, \theta_k\}$, by using an iterative method similar to (3.5). The parameter values that maximize L under the above constraints are obtained by iterating systems of equations

(2.11) and (A.2). For a fixed value of Δ_r this solution gives optimal values of parameters (θ_i^*, δ_i^*), $i = 1, 2, \ldots, k$, which correspond to a value of the survival function fixed at

$$S_r(t_0) = exp\Big[- (t_0/\theta_r^*)^{\delta_r^*} \Big], \tag{A.3}$$

Denote the corresponding value of L by L^* (Δ_r); then the problem of constructing an upper $(1 - \alpha) \times 100\%$ confidence bound reduces to finding $\Delta_r > 0$ for which

$$\Psi(\Delta_r) = -2\{\log L^*(\Delta_r) - \log \hat{L}\} = \chi^2_{1-2\alpha}. \tag{A.4}$$

Other types of inference are analogous.

Now consider the case where the shape parameters are unknown. Then one needs to include an additional group of equations in the system, namely,

$$\delta_i \frac{\partial G_r}{\partial \delta_i} = -\delta_i \, \Sigma_{j=1}^N (\bar{t}_j/\theta_i)^{\delta_i} log(\bar{t}_j/\theta_i) + \lambda_i + \delta_i \tilde{\lambda}_i = 0, \quad i = 1, 2, \ldots, k, i \neq r$$

$$\delta_r \frac{\partial G_r}{\partial \delta_r} = -\delta_r \, \Sigma_{j=1}^N (\bar{t}_j/\theta_r)^{\delta_r} log(\bar{t}_j/\theta_r) + \lambda_r + \delta_r \tilde{\tilde{\lambda}}_r = \delta_r \Delta_r \log(t_0/\theta_r). \tag{A.5}$$

Obtaining L^* (Δ_r) now requires cycling through three systems of equations, (2.11), (A.2) and (A.5). The inference is then performed, as in the previous case, on the basis of the equation (A.4).

Finally, to perform inference on the overall system survival function at t_0, i.e., $S(t_0) = \Pi_{i=1}^k S_i(t_0)$, we need to maximize the Lagrangian

$$G = \log L + \sum_{i=1}^k \lambda_i \times \left(1 - P_i - \sum_{g \supset i} P_{g|i} \right) + \Delta \left(\log c - \log \left[\sum_{i=1}^k (t_0/\theta_i)^{\delta_i} \right] \right) \tag{A.6}$$

This leads to the systems of equations

$$\left(\frac{\theta_i}{\delta_i} \right) \frac{\partial G}{\partial \theta_i} = \sum_{j=1}^N (\bar{t}_j/\theta_i)^{\delta_i} - \lambda_i = -\Delta \phi_i(t_0), \quad i = 1, 2, \ldots, k, \tag{A.7}$$

and

$$\begin{aligned} \delta_i \frac{\partial G}{\partial \delta_i} &= -\delta_i \sum_{j=1}^N (\bar{t}_j/\theta_i)^{\delta_i} log(\bar{t}_j/\theta_i) + \lambda_i + \delta_i \tilde{\lambda}_i \\ &= \delta_i \Delta \phi_i(t_0) \log(t_0/\theta_i), \; i = 1, 2, \ldots, k, \end{aligned} \tag{A.8}$$

where

$$\phi_i(t_0) = \frac{(t_0/\theta_i)^{\delta_i}}{\Sigma_{i=1}^{k}(t_0/\theta_i)^{\delta_i}} \tag{A.9}$$

is the probability that a failure at time t_0 is due to cause i. One can now obtain estimates of the parameters and $L^*(\Delta)$ by cycling through the systems of equations (2.11), (A.8) and (A.9), and carry out inference based on (A.4).

Appendix B. Data used in Example 1

Fail #	Time	Outcome	Masking			Diagnostic Prob.		
1	.0183	1						
2	.0357	1	3					
3	.0427	-1	1	3		0.992		0.008
4	.0447	1						
5	.0735	-1	1	2	3	0.765	0.217	0.017
6	.119	1	2	3				
7	.131	-1	1	2	3	0.718	0.245	0.037
8	.143	-1	1	3		0.957		0.043
9	.171	1	3					
10	.261	-1	1	3		0.901		0.099
11	.316	1	3					
12	.334	1						
13	.368	2						
14	.475	1						
15	.484	2						
16	.489	2						
17	.494	-1	1	3		0.783		0.217
18	.594	1						
19	.604	1	2	3				
20	.664	2						
21	.697	1						
22	.712	1	3					
23	.743	1	2	3				
24	.743	2	1	3				
25	.749	-1	1	2	3	0.444	0.262	0.294
26	.752	2						
27	.789	2						
28	.831	1	2	3				

29	.833	3	1	2				
30	.867	-1	1	2	3	0.410	0.253	0.336
31	.874	2						
32	.890	-1	1	3		0.604		0.396
33	.955	3						
34	1.03	-1	1	3		0.553		0.447
35	1.04	3						
36	1.10	3	1	2				
37	1.11	1	3					
38	1.17	-1	1	2	3	0.340	0.230	0.430
39	1.21	-1	1	2	3	0.332	0.227	0.441
40	1.30	1	2	3				
41	1.31	1						
42	1.31	3	1	2				
43	1.33	3						
44	1.35	-1	1	2	3	0.305	0.217	0.478
45	1.35	-1	1	2	3	0.305	0.217	0.478
46	1.37	1	3					
47	1.42	3	1					
48	1.46	3	1					
49	1.49	1	2	3				
50	1.53	1						
51	1.57	-1	1	3		0.400		0.600
52	1.58	-1	1	3		0.398		0.602
53	1.58	-1	1	3		0.397		0.603
54	1.61	-1	1	3		0.392		0.608
55	1.61	-1	1	2	3	0.266	0.199	0.535
56	1.63	-1	1	2	3	0.263	0.198	0.539
57	1.64	3	1	2				
58	1.66	3						
59	1.67	-1	1	2	3	0.257	0.195	0.548
60	1.68	1	3					
61	1.68	3	1					
62	1.68	2	1	3				
63	1.71	3	1	2				
64	1.72	3	1	2				
65	1.73	1						
66	1.75	2						
67	1.82	3	1	2				
68	1.88	-1	1	3		0.338		0.662
69	2.00	-1	1	3		0.318		0.682
70	2.03	-1	1	2	3	0.215	0.174	0.611
71	2.04	-1	1	3		0.313		0.687
72	2.04	3	1	2				

73	2.05	1	3					
74	2.05	1	3					
75	2.07	-1	1	3		0.307		0.693
76	2.09	3	1					
77	2.12	3	1	2				
78	2.15	2						
79	2.18	3	1	2				
80	2.22	3						
81	2.23	1						
82	2.26	3						
83	2.26	-1	1	3		0.281		0.719
84	2.27	-1	1	2	3	0.193	0.161	0.646
85	2.29	-1	1	2	3	0.192	0.161	0.648
86	2.30	1						
87	2.35	-1	1	2	3	0.186	0.158	0.656
88	2.37	-1	1	2	3	0.185	0.157	0.658
89	2.39	1						
90	2.40	2	1	3				
91	2.41	-1	1	3		0.263		0.737
92	2.41	3						
93	2.41	2						
94	2.43	-1	1	2	3	0.181	0.154	0.665
95	2.49	3	1					
96	2.52	-1	1	3		0.250		0.750
97	2.57	-1	1	2	3	0.170	0.148	0.682
98	2.64	3	1					
99	2.66	-1	1	2	3	0.164	0.144	0.692
100	2.67	3	1					
101	2.70	-1	1	3		0.232		0.768
102	2.74	-1	1	3		0.223		0.772
103	2.75	-1	1	3		0.227		0.773
104	2.75	3	1					
105	2.76	3	1					
106	2.79	3						
107	2.82	3	1					
108	2.83	-1	1	3		0.220		0.780
109	2.85	3	1					
110	2.86	3	1					
111	2.87	-1	1	2	3	0.152	0.136	0.712
112	2.95	2						
113	2.96	-1	1	2	3	0.146	0.133	0.721
114	2.98	3						
115	3.03	-1	1	3		0.204		0.796
116	3.04	1						

117	3.07	1	2	3				
118	3.07	1	2	3				
119	3.08	-1	1	2	3	0.140	0.129	0.731
120	3.11	2	1	3				
121	3.12	-1	1	3		0.196		0.804
122	3.13	3	1					
123	3.15	3	1					
124	3.15	-1	1	3		0.194		0.806
125	3.16	3	1	2				
126	3.16	2						
127	3.18	3	1					
128	3.19	3	1					
129	3.21	-1	1	2	3	0.134	0.125	0.741
130	3.24	3	1					
131	3.25	2						
132	3.27	-1	1	3		0.186		0.814
133	3.28	-1	1	2	3	0.130	0.123	0.747
134	3.37	-1	1	2	3	0.127	0.120	0.753
135	3.42	2	1	3				
136	3.43	-1	1	2	3	0.124	0.118	0.758
137	3.46	1	3					
138	3.51	3						
139	3.63	3	1	2				
140	3.65	3	1	2				
141	3.65	3	1					
142	3.65	-1	1	3		0.163		0.837
143	3.67	-1	1	3		0.162		0.838
144	3.67	1	3					
145	3.70	3	1					
146	3.72	1						
147	3.77	3	1	2				
148	3.77	3	1	2				
149	3.77	-1	1	2	3	0.111	0.109	0.780
150	3.78	3	1					
151	3.79	-1	1	3		0.155		0.845
152	3.79	-1	1	2	3	0.111	0.109	0.781
153	3.80	1	2	3				
154	3.81	3	1	2				
155	3.81	-1	1	2	3	0.110	0.108	0.782
156	3.83	1						
157	3.83	-1	1	2	3	0.109	0.108	0.783
158	3.84	-1	1	2	3	0.109	0.108	0.783
159	3.84	2	1	3				
160	3.85	3	1	2				

161	3.86	-1	1	2	3	0.109	0.107	0.784
162	3.86	-1	1	3		0.152		0.848
163	3.87	-1	1	2	3	0.108	0.107	0.785
164	3.89	-1	1	3		0.150		0.850
165	3.91	-1	1	3		0.150		0.850
166	3.91	-1	1	3		0.149		0.851
167	3.93	3	1	2				
168	3.95	3	1	2				
169	3.95	-1	1	2	3	0.105	0.105	0.790
170	3.96	3	1	2				
171	3.98	3						
172	3.98	3	1					

Devices censored at 4 years: 9828

Acknowledgments

We thank the Editor, Associate Editor and two anonymous referees for help in improving the quality of the manuscript. We also appreciate Stephen K. Magie's (IBM, San Jose) help in construction of the Example 1.

Note

1. Betty J. Flehinger passed away while this paper was being prepared.

References

M. H. Aboul-Seoud and J. S. Usher, "The effect of modular system structures on component reliability estimation," *5th Industrial Engineering Research Conference Proc.* pp. 464–467, 1996.

J. W. Andersen and L. M. Ryan, "Analysis of Competing Risk Data with Missing Failure Type when Missingness Depends on Covariates," *Invited Paper, presented at the Joint Statistical Meetings*, Dallas, TX., 1998.

S. Basu, A. Basu, and C. Mukhopadhyay, "Bayesian analysis for masked system failure data using non-identical Weibull models," *J. Stat. Planning and Inference* vol. 78 pp. 255–275, 1999.

G. E. Dinse, "Nonparametric prevalence and mortality estimators for animal experiments with incomplete cause-of-death data," *J. Amer. Statist. Assoc.* vol. 81 pp. 328–336, 1986.

N. Doganaksoy, "Interval estimation from censored and masked system failure data," *IEEE Trans. Reliability* vol. 40 pp. 280–285, 1991.

B. J. Flehinger, B. Reiser, and E. Yashchin, "Inference about defects in the presence of masking," *Technometrics* vol. 38 pp. 247–256, 1996.

B. J. Flehinger, B. Reiser, and E. Yashchin, "Survival with competing risks and masked causes of failures," *Biometrika* vol. 85 pp. 151–164, 1998.

E. Goetghebeur and L. Ryan, "A modified log rank test for competing risks with missing failure type," *Biometrika* vol. 77 pp. 207–211, 1990.

E. Goetghebeur and L. Ryan, "Analysis of competing risks survival data when some failure types are missing," *Biometrika* vol. 82 pp. 821–833, 1995.

F. M. Guess, J. S. Usher, and T. J. Hodgson, "Estimating system and component reliabilities under partial information on the cause of failure," *J. Statist. Plann. Inference* vol. 29 pp. 75–85, 1991.

I. Guttman, D. K. Lin, B. Reiser, and J. S. Usher, "Dependent masking and system life data analysis: Bayesian inference for two-component systems," *Lifetime Data Analysis* vol. 1 pp. 87–100, 1995.

R. K. Kodell and J. J. Chen, "Handling cause of death in equivocal cases using the EM algorithm," *Commun. Statist. A – Theory Methods* vol. 16 pp. 2565–2603, 1987.

L. Kuo and T. Y. Yang, "Bayesian reliability modeling for masked system lifetime data," *Stat. and Prob. Letters* vol. 47 pp. 229–241, 2000.

M. J. van der Laan and I. W. McKeague, "Efficient estimation from right - censored data when failure indicators are missing at random," *Annals of Statistics* vol. 26:1 pp. 164–182, 1998.

G. Lapidus, M. Braddock, R. Schwartz, L. Banco, and L. Jacobs, "Accuracy of fatal motorcycle injury reporting on death certificates," *Accident Anal. and Prevention* vol. 26:4 pp. 535–542, 1994.

D. K. J. Lin and F. M. Guess, "System life data analysis with dependent partial knowledge on the exact cause of system failure," *Microelectron. Reliability* vol. 34 pp. 535–544, 1994.

S. H. Lo, "Estimating a survival function with incomplete cause-of-death data," *J. Multivariate Anal.* vol. 39 pp. 217–235, 1991.

G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, Wiley: New York, 1997.

W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*, Wiley: New York, 1998.

M. Miyakawa, "Analyses of incomplete data in competing risk model," *IEEE Trans. Reliability* vol. 33 pp. 293–296, 1984.

C. Mukhopadhyay and A. P. Basu, "Bayesian analysis of incomplete time and cause of failure data," *J. Stat. Planning and Inference* vol. 59 pp. 79–100, 1997.

B. Reiser, I. Guttman, D. K. J. Lin, J. S. Usher, and F. M. Guess, "Bayesian inference for masked system life time data," *Applied Statistics* vol. 44 pp. 79–90, 1995.

B. Reiser, B. J. Flehinger, and A. R. Conn, "Estimating component defect probability from masked system success/failure data," *IEEE Trans. Reliability* vol. 45 pp. 238–243, 1996.

B. Reiser, E. Yashchin, and B. J. Flehinger, "Statistical inference for masked data," *Proc. 2nd World Congress of Nonlinear Analysis*, pp. 4425–4432, 1997.

D. R. Thomas and G. L. Grunkemeier, "Confidence interval estimation of survival probabilities for censored data," *J. Amer. Statist. Assoc.* vol. 70 pp. 865–871, 1975.

P. Tobias and D. Trinidade, *Applied Reliability*, Van Nostrand Reinhold: New York, 1986.

J. S. Usher, "Weibull component reliability - prediction in the presence of masked data," *IEEE Trans. Reliability* vol. 45 pp. 229–232, 1996.

J. S. Usher and F. M. Guess, "An iterative approach for estimating component reliability from masked system life data," *Qual. Reliab. Engineering Intnl.* vol. 5 pp. 257–261, 1989.

J. S. Usher and T. J. Hodgson, "Maximum likelihood estimation of component reliability using masked system life test data," *IEEE Trans. Reliability* vol. 37 pp. 550–555, 1988.