COMPETING RISKS METHODS

Núria Porta[(1)], Guadalupe Gómez[(1)], M.Luz Calle[(2)] and Núria Malats[(3)]

[(1)] *Dept. of Statistics and Operations Research*
Universitat Politècnica de Catalunya, Barcelona (Spain)
[(2)] *Dept. of Systems Biology, Universitat de Vic, Vic (Spain)*
[(3)] *Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid (Spain)*

DR 2007/14
10 December 2007

Copies of this report may be downloaded at http://www-eio.upc.es/%7Enporta/.

corresponding author: Núria Porta
Dept. Statistics and Operations Research,
UPC, Campus Nord, C5-224,
c. Jordi Girona 1-3, 08034 Barcelona,
tel. +34 934054095 fax +34 934015855
email:nuria.porta-bleda@upc.edu

Contents

Competing Risks Methods[1]

N. Porta, G. Gómez, M.L. Calle and N.Malats

Abstract: *Competing risks data usually arises in studies in which the failure of an individual may be classified into one of k ($k > 1$) mutually exclusive causes of failure. When competing risks are present, classical survival analysis techniques may not be appropriate to use. The main goal of this paper is to review the specific methods to deal with competing risks. To this aim, we first focus on how to specify a competing risks model, which is the structure of observed data in this framework, and how components of the model are estimated from a given random sample. In addition, we discuss how to correctly interpret probabilities in the presence of competing risks, and regression models are considered in detail. To conclude, we illustrate the problem with data from a bladder cancer study.*

Keywords: *Competing risks; cause-specific hazards; cumulative incidence function.*

1 Introduction

Though classical survival analysis methods for exploring time-to-event data are well developed, there are complex situations where such techniques are not appropriate. One of such situations is competing risks. Competing risks data arises when an individual may fail from different causes. The occurrence of a failure due to a specific cause may or may not preclude the occurrence of failures due to other causes. Competing risks methods deal with both situations by analyzing the time to the first event happening.

Examples of competing risks data are found nowadays in many fields. In a demographic study where the leading causes of death are registered -heart disease, cancer,...- the interest might focus on analyzing each of them separately. In a clinical trial addressed to find the benefits of a new drug to prevent myocardial infarction, patients with coronary heart disease are followed during two years. The failure of interest is myocardial infarction though patients may die from other causes. In reliability, failure may correspond, for example, to breakdown of a mechanical device where there are several causes for the failure, such as vibration or corrosion. Hence, the distinguishing feature of a competing risks setting is that for each individual, besides a lifetime T, there is a mode failure C, and a joint model for T and C is needed.

The joint distribution of (T, C) might be completely specified through the cause-specific hazard, that is, the instantaneous risk of failing at a given time from a given cause, among all individuals at risk at that time. The joint distribution can also be specified through the cumulative incidence function, representing the probability of failing from a given cause before a specific time. These two functions represent distinct quantities, and estimating one or the other depends on the research question of interest. Moreover, modelling these two functions leads to different types of regression models when covariates are present.

This report is organized as follows. In Section 2, we specify the competing risks model. We introduce the notation and the key concepts in Subsections 2.1 and 2.2. We derive the likelihood function and obtain non-parametric estimates given a random sample in Subsections 2.3 and 2.4. In Section 3 we define survival-like functions in the framework of competing risks, and discuss how to correctly interpret these quantities. In Section 4, we revise regression modelling of competing risks when covariates are present. Specifically, in Subsection 4.1, models based on the cause-specific hazards are considered. In Subsection 4.2, the most popular model based on cumulative incidence function,

[1]Partially supported by Grant 050831 from La Marató de TV3 Foundation and by grant MTM2005-0886 from the Ministerio de Ciencia y Tecnología. Núria Porta is a recipient of a research fellowship from DURSI.

Fine and Gray (1999)'s model, is explained with some detail. Finally, in Section 5, the reviewed methodologies are illustrated with data from a bladder cancer study.

2 Competing risk data

2.1 Model specification

Define, for each individual, the pair (T, C), where T is the failure time, and C is the failure cause. T is assumed to be a continuous and positive random variable, while C takes values in the finite set $\{1, \ldots, k\}$. Assume that the individual fails from one and only one cause. The joint distribution of (T, C) is completely specified through either the cause-specific hazards, $\lambda_j(t)$, or through the cumulative incidence functions, $F_j(t)$.

The cause-specific hazard function for the j^{th} cause is defined as

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{Pr\Big(T < t + \Delta t, C = j | T \geq t\Big)}{\Delta t} \qquad j = 1, \ldots, k$$

and represents the rate of occurrence of the j^{th} failure.

The cumulative incidence function from type j failure is defined by

$$F_j(t) = Pr(T \leq t, C = j) \qquad j = 1, \ldots, k, \tag{2.1}$$

and corresponds to the sub-distribution function for the probability of a subject failing from cause j in the presence of all the competing risks.

The cause-specific cumulative hazards $\Lambda_j(t)$, the overall hazard $\lambda(t)$, the overall cumulative hazard $\Lambda(t)$ and the overall survival function $S(t)$ are defined, respectively, as:

$$\begin{aligned}
\Lambda_j(t) &= \int_0^t \lambda_j(t) \qquad j = 1, \ldots, k, \\
\lambda(t) &= \lim_{\Delta t \to 0} \frac{Pr(T < t + \Delta t | T \geq t)}{\Delta t} = \sum_{j=1}^{k} \lambda_j(t), \\
\Lambda(t) &= \int_0^t \lambda(u) du = \sum_{j=1}^{k} \Lambda_j(t), \text{ and} \\
S(t) &= Pr(T > t) = e^{-\Lambda(t)}.
\end{aligned}$$

The survival function can be factorized into the following k functions $S_j(t) = e^{-\Lambda_j(t)}$ as follows

$$S(t) = e^{-\sum_{j=1}^{k} \Lambda_j(t)} = \prod_{j=1}^{k} e^{-\Lambda_j(t)} = \prod_{j=1}^{k} S_j(t). \tag{2.2}$$

Caution is needed when interpreting functions $S_j(t)$. Despite having the mathematical properties of continuous survivor functions, they are not the survivor functions of any observable random variables. Moreover, $S_j(t) \neq 1 - F_j(t)$, as we shall see in future sections 2.3 and 3, where more details are given on their interpretation.

The subdensity functions $f_j(t)$ from cause j, the marginal distribution $F(t)$ of T, and the marginal distribution of C are respectively given by:

$$f_j(t) = \frac{d}{dt}F_j(t) = \lambda_j(t)S(t),$$
$$F(t) = P(T \leq t) = \sum_{j=1}^{k} F_j(t), \text{ and}$$
$$\pi_j(t) = Pr(C = j) = \lim_{t\to\infty} F_j(t) \qquad j = 1, \ldots, k.$$

The cumulative incidence function for cause j, $F_j(t)$, can be obtained from the cause specific hazard λ_j and the overall survival function $S(t)$ from the relationship:

$$F_j(t) = \int_0^t \lambda_j(u)S(u)du \qquad j = 1, \ldots, k. \tag{2.3}$$

2.2 Alternative representations of competing risks

A different way of describing a competing risks situation with k causes of failure is to consider a failure time T_j for each cause, $j \in \{1, \ldots, k\}$. These times are latent variables corresponding to the hypothetical failure times if the other causes of failure were not present. It has been argued that multivariate models $F(t_1, \ldots, t_k)$ could be specified for the joint distribution of $T_1, \ldots, T_k$ (see Kalbfleisch and Prentice, 2002; Lawless, 2003; Andersen *et al.*, 2002 for further references). However, when all risks are present only $T = \min(T_1, \ldots, T_k)$ can be observed, together with $C = j$, such as $T = T_j$, and an identifiable problem is found (Tsiatis, 1975; Cox and Oakes, 1984). $F(t_1, \ldots, t_k)$ is inestimable solely based on these observations. Two different distributions for $F(t_1, \ldots, t_k)$ may result in the same marginal for (T, C). Only under strong assumptions such as independence the multivariate distribution is identifiable. However, this assumption is untestable based solely on observed competing risk data.

On the other hand, competing risks can be viewed as a special case of a multi-state model (Andersen *et al.*, 2002). In this case the multi-state model has one transient state 'Alive' and k absorbing states 'Failure from cause 1', ..., 'Failure from cause k'. The process is Markovian, and in this setting, the goal is to model the transitions between states, through the probabilities $P_{hj}(s, t)$, probability of being in state j at t, provided that at time s, the state h was occupied. Note that $P_{0j}(0, t) = P(T \leq t, C = j)$ are the cumulative incidence functions as defined in section 2.1, whereas the intensity transition functions are the cause-specific hazards.

2.3 Likelihood function

Consider a random sample of n individuals, $(T_1, C_1), \ldots, (T_n, C_n)$, where T_i is the time of failure and C_i is the cause of failure for subject i. For each individual, there exists a non-negative right censoring time V_i. Let $\delta_i = I(T_i \leq V_i)$ be the censoring indicator, and define $\tilde{C}_i = \delta_i C_i$. $\tilde{C}_i$ is the cause of failure for failing individuals or 0 for censored individuals. The observed data for each individual are given by

$$\{Y_i = \min(T_i, V_i), \delta_i, \tilde{C}_i, \ i = 1, \ldots, n\}.$$

In the following, these conditions are assumed:

H1) V_i is independent of (T_i, C_i).

H2) If $Y_i = T_i$ (T_i is not censored), then C_i is observed (we exclude cases when the time of failure is observed, but no information about the cause of failure is available).

H3) The supports of T and V are disjoint.

To derive the likelihood function, the contribution of each individual must be taken into account. To further clarify this point, we first develop the likelihood function for $n = 1$, and then for $n > 1$.

Sample of size $n = 1$:

Only one individual is observed with data $(Y, \delta, \tilde{C})$. Two scenarios are then possible:

A) The individual is not censored, $\delta = 1 \Longrightarrow (Y, \delta, \tilde{C}) = (y, 1, j)$, $j \neq 0$ (H2):
For the discrete case,

$$P\{y, 1, j\} = P(Y = y, \delta = 1, C = j) = P(T = y, T \leq V, C = j) = P(T = y, V \geq y, C = j)$$
$$\underset{H1}{=} P(T = y, C = j)P(V \geq y).$$

If T and V are continuous random variables, and $q(v)$ and $Q(v)$ are the density and survival functions for V, the contribution of the individual to the likelihood function would be given by $f_j(y)Q(y)$.

B) If the individual is censored, $\delta = 0 \Longrightarrow (Y, \delta, \tilde{C}) = (y, 0, 0)$:
For the discrete case,

$$P\{y, 0, 0\} = P(Y = y, \delta = 0, \tilde{C} = 0) = P(Y = y, \delta = 0) = P(V = y, T > V) = P(V = y, T > y)$$
$$\underset{H1}{=} P(T > y)P(V = y).$$

If T and V are continuous random variables as defined in A), the contribution of the individual to the likelihood function would be $S(y)q(y)$.

Therefore the likelihood function for a given individual has the form $\mathcal{L} = (f_c(y)Q(y))^{\delta}(S(y)q(y))^{1-\delta}$.

Sample of size $n > 1$:

We evaluate the contribution of each individual to the likelihood. If subject i fails at y_i by cause $c_i = j$, then his/her contribution to the likelihood will be given by $f_j(y_i)$, the cause-specific density function for cause j. On the other hand, if individual i is censored at time y_i, the individual is still at risk for any cause and hence his/her contribution to the likelihood is given by the overall survival $S(y_i)$.

The likelihood function for the sample is given by

$$\mathcal{L} = \prod_{i=1}^{n} f_{c_i}(y_i)^{\delta_i} S(y_i)^{1-\delta_i} \prod_{i=1}^{n} Q(y_i)^{\delta_i} q(y_i)^{1-\delta_i}.$$

Since the censoring time V is independent from the failure time T, and their supports are disjoint (assumptions H1 and H3 respectively), the censoring terms in the likelihood do not provide information on the failure process and can be removed. The likelihood function is then proportional to

$$\mathcal{L} \propto L = \prod_{i=1}^{n} f_{c_i}(y_i)^{\delta_i} S(y_i)^{1-\delta_i}.$$

Denote by $\delta_{ij} = I(C_i = j)$, where $\delta_i = \sum_{j=1}^k \delta_{ij}$. If $\delta_i = 1$, then it exists some j with $\delta_{ij} = 1$. From the factorization of the survival $S(t) = \prod_{j=1}^k S_j(t)$ (see (2.2)), and defining $g_j(t) = -S_j'(t) = \lambda_j(t)S_j(t)$, the likelihood function can be rewritten as product of k separate components for each failure cause:

$$
\begin{aligned}
L &= \prod_{i=1}^n \left(\prod_{j=1}^k f_j(y_i)^{\delta_{ij}} \right) S(y_i)^{1-\delta_i} = \prod_{i=1}^n \left(\prod_{j=1}^k \left(\lambda_j(y_i) S(y_i) \right)^{\delta_{ij}} \right) S(y_i)^{1-\delta_i} \\
&= \prod_{i=1}^n \left\{ \left(\prod_{j=1}^k \left(\lambda_j(y_i) S(y_i) \right)^{\delta_{ij}} \right) \left(\prod_{j=1}^k S_j(y_i)^{1-\delta_i} \right) \right\} \\
&= \prod_{i=1}^n \prod_{j=1}^k \left(\lambda_j(y_i)^{\delta_{ij}} \left[\prod_{\ell=1}^k S_\ell(y_i)^{\delta_{ij}} \right] S_j(y_i)^{1-\delta_i} \right) \\
&= \prod_{i=1}^n \prod_{j=1}^k \lambda_j(y_i)^{\delta_{ij}} S_j(y_i)^{\delta_{ij}} S_j(y_i)^{\sum_{\ell \neq j} \delta_{i\ell} + 1 - \delta_i} \\
&= \prod_{i=1}^n \prod_{j=1}^k g_j(y_i)^{\delta_{ij}} S_j(y_i)^{1-\delta_{ij}} = \prod_{j=1}^k \left(\prod_{i=1}^n g_j(y_i)^{\delta_{ij}} S_j(y_i)^{1-\delta_{ij}} \right) = \prod_{j=1}^k L_j. \qquad (2.4)
\end{aligned}
$$

Expression (2.4) provides a factorization of the overall likelihood L in terms of cause-specific likelihoods L_j. Note that L_j corresponds to the likelihood it would be obtained from sample $\{(Y_i, \delta_{ij}),\ i = 1 \ldots, n\}$ if failure times from other causes were considered as censoring times, and where the corresponding hazard, density and survival functions being, respectively, $\lambda_j(t)$, $g_j(t)$ and $S_j(t)$. This factorization shows how $\lambda_j(t)$ and $\Lambda_j(t)$ are directly estimable from data (Y_i, δ_{ij}), by treating failures from other causes at Y_i as censored observations. However, as we have already mentioned, $S_j(t)$ does not correspond to ANY observable random variable, that is, it does not exist any observed random variable U_j such that $P(U_j > t) = S_j(t)$.

2.4 Nonparametric estimation

Consider a random sample with observed data $\{(Y_i, \delta_i, \tilde{C}_i), i = 1, \ldots, k\}$, defined as in the previous section 2.3. Let $0 < y_1 < \cdots < y_N$ be the ordered distinct observed time points. We define the following quantities:

- d_{ij} is the number of subjects failing from cause j at time y_i,
- $d_i = \sum_{j=1}^k d_{ij}$ is the number of subjects failing at time y_i from any cause,
- $n_i = \sum_{\ell=1}^n I_\ell(y_i)$, with $I_\ell(t) = I(y_\ell \geq t)$, is the number of individuals at risk at y_i, that is, alive and uncensored just prior to this time.

The estimate of the cause-specific hazard for cause j at time y_i is given by $\hat{\lambda}_j(y_i) = \frac{d_{ij}}{n_i}$, and it is 0 at any other time. Hence, the Nelson-Aalen estimator for the cumulative cause-specific hazard function is given by

$$
\hat{\Lambda}_j(t) = \sum_{i: t_i \leq t} \frac{d_{ij}}{n_i} \qquad j = 1, \ldots, k,
$$

with variance estimated by

$$\widehat{\text{Var}}[\hat{\Lambda}_j(t)] = \sum_{i:t_i \le t} \frac{d_{ij}}{n_i^2} \qquad j = 1, \ldots, k.$$

The overall survival function for T can be obtained by using the Kaplan-Meier estimate:

$$\hat{S}(t) = \prod_{i:y_i<t} \left(1 - \frac{d_i}{n_i}\right)^{\delta_i}.$$

Alternatively, $S(t)$ can be obtained through $\hat{S}(t) = \exp\left[-\sum_{j=1}^{k} \hat{\Lambda}_j(t)\right]$.
Since the cumulative incidence function for cause j can be obtained from the cause specific hazard trough $F_j(t) = \int_0^t \lambda_j(u)S(u)du$ (2.3), a natural non-parametric estimate of $F_j(t)$ is

$$\hat{F}_j(t) = \int_0^t \hat{\lambda}_j(u)\hat{S}(u)du = \sum_{i:y_i \le t} \frac{d_{ij}}{n_i}\hat{S}(y_i^-) \qquad j = 1, \ldots, k.$$

3 Interpreting probabilities in competing risks

We have reported in the previous sections how to specify a competing risk model through the cause-specific hazards $\lambda_j(t)$ or via the cumulative incidence functions $F_j(t)$, and how to estimate non-parametrically them. Compared to classical survival analysis, it seems odd to use the cumulative incidence function $F_j(t)$ instead of some type of *cause-specific survival function* for cause j. In classical survival analysis, a lifetime endpoint T is usually described by its survival function $S(t) = P(T > t)$, which satisfies that $S(t) = 1 - F(t)$, $F(t)$ being its distribution function. The survival function could be derived from the hazard function of T, $\lambda(t)$, by

$$S(t) = e^{-\int_0^t \lambda(u)du}.$$

By analogy, in competing risks, given the cause-specific hazard for cause j, $\lambda_j(t)$, a similar function $S_j(t)$ could be considered for each cause of failure:

$$S_j(t) = e^{-\int_0^t \lambda_j(u)du} = e^{-\Lambda_j(t)},$$

(see (2.2)). However, these functions do not have the usual meaning of a survival function in the classical approach. Furthermore, the functions $S_j(t)$ do not correspond to the complementary of the incidence function $F_j(t)$, that is, $S_j(t) \neq F_j(t)$, neither to the joint probability of failing from cause j after t, $P(T > t, C = j)$. These considerations lead us to define two more functions that may play the role of cause-specific survivals.

On one hand, we define $S^*(t)$ as the complement of the cumulative incidence function

$$S_j^*(t) = 1 - F_j(t),$$

on the other hand, and by analogy with the definition of $F_j(t)$, we define

$$\tilde{S}_j(t) = P(T > t, C = j).$$

These two new functions are, as it was $S_j(t)$, survival-like functions, that is, functions which satisfy the mathematical properties of a survival function. In the following, we will deepen on the interpretation of these three functions,$S_j(t)$, $S_j^*(t)$ and $\tilde{S}_j(t)$, and argue why they are not proper survival functions, and which is the relationship among them. For this matter it is worthwhile to remind that a function $S(t)$ is a survival function if

i. it is defined in $[0, \infty)$,

ii. it is non-negative and non-increasing,

iii. it is right-continuous,

iv. $S(0) = 1$ and $lim_{t\to\infty} S(t) = 0$.

In addition, $S(t)$ is a survival function of a random variable T if $S(t) = P(T > t)$.

3.1 Interpretation of function $\mathbf{S_j^*(t) = 1 - F_j(t)}$

$S_j^*(t) = 1 - F_j(t)$ represents the probability of not failing from cause j before t. It is not a proper survivor function because

$$\lim_{t\to\infty} S_j^*(t) = 1 - \lim_{t\to\infty} F_j(t) = 1 - P(C = j),$$

which is strictly positive if there are at least two causes of failure. Moreover,

$$S_j^*(t) = 1 - F_j(t) = 1 - F(t) + \sum_{\ell\neq j} F_\ell(t) = S(t) + \sum_{\ell\neq j} P(T \leq t, C = \ell).$$

That is, the probability of not failing from cause j before t is the sum of the probability of having not failed for any cause by t plus the probability of having failed before t from other causes than j. This probability $S_j^*(t)$ is used to build Fine and Gray's regression model for the cumulative incidence function (see 4.2.1).

3.2 Interpretation of function $\mathbf{\tilde{S}_j(t) = P(T > t, C = j)}$

By analogy with the way cumulative incidence functions F_j were defined, $\tilde{S}_j(t) = P(T > t, C = j)$ represents the probability of failing from cause j after t. It is not a proper survivor function because

$$\tilde{S}_j(0) = P(C = j)$$

which is strictly below 1 if there are at least two causes of failure.

The relationship with $F_j(t)$ is given by

$$\begin{aligned}\tilde{S}_j(t) &= P(T > t, C = j) = Pr(T > j|C = j)P(C = j) = \left[1 - P(T \leq t|C = j)\right] P(C = j) \\ &= P(C = j) - P(T \leq t, C = j) = P(C = j) - F_j(t).\end{aligned}$$

Hence, it behaves like a complementary probability for $F_j(t)$, complementary on the probability of failing from cause j, $P(C = j)$. Note as well that $S(t)$ could be decomposed in terms of $\tilde{S}_j(t)$ as follows:

$$S(t) = 1 - F(t) = 1 - \sum_{j=1}^{k} F_j(t) = 1 - \sum_{j=1}^{k} P(C = j) + \sum_{j=1}^{k} P(T > t, C = j) = \sum_{j=1}^{k} \tilde{S}_j(t).$$

The expression of $S(t)$ as a sum of $\tilde{S}_j(t)$ is indeed different from the alternative decomposition $S(t) = \prod_{j=1}^{k} S_j(t)$ (see (2.2)), and shows that $\tilde{S}_j(t)$ and $S_j(t)$ are different.

A consistent estimate for $\tilde{S}_j(t)$ is given by

$$\widehat{\tilde{S}}_j(t) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i > t, C_i = j) \qquad j = 1, \ldots, k.$$

Despite these are estimable functions and could specify the competing risks model, they have been scarcely used in the competing risks literature (Peterson, 1976).

3.3 Interpretation of function $\mathbf{S_j(t) = e^{-\Lambda_j(t)}}$

We have came across functions $S_j(t)$ repeatedly in the previous sections. Firstly, we encountered them in the factorization of the survival function $S(t)$ (2.2). Later, in the factorization of the likelihood function (2.4), where $S_j(t)$ corresponds to the survival function obtained from the the cumulative hazard function $\Lambda_j(t)$, the cumulative risk when failure times from other causes are treated as censoring times. The functions $S_j(t)$ hold the mathematical properties of a survival function, however they are not survival functions of any observable random variable.

When failures from other causes are treated as censored observations, the assumption of independence between failure time and censoring time is possibly violated. Thus, only when distinct causes of failure are assumed to be independent, $1 - S_j(t)$ is fully interpretable as the probability of failing from cause j if the other causes of failure were removed (Gooley *et al.*, 1999).

Often $1 - S_j(t)$ has been used incorrectly to estimate $F_j(t)$, partly because of the availability of software to obtain the Kaplan-Meier estimate for $S_j(t)$:

$$\mathrm{KM}_j(t) = \prod_{i:y_i<t} \left(1 - \frac{d_{ji}}{n_i}\right)^{\delta_{ij}},$$

where d_{ij}, n_i and δ_{ij} are defined in section 2.4 and failures from other causes are treated as censored observations. However, $1 - \mathrm{KM}_j(t)$ provides a biased estimate of the cumulative probability of failure from type j, $F_j(t)$ (Putter *et al.*, 2007). This is clear intuitively since $S_j(t)$ only depends on the cause-specific hazard $\lambda_j(t)$, whereas $F_j(t)$ depends on all cause-specific causes $\hat{\lambda}_\ell(t)$, $\ell \in \{1, \ldots, k\}$ through the survival function $S(t)$ (see 2.3). Moreover, $1 - \mathrm{KM}_j(t)$ as an estimate of $1 - S_j(t)$ overestimates the probability of failure from cause j, $F_j(t)$. This is reasonable, because if an individual failing from other causes is treated as a censored observation, one assumes that the individual WILL fail from the cause of interest j somewhen in the future, which in some situations may be unfeasible: if an individual dies due to cancer, he/she would not certainly die (again) due to a heart attack. By censoring individuals, we expect a higher incidence of failures. In effect, there always exist $t^* > 0$ such as

$$F_j(t^*) = \int_0^{t^*} S(u)\lambda_j(u)du < \int_0^{t^*} S_j(u)\lambda_j(u) = 1 - S_j(t^*).$$

Proof:

- It always exists $\ell \neq j$ and $t^* > 0$ such as $\Lambda_\ell(t^*) = \int_0^{t^*} \lambda_\ell(u)du > 0$. That is, there exists at least one other cause of failure with at least one failure. Otherwise, there are not competing risks in our data.
- Therefore, $\Lambda(t^*) = \sum_{m=0}^{k} \Lambda_m(t^*) > \Lambda_j(t^*)$.
- Being $g(u) = e^{-u}$ non-increasing and $\Lambda_j(u)$ non-negative,

$$S_j(t^*) = e^{-\Lambda_j(t^*)} > e^{-\Lambda(t^*)} = S(t^*). \qquad \square$$

4 Regression modelling of Competing risks

In a survival analysis with competing risks, two different regression modelling strategies are possible: modelling the cause-specific hazards or modelling the cumulative incidence functions.

When the cause-specific hazards are modelled, each hazard is analysed separately by treating individuals failing from other causes as censored observations, as follows from the factorization of the likelihood function (see 2.3). This approach is appropriate when determining factors associated to the risk of a specific cause of failure is of interest.

On the other hand, the cumulative incidence functions are used to determine factors associated to the incidence of a given cause. This analysis does not treat individuals failing from other causes as censored observations.

In the following sections, three specific models are revised: Cox's proportional hazards model (Prentice *et al.*, 1978) and Aalen's additive model (Aalen, 1993), which specify models for the cause-specific hazards, and the approach given by Fine and Gray (1999), which is based on the cumulative incidence functions.

4.1 Modelling the cause-specific hazards $\lambda_j(t)$

4.1.1 Cox's proportional hazards model

The classical regression analysis of competing risks establishes a Cox proportional hazards (PH) model (Prentice *et al.*, 1978) for each cause-specific hazard:

$$\lambda_j(t|\mathbf{Z}) = \lambda_{0j} e^{\beta_j' \mathbf{Z}} \qquad j = 1, \ldots, k,$$

where $\mathbf{Z}$ is a $p \times 1$ vector of covariates and $\boldsymbol{\beta_j}$ is a $p \times 1$ vector of regression coeficients for each cause. Each cause of failure is analysed separately, treating individuals failing from other causes as censored observations. The effect of the covariates is assumed to act multiplicatively on an unknown baseline hazard function λ_{0j}. As in classical PH analysis, the validity of the models does not depend on the true form of the baseline hazard, provided the multiplicative form of the model is correct. The PH assumption is a strong one that must be carefully checked for each cause.

Estimation of the regression parameters $\boldsymbol{\beta}_j$ is based on the partial likelihood approach. Let's suppose that a censored random sample $(y_i, \delta_i, \delta_i c_i)$, $i = 1, \ldots, n$, yields N distinct observed times of failure $t_1 < \cdots < t_N$ and $n - N$ censored times (no ties considered here). Consider the probability that an individual fails by cause j at time t_i , given that one of the individuals at risk (alive and uncensored) at time t_i fails by cause j:

$$\frac{e^{\beta_j' \mathbf{Z}_i}}{\sum_{\ell=1}^{n} Y_\ell(t_i) e^{\beta_j' \mathbf{Z}_\ell}},$$

where $Y_\ell(t) = I(t_\ell \geq t)$. The partial likelihood function is defined only in the N times of failure, yielding:

$$L(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k) = \prod_{i=1}^{n} \prod_{j=1}^{k} \left(\frac{e^{\beta_j' \mathbf{Z}_i}}{\sum_{\ell=1}^{n} Y_\ell(t_i) e^{\beta_j' \mathbf{Z}_\ell}} \right)^{\delta_{ij}} = \prod_{j=1}^{k} L_j(\boldsymbol{\beta}_j) \tag{4.1}$$

where $\delta_{ij} = I(C_i = j)$. The risk set can be diminished by the occurrence of an event from any cause.

Maximizing each factor in (4.1) provides an estimator $\hat{\boldsymbol{\beta}}_j$ consistent and asymptotically normal under suitable conditions, and score, information and likelihood ratio statistics based on $L(\hat{\boldsymbol{\beta}}_j)$ behave as if they were deduced from ordinary likelihood.

Given $\hat{\boldsymbol{\beta}}_j$, the generalized Nelson-Aalen estimates for the cause-specific baseline cumulative hazard functions are:

$$\hat{\Lambda}_{0j}(t) = \sum_{i:t_\ell \le t} \left(\frac{\delta_{ij}}{\sum_{\ell=1}^{n} Y_\ell(t_i) e^{\hat{\beta}_j' \mathbf{Z}_\ell}} \right) \qquad j = 1, \ldots, k.$$

Inference for the $\boldsymbol{\beta}_j$'s and for the Λ_{0j}'s can be conducted then as in the standard Cox model where a single cause of failure is considered. Overall survival and cumulative hazard functions for T given $\boldsymbol{Z}$ are obtained by

$$\hat{S}(t|\boldsymbol{Z}) = \exp\left\{ -\sum_{j=1}^{k} \hat{\Lambda}_{0j}(t) e^{\hat{\beta}_j' \mathbf{Z}} \right\} \text{ and}$$

$$\hat{\Lambda}_j(t|\boldsymbol{Z}) = \hat{\Lambda}_{0j}(t) e^{\hat{\beta}_j' \mathbf{Z}} \qquad j = 1, \ldots, k.$$

Finally, the cumulative incidence function $F_j(t|\boldsymbol{Z})$ can be obtained by plugging-in the previous estimates in equation (2.3):

$$\begin{aligned} \hat{F}_j(t|\boldsymbol{Z}) &= \int_0^t \hat{S}(u|\boldsymbol{Z}) d\hat{\Lambda}_j(u|\boldsymbol{Z}) \\ &= \sum_{i:t_i \le t} \delta_{ij} \exp\left\{ -\sum_{\ell=1}^{k} \hat{\Lambda}_{0\ell}(u) e^{\hat{\beta}_\ell' \mathbf{Z}} \right\} \frac{e^{\hat{\beta}_j' \mathbf{Z}}}{\sum_{r=1}^{n} Y_r(t_i) e^{\hat{\beta}_j' \mathbf{Z}_r}} \end{aligned}$$

The methodology proposed is completely standard, but some caution is needed when interpreting the models. Even when only one cause of failure is of interest, it is not sufficient to perform a single analysis for this cause. It is necessary to model all causes of failures in order to perform a full and appropriate interpretation of the failure process (see section 5.2.1 for an example).

4.1.2 Aalen's additive hazards model

Cox methodology has been widely discussed, becoming the standard analysis to perform in regression modelling (see Lawless, 2003, for example). However, in long follow up studies, it is natural to think that the effect of a covariate in the hazard can change over time. In this situation, models with constant parameters along time, such as the Cox model, may be inappropriate. An alternative is found in the methodology proposed by Aalen (1993, 2001), in which an additive hazards model for each cause-specific hazard is specified:

$$\lambda_j(t|\boldsymbol{Z}(t)) = \beta_{j0}(t) + \boldsymbol{Z}(t)^t \boldsymbol{\beta}_j(t) \qquad j = 1, \ldots, k$$

where $\boldsymbol{Z}^t(t) = [Z_1(t), \ldots, Z_p(t)]$ is a p-vector of (possibly time-dependent) covariates and $\boldsymbol{\beta}_j^t(t) = [\beta_{j1}(t), \ldots, \beta_{jp}(t)]$ are unknown parameter functions. Unlike in the Cox setting, this model assumes that the covariates act in an additive manner on the unknown baseline hazard function. This effect is assessed through the time-dependent functions $\boldsymbol{\beta}_j(t)$, so its variation along time can be explored.

Least-squares techniques are used to derive estimates of the cumulative effects

$$\mathbf{B}(t) = (B_0(t), B_1(t), \ldots, B_p(t)) = \left(\int_0^t \beta_0(u) du, \int_0^t \beta_1(u) du, \ldots, \int_0^t \beta_p(u) du \right)$$

because they are easier to obtain, and then estimates of $\boldsymbol{\beta}(t)$ can be obtained by the slopes of $\hat{\mathbf{B}}(t)$, or by using a kernel smooth estimator based on them.

Counting processes notation will be used to derive those estimates. Consider the i^{th} individual, and define:

- For each cause j, $N_{ji}(t) = I(Y_i \leq t, C_i = j)$ is the function determining if the j^{th} cause of failure already occurred at time t.
- Consider $\mathbf{N}_j(t) = (N_{ji}(t), i = 1, \ldots, n)'$.
- The risk indicator for the i^{th} individual is defined as $Y_i(t) = I(Y_i \geq t)$.
- Consider the $n \times (p+1)$ design matrix $\boldsymbol{X}(t)$ with the i^{th} row given by

$$\boldsymbol{X}_i(t) = Y_i(t)(1, Z_{i1}(t), \ldots, Z_{ip}(t)).$$

  So $\boldsymbol{X}_i(t) = (1, Z_{i1}(t), \ldots, Z_{ip}(t))$ if subject i is at risk at time t, and a $(p+1)$-zero vector if the subject is not at risk.

The least-squares estimates of $\mathbf{B}_j(t) = (B_{j0}(t), B_{j1}(t), \ldots, B_{jp}(t))$ for the regression parameters for cause j are given by:

$$\hat{\mathbf{B}}_j(t) = \int_0^t \boldsymbol{X}^-(u) dN_j(u) \approx \sum_{y_i \leq t} \boldsymbol{X}^-(t_i)\mathbf{I}_j(y_i), \qquad j = 1, \ldots, k$$

where $\boldsymbol{X}^-(t) = (\boldsymbol{X}(t)'\boldsymbol{X}(t))^{-1}\boldsymbol{X}(t)'$ is a generalized inverse of $\boldsymbol{X}(t)$, and $\mathbf{I}_j(t)$ is the $n \times 1$ vector with the i^{th} component equal to 1 if the i^{th} subject fails due to cause j at t, 0 otherwise. Note that the estimator $\hat{\mathbf{B}}_j(t)$ only exists up to the smallest time t_r at which $\boldsymbol{X}(t_r)'\boldsymbol{X}(t_r)$ becomes singular.

The estimated variance-covariance matrix of $\mathbf{B}_j(t)$ is

$$\hat{\boldsymbol{\Sigma}}_j(t) = \sum_{y_i \leq t} \boldsymbol{X}^-(y_i)\mathbf{I}_j^D(t_i)\boldsymbol{X}^-(y_i)'$$

where $\mathbf{I}_j^D(y_i)$ is a diagonal matrix, and the elements in the diagonal are $\mathbf{I}_j$. Other choices of the generalized inverse of $\boldsymbol{X}(t)$ can be used, including those that leads to weighted least squares. As in the previous section, the cumulative incidence function is derived by using the cause-specific hazards and baseline cumulative hazards estimates from the additive model.

Consider the problem of testing the hypothesis of no regression effect for one or more covariates for the j^{th} cause of failure: $H_0 : \beta_{j\ell}(t) = 0$ for all t in the observed period, and ℓ in some set $L \subset \{1, \ldots, p\}$. Aalen (1993) proposed the following test statistic vector

$$\mathbf{U_j} = \sum_{y_i} \mathbf{W}(y_i)\boldsymbol{X}^-(y_i)\mathbf{I}_j(y_i)$$

where $\mathbf{W}(t)$ is a $(p+1) \times (p+1)$ diagonal weight matrix . The $(\ell+1)$ element of $\mathbf{U_j}$ is the test statistic used for $H_0 : \beta_{j\ell}(t) = 0$. The covariance matrix of $\mathbf{U_j}$ is given by

$$\mathbf{V}_j = \sum_{y_i} \mathbf{W}(y_i)\boldsymbol{X}^-(y_i)\mathbf{I}_j^D(y_i)\boldsymbol{X}^-(y_i)'\mathbf{W}(y_i)$$

Different choices of the weight matrix can be considered. Aalen (1993) proposed

$$\mathbf{W}(t) = \{\text{diag}[[\boldsymbol{X}'(t)\boldsymbol{X}(t)]^{-1}]^{-1}\},$$

while Klein and Moeschberger (1997) proposed to use a weight function constant for all sub-hypotheses, such as number at risk at time t or any other constant.

Additive models are a convenient option when proportional hazards do not hold. The effect of the covariates along time can be assessed graphically by plotting the excess cumulative risk $\hat{B}(t)$ for each covariate and comparing the estimates for each competing risks and for overall survival. Though these models are very flexible and easy to implement, they are less used than Cox model because inference regarding its nonparametric terms is not fully developed and it is not included in standard statistical software. See Martinussen and Scheike (2006) for further details.

4.2 Modelling the cumulative incidence functions $F_j(t)$

The modelling of the cause-specific hazards applies when the goal is to assess if a factor is associated with the risk of a specific cause of failure. However, when the goal is to compare the observed incidence of events from a given cause between groups, the cumulative incidence functions should be used.

Estimates of these functions can be obtained via

$$\hat{F}_j(t|\mathbf{Z}) = \sum_{t_i \le t} \hat{\lambda}_j(t_i|\mathbf{Z})\hat{S}(t|\mathbf{Z}),$$

where $\hat{\lambda}_j(t|\mathbf{Z})$ are the estimated hazards resulting from Cox's or Aalen's analyses, and t_i the distinct failure times. The overall survivor function is

$$\hat{S}(t|\mathbf{Z}) = \exp\left\{-\sum_{j=1}^{k}\sum_{t_i \le t}\hat{\lambda}_j(t_i|\mathbf{Z})\right\}.$$

The problem with this approach is that no direct estimate for the effect of a covariate in the cumulative incidence function $F_j(t)$ is given. Although the effect of the covariates on the cause-specific hazard $\lambda_j(t|\mathbf{Z})$ is directly given by β_j, the effect on the cumulative incidence function $F_j(t)$ combines the effect β_j together with the overall effect on $\hat{S}(t|\mathbf{Z})$. Moreover, it is not possible to test for significant effects on the sub-distribution functions, because some covariates can have a significant effect on the hazard, but not on the $F_j's$. In order to be able to perform model selection and obtain estimates for the effects of the covariates on the cumulative incidence functions, models based directly on the sub-distribution functions have been proposed.

4.2.1 Fine and Gray's model

Fine and Gray (1999) considers a new function, the sub-hazard $\gamma_j(t)$ derived from the sub-distribution function:

$$\begin{aligned}\gamma_j(t|\mathbf{Z}) &= \lim_{\Delta t \to 0} \frac{Pr\Big(T < t + \Delta t, C = j \Big| \mathbf{Z}, \{T \ge t \text{ or } (T < t \text{ and } C \ne j)\}\Big)}{\Delta t} \\ &= \frac{f_j(t|\mathbf{Z})}{1 - F_j(t|\mathbf{Z})} \qquad j = 1, \ldots, k.\end{aligned}$$

This would be the hazard obtained from F_j if it were a proper distribution. The conditional expression includes two different scenarios: i) the event has not occurred at time t, ii) the event has occurred from a different cause before t. Thus, the risk set at time t is formed by two types of individuals, corresponding to the two different scenarios. Contrary to the analyses based on the cause-specific hazards, a patient failing from other causes would not be removed from the risk set at his/her time of failure. The sub-distribution function is expressed in terms of the sub-hazards as $F_j(t|\mathbf{x}) = 1 - \exp\left(-\int_0^t \gamma_j(t|\mathbf{x})\right)$.

Fine and Gray proposed to fit the subhazard with a Cox model, that is

$$\gamma_j(t|\mathbf{x}) = \gamma_{0j}(t)e^{\beta_j'\mathbf{x}}, \qquad j = 1, \ldots, k,$$

where the covariates are linear on a complementary log-log transformed cumulative incidence function. When censoring is absent or is always observable, Fine and Gray (1999) showed that the partial likelihood approach is valid for estimation. In the case of right-censoring, they developed a weighted score function based on the non-censored case to deal with dependent censoring. If there are N failures at the times $t_1 < t_2 < \cdots < t_N$, the partial likelihood was defined by

$$\tilde{L}(\boldsymbol{\beta}_j) = \prod_{i=1}^{N} \left(\frac{e^{\beta_j'\mathbf{Z}_i}}{\sum_{\ell \in \tilde{R}_i} w_{i\ell} e^{\beta_j'\mathbf{Z}_\ell}} \right).$$

Now the risk set for cause j at time t_i is $\tilde{R}_i = \{\ell : t_\ell \geq t_i \text{ or } (t_\ell \leq t_i \text{ and } C \neq j)\}$, where subjects experiencing a competing cause remain in the risk set. The weight $w_{i\ell}$ given to such an individual is $\tilde{G}(t_i)/\tilde{G}(\min(t_\ell, t_i))$, where $\tilde{G}$ is the survivor function for the censoring distribution. An individual satisfying $t_\ell \geq t_i$ is given a weight of 1.

4.2.2 Other models based on F_j

Other models have been proposed in line with Fine and Gray's work. Fine (2001) considered a model for the transformation of the cumulative incidence functions

$$g(F_j(t|\mathbf{Z})) = \alpha_j(t) - \boldsymbol{\beta}_j'\mathbf{Z},$$

where $g(\cdot)$ is a known differentiable function, and $\alpha_j(t)$ is the baseline failure probability when $\mathbf{Z} = 0$, which is unspecified, invertible and strictly increasing in t. Least-squares techniques are used for estimation, and inverse weighting methods are used to take into account dependent censoring. Scheike and Zhang (2005b) proposed a Cox-Aalen model for the sub-distribution hazards $\gamma_j(t|\mathbf{Z})$, and in Scheike and Zhang (2005a), they used binomial regression methods to estimate coefficients. More recently, Klein and Andersen (2005) and Klein (2006) have proposed pseudo-values regression models to approximate the sub-distribution functions.

5 An application to the Spanish Bladder Cancer Study

5.1 Competing risks data in the Spanish Bladder Cancer Study

The "Spanish Bladder Cancer Study" is a multicenter study with 1356 newly diagnosed bladder cancer cases, recruited between 1997 and 2001 in 18 Spanish hospitals and followed until March 2006. In this paper, only 994 superficial bladder cancer cases, where tumour was confined to the lining of the bladder, are considered. Recurrences of the tumour remain common among those patients, and efforts to reduce them are of paramount clinical importance.

The aim of the study is to characterize different courses of the disease. After the start of first-line therapy, transurethral resection of the tumour (TUR), patients may suffer a number of different adverse events during their follow-up: (i) recurrence, if the tumour reappears and is classified as superficial; (ii) progression, if the new tumour is classified as invasive and (iii) death, if the subject dies due to bladder cancer. About 32% of the patients experienced a recurrence, 4.9% progressions, only 1.3% died due to cancer and 13.6% died from other causes (see Figure (1)).

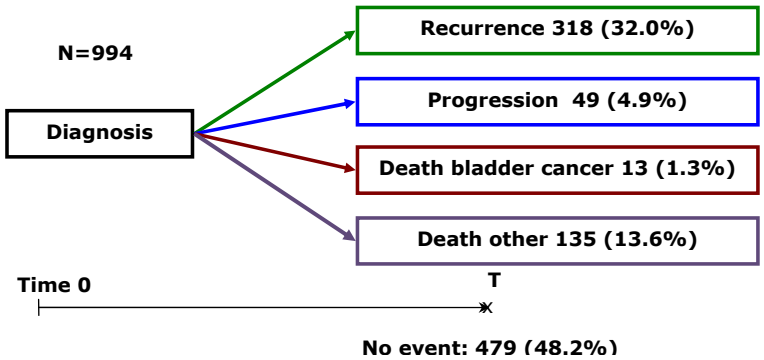

Figure 1: Competing risks structure from the Spanish Bladder Cancer Study

One important question to consider is why some patients experience a progression as a first event after diagnosis instead of a recurrence. The median time to develop a recurrence or a progression as first event is similar, 9.3 and 9.7 months, respectively suggesting that there exist distinct courses/aggressiveness of the tumour development. In order to answer this question, it is important to analyse the event-free survival time, defined as the time to the occurrence of the first event, distinguishing the different kind of events.

Let T be the time from diagnose to the first event, and C the cause of failure. In order to identify distinct patterns of the disease it is necessary to explore the joint distribution of (T, C) through a competing risks model. We will first obtain the non-parametric estimates of the cumulative incidence functions and the hazard functions specific of each cause of failure. In figures 2(a) and 2(b) we observe that both the cumulative risk and incidence of experiencing a recurrence is higher than the risk and incidence of experiencing a progression or death. Notice that the risk of dying due to other causes increases more rapidly than other causes along with time, which is logical because the risk of dying of other causes increases with age in our cohort. Both progression and death due to bladder cancer have a low cumulative hazard and cumulative incidence of occurring.

5.2 Regression modelling of competing risks

In this section, regression models are used in order to identify prognostic factors which characterize and differentiate patients who progress from those experiencing a recurrence, for example. The three methodologies revised in section 4 are applied to the bladder data. The analysis has been implemented in the freeware statistical package `R` (http://cran.r-project.org), using the libraries `survival` (Lumley and Therneau, 2003) for the Cox model, `addreg` (Fekjær, 1997) and `timereg` (Scheike and Martinussen, 2006) for the additive model, and `cmprsk` (Gray, 2004) for the subdistribution hazards approach. To illustrate such methods, we chose a subset of covariates which resulted significant in the literature available on bladder cancer.

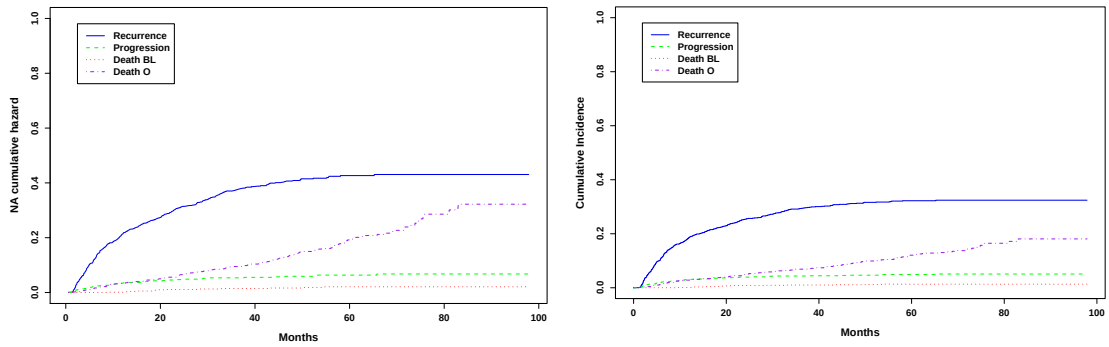

Figure 2: Non-parametric estimates of the (a) cumulative hazard function and (b) the cumulative incidence functions

Four factors are considered which may affect the risk of reappearance of the tumour: stage+grade, tumour multiplicity, Spanish region and treatment. Table 1 summarizes the distribution of these factors across the population of patients for recurrence, progression, death due to cancer, death due to other causes and total failures. For the analysis of competing risks, patients diagnosed with stage+grade Ta Benign and Tis tumours were not considered. In addition, TaGII and T1GII tumours were joined in a single category.

Table 1: Characteristics of the patients and the failure

			Cause of failure			
Factor		**Patients**	**Recurrence**	**Progression**	**Death BL**	**Death O**
Spanish region	Total	995 (100.0%)	318 (100.0%)	49 (100.0%)	13 (100.0%)	136 (100.0%)
	Barcelona	225 (22.6%)	57 (17.9%)	15 (30.6%)	2 (15.4%)	32 (23.5%)
	Vallès	161 (16.2%)	64 (20.1%)	9 (18.4%)		17 (12.5%)
	Alacant	84 (8.4%)	22 (6.9%)	3 (6.1%)		12 (8.8%)
	Tenerife	153 (15.4%)	39 (12.3%)	7 (14.3%)	3 (23.1%)	28 (20.6%)
	Asturias	372 (37.4%)	136 (42.8%)	15 (30.6%)	8 (61.5%)	47 (34.6%)
Multiplicity	Total	942 (100.0%)	302 (100.0%)	43 (100.0%)	12 (100.0%)	127 (100.0%)
	1 tumor	660 (70.1%)	186 (61.6%)	22 (51.2%)	6 (50.0%)	96 (75.6%)
	>1 tumor	282 (29.9%)	116 (38.4%)	21 (48.8%)	6 (50.0%)	31 (24.4%)
Stage+Grade	Total	995 (100.0%)	318 (100.0%)	49 (100.0%)	13 (100.0%)	136 (100.0%)
	Ta Benign	50 (5.0%)	17 (5.3%)			1 (0.7%)
	TaGI	374 (37.6%)	113 (35.5%)	9 (18.4%)	3 (23.1%)	46 (33.8%)
	TaGII	306 (30.8%)	119 (37.4%)	6 (12.2%)		41 (30.1%)
	TaGIII	98 (9.8%)	35 (11.0%)	7 (14.3%)	1 (7.7%)	14 (10.3%)
	T1GII	25 (2.5%)	7 (2.2%)	2 (4.1%)	1 (7.7%)	3 (2.2%)
	T1GIII	136 (13.7%)	24 (7.5%)	24 (49.0%)	7 (53.8%)	31 (22.8%)
	Tis	6 (0.6%)	3 (0.9%)	1 (2.0%)	1 (7.7%)	
Treatment	Total	983 (100.0%)	314 (100.0%)	47 (100.0%)	12 (100.0%)	135 (100.0%)
	RTU	401 (40.8%)	148 (47.1%)	13 (27.7%)	6 (50.0%)	64 (47.4%)
	RTU + BCG	286 (29.1%)	69 (22.0%)	24 (51.1%)	2 (16.7%)	27 (20.0%)
	TUR+Chemo	214 (21.8%)	71 (22.6%)	5 (10.6%)		32 (23.7%)
	TUR+BCG+Chemo	51 (5.2%)	21 (6.7%)	3 (6.4%)		8 (5.9%)
	Other	31 (3.2%)	5 (1.6%)	2 (4.3%)	4 (33.3%)	4 (3.0%)

Cells show Frequencies (Percentages of total cases)
Death BL: death due to bladder cancer, *Death O:* Due to other causes

5.2.1 Modelling the cause-specific hazards

Cox proportional hazards model and Aalen's additive hazards model were fitted to our data to describe cause-specific hazards. Both models involve censoring individuals failing from other causes. In Cox model, proportional hazards assumption was tested by graphical exploration of Schoenfield residuals and diagnostic tests based on them. Validation of Aalen's additive model was assessed via graphical exploration of the cumulative martingale residuals, as stated in Aalen (1993).

The parameter estimates for each model are given in Table 2. In both models, the four factors considered resulted statistically significant for recurrence, whereas only stage+grade and tumour multiplicity were so for progression. Stage+grade for death due to bladder cancer was significant in the proportional hazards model, while its significance was not clear in the additive model. Ta_T1GII tumours are associated to the risk of recurrence, while progression and death are more frequent in TaGIII and T1GIII tumours. Even if only recurrence was of interest for the researchers, it is important to model all causes of failure. Only by the results of recurrence, one may think that TaGIII and T1GIII tumours are 'protective' of recurrence, which is strange, given that these tumours are more serious. Since these two type of tumour are strongly associated to progression and death, they induce a protective effect on recurrence. Therefore, all causes of failure must be modelled in order to perform an interpretation of disease.

In Cox model, the hazard ratio $e^{\hat{\beta}_j}$ for covariate z is interpreted as the increase of hazard relative to the reference level of the z. Aalen's model provide an estimate of $\mathbf{B}_j(t) = \int_0^t \beta_j(u)du$ for a given z, where $\beta_j(t)$ is interpreted as the increase on absolute risk relative to the baseline hazard at instant t. A graphical exploration of these parameter estimates can be obtained by observing the slopes of $\hat{\mathbf{B}}_j(t)$ versus time, and how they vary. As an example, Figure 2 shows the cumulative regression functions $\hat{B}_j(t)$ for covariates multiplicity and Ta_T1GII when studying recurrence. The slope for tumour multiplicity remains constant for the first 60 months approximately, indicating that its effect on recurrence is constant over time. The slope for Ta_T1GII tumours varies over time, indicating a non-constant effect.

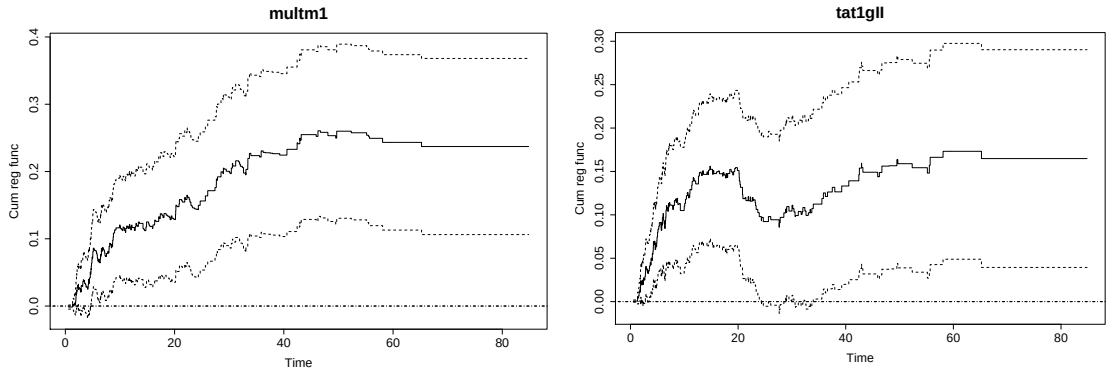

Figure 3: Cumulative regression functions B(t) for recurrence of (a) the multiplicity of the tumour, and (b) Ta_T1GII tumours.

5.2.2 Modelling the cumulative incidence functions

The results for Fine and Gray's (1999) approach are also shown in Table 2. This methodology does not need to censor individuals failing from other causes. The same factors as in the above models resulted significant, but now the subdistribution hazard ratio $e^{\hat{\beta}_j}$ for covariate z has a direct

interpretation in terms of the cumulative incidence function. The effect of T1GIII tumours over the incidence of recurrence is -0.343 while the effect of this covariate on the rate of recurrence, censoring by other causes, is -0.1764. Though similar estimates are found in our data, this situation may be dramatically different.

Table 2: Parameter estimates for the three models of competing risks.

	Cox PH for $\lambda_j(t\|Z)$			Additive for $\lambda_j(t\|Z)$			Cox PH for $\gamma_j(t\|Z)$		
Factor	$\hat{\beta}$	**SE**	**p−value**	$\hat{\beta}$	**SE**	**p−value**	$\hat{\beta}$	**SE**	**p−value**
Recurrence									
Stage+Grade Ref.level= TaGI									
Ta_T1GII	0.3987	0.135	**0.003**	0.165	0.064	**0.004**	0.408	0.133	**0.002**
TaGIII	0.1944	0.207	0.350	0.061	0.100	0.354	0.194	0.215	0.370
T1GIII	-0.1764	0.247	0.480	-0.074	0.071	0.522	-0.343	0.247	0.170
Tumour multiplicity Ref.level= 1 tumour									
>1 tumour	0.5358	0.128	**<0.001**	0.237	0.067	**0.000**	0.500	0.129	**<0.001**
Spanish Region Ref.level= Barcelona									
Valles	0.5439	0.237	**0.022**	0.202	0.086	**0.026**	0.576	0.242	**0.017**
Alicante	0.1070	0.312	0.730	0.013	0.092	0.976	0.155	0.311	0.620
Tenerife	0.0497	0.260	0.850	0.000	0.079	0.834	0.097	0.263	0.710
Asturias	0.4009	0.221	**0.007**	0.157	0.071	**0.061**	0.436	0.228	**0.055**
Treatment Ref.level= TUR									
TUR+BCG	-0.5836	0.171	**<0.001**	-0.248	0.072	**0.001**	-0.584	0.176	**0.001**
TUR+Chemo	-0.3125	0.157	**0.046**	-0.148	0.077	**0.044**	-0.292	0.154	**0.058**
TUR+BCG+Chemo	0.4139	0.316	0.190	0.157	0.152	0.332	0.421	0.319	0.190
Other	-1.1193	0.592	**0.059**	-0.338	0.101	**0.001**	-1.080	0.570	**0.058**
Progression									
Stage+Grade Ref.level= TaGI									
Ta_T1GII	-0.128	0.540	0.810	0.002	0.015	0.768	-0.231	0.537	0.670
TaGIII	1.079	0.524	**0.040**	0.115	0.073	0.129	2812	0.524	**0.048**
T1GIII	1.999	0.423	**<0.001**	0.197	0.055	**<0.001**	2.001	0.428	**<0.001**
Tumour multiplicity Ref.level= 1 tumour									
>1 tumour	0.764	0.321	**0.017**	0.044	0.028	**0.038**	0.689	0.324	**0.034**
Death due to bladder cancer									
Stage+Grade Ref.level= TaGI									
TaGI	Ref			Ref			Ref		
Ta_T1GII	-0.872	-0.755	0.450	-0.008	0.008	0.417	-0.987	1.153	0.390
TaGIII	0.355	0.308	0.760	0.013	0.026	0.768	0.213	1.150	0.850
T1GIII	1.784	2.442	**0.015**	0.050	0.029	**0.068**	1.678	0.724	**0.020**

Acknowledgements

We specially thank Àlex Amorós, since this work is inspired in his master thesis project entitled 'Supervivencia en presencia de eventos compitiendo', presented on June 2006 at the Faculty of Mathematics and Statistics. We thank Àlex and also Cristiane Murtra from IMIM, Barcelona, Spain, for their previous work on the Spanish Bladder Cancer Data, as well as for the data management of the study. This work was partially supported by Grant 050831 from La Marató de TV3 Foundation and by grant MTM2005-0886 from the Ministerio de Ciencia y Tecnología. Núria Porta is a recipient of a doctoral research fellowship from the Catalan Ministry of Innovation, Universities and Enterprise.

A Appendix

A.1 Members of the participating centres

Institut Municipal d'Investigació Mèdica, Universitat Pompeu Fabra (Barcelona) (coordinating center): M. Kogevinas, N. Malats, F.X. Real, M. Sala, G. Castaño, M. Torà, D. Puente, C. Villanueva, C. Murta-Nascimento, J. Fortuny, E. López, S. Hernández, R. Jaramillo, F Fernandez, A. Amorós, G Vellalta, L Palencia, A Alfaro, G Carretero. Hospital del Mar, Universitat Autònoma de Barcelona (Barcelona): J. Lloreta, S. Serrano, L. Ferrer, A. Gelabert, J. Carles, O. Bielsa, K. Villadiego. Hospital Germans Trias i Pujol (Badalona, Barcelona): L. Cecchini, J.M. Saladié, L. Ibarz. Hospital de Sant Boi (Sant Boi, Barcelona): M. Céspedes. Centre Hospitalari Parc Taulí (Sabadell, Barcelona): C. Serra, D. García, J. Pujadas, R. Hernando, A. Cabezuelo, C. Abad, A. Prera, J. Prat. Centre Hospitalari i Cardiològic (Manresa, Barcelona): M. Domènech, J. Badal, J. Malet. Hospital Universitario (La Laguna, Tenerife): R. García-Closas, J. Rodríguez de Vera, A.I. Martín. Hospital La Candelaria (Santa Cruz, Tenerife): J. Taño, F. Cáceres. Hospital General Universitario de Elche, Universidad Miguel Hernández (Elche, Alicante): A. Carrato, F. García-López, M. Ull, A. Teruel, E. Andrada, A. Bustos, A. Castillejo, J.L. Soto. Universidad de Oviedo (Oviedo, Asturias): A. Tardón. Hospital San Agustín (Avilés, Asturias): J.L. Guate, J.M. Lanzas, J. Velasco. Hospital Central Covadonga (Oviedo, Asturias): J.M. Fernández, J.J. Rodríguez, A. Herrero. Hospital Central General (Oviedo, Asturias): R. Abascal, C. Manzano, T. Miralles. Hospital de Cabueñes (Gijón, Asturias): M. Rivas, M. Arguelles. Hospital de Jove (Gijón, Asturias): M. Díaz, J. Sánchez, O. González. Hospital de Cruz Roja (Gijón, Asturias): A. Mateos, V. Frade. Hospital Alvarez-Buylla (Mieres, Asturias): P. Muntañola, C. Pravia. Hospital Jarrio (Coaña, Asturias): A.M. Huescar, F. Huergo. Hospital Carmen y Severo Ochoa (Cangas, Asturias): J. Mosquera.

References

Aalen, O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models. *The Annals of Statistics*, **6**(3), 534–545.

Aalen, O. O. (1993). Further results on the non-parametric linear regression model in survival analysis. *Stat Med*, **12**(17), 1569–1588.

Andersen, P. K., Abildstrom, S. Z., and Rosthøj, S. (2002). Competing risks as a multi-state model. *Stat Methods Med Res*, **11**(2), 203–215.

Cox, D. and Oakes, D. (1984). *Analysis of survival data*. London: Chapman-Hall.

Escrig, V. J. (2005). [many published survival analyses may be intrinsically flawed. the cases of transplantation and cancer]. *Cir Esp*, **77**(5), 297–298.

Fekjær, H. (1997). *Documentation for S-Plus package ADDREG in additive survival analysis*. Section of Medical Statistics, University of Oslo, Norway. http://www.med.uio.no/imb/stat/addreg/index.html.

Fine, J. and Gray, R. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, **94**(446), 496–509.

Fine, J. P. (2001). Regression modeling of competing crude failure probabilities. *Biostatistics*, **2**(1), 85–97.

Gooley, T. A., Leisenring, W., Crowley, J., and Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*, **18**(6), 695–706.

Gray, R. (2004). *The cmprsk package*. The Comprehensive R Archive network. http://cran.r-project.org/src/contrib/Descriptions/cmprsk.html.

Kalbfleisch, J. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Klein, J. and Moeschberger, M. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag New York.

Klein, J. P. (2006). Modelling competing risks in cancer studies. *Stat Med*, **25**(6), 1015–1034.

Klein, J. P. and Andersen, P. K. (2005). Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics*, **61**(1), 223–229.

Lawless, J. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Lin, D. Y. (1997). Non-parametric inference for cumulative incidence functions in competing risks studies. *Stat Med*, **16**(8), 901–910.

Llorca, J. and Delgado-Rodríguez, M. (2004). [survival analysis with competing risks: estimating failure probability.]. *Gac Sanit*, **18**(5), 391–397.

Lumley, T. and Therneau, T. (2003). *The survival package*. The Comprehensive R Archive network. http://cran.r-project.org/src/contrib/Descriptions/survival.html.

Martinussen, T. and Scheike, T. H. (2006). *Dynamic regression models for survival data*. Statistics for Biology and Health. Springer, New York.

Pepe, M. S. and Mori, M. (1993). Kaplan-meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Stat Med*, **12**(8), 737–751.

Peterson, A. V. (1976). Bounds for a joint distribution function with fixed sub-distributions functions: Applications to competing risks. *Proceedings of the National Academy of Sciences of the USA*, **73**, 11–13.

Pintilie, M. (2006). *Competing risks: a practical perspective*. Wiley.

Pintilie, M. (2007). Analysing and interpreting competing risk data. *Stat Med*, **26**(6), 1360–1367.

Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., J., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, **34**(4), 541–554.

Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*, **26**(11), 2389–2430.

Scheike, T. and Martinussen, T. (2006). *Semi-parametric timevarying regression for R. Survival regression software*. http://staff.pubhealth.ku.dk/t̃s/timereg.html.

Scheike, T. and Zhang, M. (2005a). Predicting cumulative incidence probability by direct binomial regression. Technical report, Institute of Public Health, University of Copenhagen.

Scheike, T. and Zhang, M. (2005b). Predicting cumulative incidence probability: Marginal and cause-specific modelling. Technical report, Institute of Public Health, University of Copenhagen.

Tai, B. C., Machin, D., White, I., Gebski, V., and behlf of the EOI (The European Osteosarcoma Intergroup), O. (2001). Competing risks analysis of patients with osteosarcoma: a comparison of four different approaches. *Stat Med*, **20**(5), 661–684.

Therneau, T. and Grambsch, P. (2000). *Modeling Survival Data: Extending the Cox Model*. Statistics for Biology and Health. Springer.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc Natl Acad Sci U S A*, **72**(1), 20–22.