Non-parametric estimation of the marginal distribution of time to specific events in the presence of competing risks

Ilaria Ardoino[1], Elia Biganzoli[1,2], Patrizia Boracchi[2]

[1] Unit of Medical Statistics and Biometry, IRCCS National Cancer Institute, Milan, Italy
[2] "G.A. Maccacaro" Institute of Medical Statistics and Biometry, University of Milan, Italy

Corresponding Author:
Ilaria Ardoino, Unit of Medical Statistics and Biometry, Fondazione IRCCS Istituto Nazionale Tumori
Campus Cascina Rosa - Via Venezian 1, 20133 Milano, Italy
E-mail: ilaria.ardoino@istitutotumori.mi.it

Summary
In clinical studies multiple failure times are often of interest. Different events can be thought as "competing" with each other in inducing the first failure (*competing risks*). It may be that the outcome of interest is the time of occurrence of a specific, clinically important, event; thus the occurrence of other fatal and non-fatal events, precluding its observation, acts as a "nuisance".
In practice it is almost never possible to observe the whole event history, and nor it is possible to observe each single event in isolation. The occurrence of a specific event in a follow-up setting cannot be reduced to a standard survival context, without a strong and unrealistic working assumption of independence between events. The marginal survival distribution cannot be estimated using a standard technique.
A strategy based on a possible generalisation of the best known non-parametric estimators, accounting for the simultaneous action of competing events and independent censoring, is proposed. A recent approach based on the concept of semi-competing risks (Fine et al., 2001) is introduced. An application of the proposed methodology to a real dataset of 2233 breast cancer patients submitted to conservative surgery is also shown, focusing on the estimate of distant metastasis disease-free survival.

KEY WORDS: *breast cancer, competing risks, marginal distribution, non-parametric estimator, survival analysis*.

Introduction

Situations involving multiple correlated failure times frequently arise in medical research. The course of chronic diseases may, in fact, be characterised by the possible occurrence of several non-fatal events prior to death, landmarks in the disease dynamic. The first event a patient experiences, regardless of whether it is fatal or not, is often of special interest, as it may be considered the first evidence of "treatment failure". Different kinds of event generally demand different therapeutic strategies and it is important to estimate the probability of their occurrence. In this situation, characterised by the presence of events competing with each other in inducing failure, reference is made to *competing risks*.
As a simple example, competing risks arise in analyses of mortality in a given population, when attention is focused on distinguishing between different causes of death (e.g. cardiovascular disease, tumour-related, or others). Generally, the causes of failure might be thought as mutually exclusive, as they compete with each other to be observed.
When aiming to establish how the incidence of each event contributes to determining mortality or the failure pattern, interest focuses on the occurrence of a specific event as the first event, and here the quantity of interest is the crude cumulative incidence (CCI).

However, it may be that different events are not of equal importance: were the aim to evaluate therapeutic strategies targeted at preventing the occurrence of a specific event, then the time to that event would be relevant, in the absence of the confounding (nuisance) influence of other events, whether or not it was the first event occurring. In this case, *specific event survival* is, the quantity of interest.

In the presence of competing risks, it is usual to interpret marginal distribution as net survival, as in a hypothetical setting in which other causes of failure, or some of them, are no longer acting (2). However, the interpretation of marginal distribution is still debated in situations in which it is meaningful to refer to events potentially occurring as second or even as subsequent events. In this case, it would be preferable to attempt to study failure time to a specific event, without ascribing it any special "physical" meaning.

An example may be useful in order to illustrate the difference between a classic competing risk approach based on observable quantity (e.g. the CCI) and those based on estimation of marginal survival distribution. When studying the progression of HIV infection, primary AIDS diagnoses and subsequent deaths are of primary concern. If the interest is in the incidence of new AIDS cases in a population, for the purpose of allocating resources for treatments and hospitalisation, AIDS diagnoses are the primary endpoint. The focus of interest is, therefore, the subjects currently under observation, and the CCI is the appropriate quantity to investigate. The situation differs with regard to the distribution of the clinical incubation period, which extends from seroconversion to the presentation of different categories of AIDS diagnosis (e.g. Kaposi sarcoma, Pneumocystiis carinii); HIV-infected individuals may present secondary and additional AIDS-defining conditions. In this context, given the different therapeutic approaches, it would be relevant to assess the marginal probabilities of each event. When aiming to study the temporal trend according to different conditions, the marginal incidence and the corresponding hazard function would be the reliable quantity of interest; see (4).

Even when competing risks can be supposed to be independent, CCI and marginal, rather than net survival, may respond functionally to different aims in surveys.

Most importantly the marginal distribution of a non-fatal event is not estimable non-parametrically on the basis of competing risks data – the "non-identifiability aspect", (5) – nor on the basis of a partially observed event history. Unless the unverifiable assumption that the risks are independent is accepted, proxy knowledge of the joint survival distribution or at least of the association among events is required. In the biomedical setting, physiological and biochemical conditions make independence among failure times in the same subject, in most situations, unrealistic. Nonetheless, in the clinical literature the Kaplan-Meier method, censoring the events other than the event of interest, thus presuming "censoring processes" to be independent and non informative, is still commonly used. This procedure generally results in a biased estimate of the marginal survival curve: the bias is not consistent in its sign and increases as censoring due to competing events increases. Classical approaches based on the use of estimable quantities, such as the CCI, are thus preferred. The above procedure is often induced by the lack of a correct framing of the statistical methods in response to the clinical problem. Historical approaches identify bounds, i.e. an upper and a lower estimate based on estimable quantities, in the attempt to delimit the portion of the plane in which the marginal curve would lie. To date, no exhaustive work has been done and the results, from a practical point of view, are unsatisfactory. Approaches that have been proposed (6-8) deal with mutually exclusive events, as though censoring due to competing events were the only process acting, thereby disregarding independent censoring, i.e. administrative conclusions and losses to follow up. Throughout common practice, it would be naive to think that dependent and independent censoring processes do not act jointly. If a more general setting is considered, such as that of longitudinal follow-up data, it is clear that patients may experience more than one type of failure before death. Koscielny and Themes (14) illustrate the usefulness of additional information in improving estimates of bounding.

The aim of this work is to present a non-parametric procedure to estimate the marginal survival distribution in the presence of both correlated events and independent censoring. In the context of a critical review of the possible interpretation and usefulness of

the marginal survival distribution (6,7), a possible methodology is proposed, as a simple generalisation of the said proposals for bounding net survival. In addition, our approach is adjusted with suggestions on how to estimate dependence between the events of interest.

Finally, semi-competing risks – a variation of competing risks referring to a situation in which a terminal event may censor a non-fatal one, but not vice versa – together with an estimation procedure, which accounts explicitly for the event history information when available, are briefly introduced. This setting makes it possible to accommodate both mortality and informative dropout: the data are often a mixture of the two scenarios. Semi-competing risks analysis constitutes, in our view, a valid and proper approach in situations in which results based on classical competing risks may be misleading and, in general, unsatisfactory.

Motivation

The present work, like others by the same authors, was conducted in order to investigate treatment effect in early breast cancer patients.

Although surgical and radiation therapies have steadily improved survival, breast cancer is still one of leading causes of cancer-related mortality for women in the western population. Heterogeneous outcomes are observed in the prognoses of many patients.

Breast conservative surgery, which may or may not be followed by adjuvant therapy, has become the preferred treatment for early breast cancer, aiming to prevent further dissemination of the neoplasm. It is often preferred to more aggressive treatment, given that the observed differences in overall survival are non significant (9). Conservative surgery, however, may be an inadequate local treatment, exposing the patient to a higher risk of ipsilateral tumours and local relapse, commonly grouped under the heading "intra-breast tumour recurrences" (IBTRs). IBTRs, on their own, are not considered severe neoplastic events, rather an added source of discomfort for the patient, who may require further mastectomy. Distant metastases (DMs) involving other anatomical sites are, instead, serious life-threatening events; they may appear in vital organs, and be highly predictive of death. Patients may also experience, during follow up, a second contralateral carcinoma (CC) or another non-mammary primary tumour (OP), prior to death. The breast cancer dynamic following conservative surgery is illustrated in Figure 1.

Although, in clinical trials, study protocols provide accurate information and events are recorded for each subject, it is often preferred, as an established practice to reduce such data in a competing risks framework, ignoring the occurrence of non-fatal and fatal events.

It is likely that the occurrence of one event may alter the risk of subsequent ones and we therefore feel that when each patient's event history is recorded all the

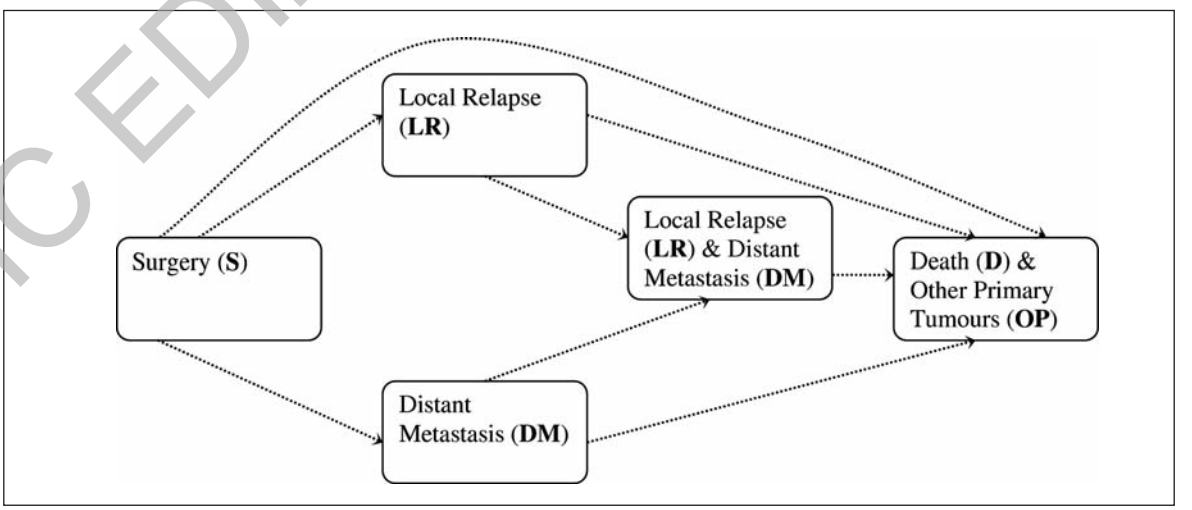

Figure 1. Breast cancer dynamic following conservative surgery.

available information on the follow up should be taken into account.

In the next section, competing risks and multiple failure times data are introduced under a unified framework. We enter main quantities of interest: their interpretation and estimation procedure. We focus on the marginal survival distribution; available approaches are reviewed, together with their current limits. A suitable methodology is then proposed. Finally, an application to a real dataset of breast cancer patients is shown. The article ends with a general conclusion and a few thoughts on future developments.

Materials and Methods

Basic notation and methodological background

Since a subject is believed to be at risk of developing each single event at a certain time point, it is usual to have recourse to a multivariate representation. Let us consider $\mathbf{T} = (T_1,...T_r,...T_R)$ to be a vector of failure times, with multivariate survival distribution $S_T(\mathbf{t}) = P(T_1 > t_1;...T_r > t_r;...T_R > t_R)$ This vector refers to a subset of non fatal, and generally unordered, events, and a final, terminal one. The treatment failure, the first failure to occur, can be described through a couple of random variables, r.v., $(T_f;\rho)$, where $T_f = \min(T_1,...T_r,...T_R)$ is the failure time and $\rho = \arg(T_f)$ designates the failure event type.

Define the net or marginal survival distribution as:

$$S_{T_r}(t) = P(T_r > t) \qquad [1]$$

In accordance with the motivation, i.e. to evaluate the true benefits and safety profile of conservative surgery for early breast cancer patients, the occurrence of DM is the main endpoint and the time to this event may be referred to as T_1. The occurrence of a local failure (IBTR), at time T_2, may well precede the occurrence of a DM without precluding its detection. Death may well occur at any time and this time will be T_3. In view of the objectives of the present work, the marginal distribution of DM is the most appropriate quantity to investigate.

Other functions are of interest in a competing risks setting. Define the CCI function as:

$$F_r(t) = P\left\{T_r \le t; \bigcap_{k \ne r} T_k > T_r\right\} = P\left(T_f \le t; \rho = r\right) \qquad [2]$$

which represents the probability of failure by time t, due to the specific event of interest, i.e. DM in any site, when it is the first event observed and the other types of event are also considered.

Many authors have attempted to render CCI as a "marginal" probability: note that it is not a true "marginal" in the sense given earlier, and it is likely to be biased when the marginal distribution is the issue of interest. It is interesting that, in estimating the function [2], no special assumption is required on the informative "nature" of the relationship between events.

Define, finally, the "treatment failure" time distribution, also named disease-free survival, as

$$S_{T_f}(t) = P(T_1 > t;...T_r > t;...T_R > t)$$

i.e. the probability of being alive at time t without any adverse event.

In the following, and without loss of generality, we consider only two correlated failure times T_r and T_R, the latter referring to the occurrence of a "terminal" event. In the attempt to refer to a more comprehensive pattern, a right censoring time, T_C, independent of both of them, is introduced. We refer explicitly to a univariate and equal censoring time, consistently with the setting introduced, where it would appear difficult to contemplate a multivariate vector of censoring times. In general, the last observation time corresponds to the minimum between the time to the "terminal" event, if observed, and T_C, $T = \min(T_R,T_C)$; $\delta = I(T_R < T_R)$; would provide the "death" indicator. Given a sample of N subjects, then, the observed data consists of N replicates of $\{(T_{ri},T_i,\delta_i,\eta_i); i = 1,2,\cdots,N\}$, where the additional $\eta_i = I(T_{ri} < T)$ is an event indicator for subject i.

Let $0 < t_1 < t_2 < \cdots < t_j < t \cdots < t_R$ be the ordered failure times for any type, observed in N subjects; denote, also, by d_{rj} the number of subjects failed for the event of interest in time t_j and by n_j the number of subjects at risk just before the same time point.

A consistent estimator of the CCI in [2] can be obtained by:

$$\hat{F}_r(t) = \sum_{j|t_j \le t} \frac{d_{rj}}{n_j} \cdot \prod_{j|t_j \le (t-1)} \left(1 - \frac{d_j}{n_j}\right) \qquad [3]$$

regardless of the assumption of independence among

failure types. Here, the specific risk of failure for the event *r* is properly "adjusted" for the probability of being at risk at each time point: $\prod_{j|t_j \le (t-1)} \left(1 - \frac{d_j}{n_j}\right)$ provides the Kaplan-Meier estimate for disease-free survival $\hat{S}_{T_f}(t-1)$. In the absence of censoring, the proposed estimator comes down simply to the proportion of those, among all the subjects, who have failed for the event of interest.

A "naïve", widely used estimator of marginal survival distribution is the Kaplan-Meier estimator for the event of interest:

$$^{B}\hat{S}_{T_r}(t) = \prod_{j|t_j \le t} \left(1 - \frac{d_{rj}}{n_j}\right) \qquad [4]$$

In [4], the usual convention is adopted whereby failures due to competing events, as well as censored observations, are removed from the set of patients at risk, in such a way that both are supposed to have the same risk of undergoing the event as those alive and under observation. Under the assumption of independence between failure times, [4] becomes an asymptotically unbiased estimator (18, 20).

The assumption of independence is completely untestable on the basis of competing risks data and rather questionable, at least in biomedical applications. The same non-identifiability aspect is transposed to the multivariate setting. Independence can be justified only on a priori subjective grounds, or as a convenient singling out of a distribution that can give rise to $(T_f;\rho)$.

Peterson in (10) identified natural empirical bounds for the marginal survival function, delimiting a region in the plane that contains the estimate of the "true" marginal distribution:

$$S^{*}_{T_f}(t) \le \hat{S}_{T_r}(t) \le S_r(t) + [1 - \pi_r] \qquad [5]$$

where $\{S_r(t) + [1 - \pi_r]\} = S^{*}_r(t)$ is really just the 1–CCI, in [2]. Obviously $S^{*}_r(t)$ is not a proper survival function. These empirical bounds, although easily computable, are not very useful in general, because they are quite broad; the stronger the action of the competing events, the wider they become. In general, they are useful for showing how one might be misled, at worst, into erroneously assuming independence. The results in [5] can easily be proved via heuristic considerations, and their meaning can be intuited as corresponding to the extreme dependence pattern (11).

Estimating the marginal survival distribution

To estimate the marginal survival distribution in competing risks settings, the main proposals reported in the literature have focused on improving the bounds proposed by Peterson, identifying a proper dependence measure, as well as an association structure, like a copula, among events, which makes the marginal survival distribution identifiable. However, these proposals seem to be of little pertinence to many researchers, and only a few remarks are found in the clinical and methodological literature. It is, in any case, a fact that bounding, on its own, is poorly informative.

In the context of these main proposals, the authors assumed that they would observe either of two dependent and mutually exclusive events, in accordance with the aforementioned notation T_r and T_R: in this context (studying breast cancer dynamic following conservative surgery), these refer to the time of occurrence of DM and of death, respectively. In this situation, Slud and Rubinstein, by defining a time dependence measure between failure times, achieved a "consistent" generalised Kaplan-Meier estimator for the marginal survival function. This dependence measure was given by the ratio of the conditional hazard of failure for the event of interest *r* according to whether an individual had previously been "censored" due to the occurrence of the competing event *R*, or was free from any events:

$$\rho(t) = \lim_{\Delta t \to 0} \frac{P\left(t < T_r \le t + \Delta t \middle| T_r > t; T_R \le t\right)}{P\left(t < T_r \le t + \Delta t \middle| T_r > t; T_R > t\right)} \qquad [6]$$

Through recourse to an alternative analytical formulation and a little mathematics, the marginal survival distribution results in a closed-form equation. Slud and Rubinstein give for their empirical estimator, in the absence of independent censoring, two equivalent formulations. Based on empirical considerations their estimator has a quite simple interpretation. The estimate of the marginal survival distribution may be obtained by adding to the overall survival distribution (Peterson's lower bound) an amount corresponding to the increasing potential contribution of the previously dependent censored observation, surviving beyond time *t*, and estimated by the means of the authors' non-parametric assumption. When independence is assumed, the Slud and Rubinstein estimator reverts to a usual Kaplan-Meier estimator.

Klein and Moeschberger proposed an alternative approach beginning from slightly different information. They claim to model the dependence between the outcomes of interest – in our example, DM and death – through the Clayton Oakes copula, defined by:

$$S_T(t_r, t_R) = \left\{ \left[\frac{1}{S_{T_r}(t_r)} \right]^{\theta-1} + \left[\frac{1}{S_{T_R}(t_R)} \right]^{\theta-1} - 1 \right\}^{-1/(\theta-1)} \quad [7]$$

where the parameter $\theta > 1$ allows for non-negative dependence between failure times. The gamma frailty model is often chosen for its computational tractability, and because the dependence parameter is intuitively and clinically meaningful and helpful in medical decision making. Parameter θ is simply interpretable as a predictive hazard ratio, which is constant over time and represents the conditional ratio between the hazard function of failure at time t_r, given that the individual is censored at time t_R (i.e. $P(t_r < T_r \leq t_r + \Delta t | T_r > t_r; T_R = t_R)$), or later than t_R (i.e. $P(t_r < T_r \leq t_r + \Delta t | T_r > t_r; T_R \geq t_R)$): this implies that the hazard rate for the event of interest, after competing events occur, is accelerated by a factor θ. When $\theta = 1$, independence is reached.

In general, $\theta = (1 + \tau_K) / (1 - \tau_K)$; so that Kendall's tau coefficient is: $\tau_K = (\theta - 1) / (\theta + 1)$.

Now, if the CCI, in [2], for the event of interest is obtained – this is done by differentiating it – and if associated differential equation for $S_{T_r}(t)$ is solved, a natural expression for the marginal survival distribution is given. For $\theta = 1$, when failure times are independent, the estimator becomes a cumulative hazard rate estimator, as proposed by Nelson Aelen for randomly censored data. More details are given in the appendix.

Proposed Methodology

In the approaches reviewed, only dependent censoring due to the action of competing events is explicitly considered; however, clinical studies are often characterised by long follow-up times and by large proportions of patients alive at the closing date, or of losses to follow up. A rough solution has been presented, in which administrative censoring is dealt with in the same way as dependent censoring. However, these two situations show important differences. An example may be helpful to illustrate these. In studying cancer-specific mortality among patients with a smoking history there emerges a higher mortality from cardiovascular diseases than among non-smoking ones. In estimating marginal cancer-specific survival, however, if no stratification with respect to smoking attitude is applied, the censoring pattern due to mortality for cardiovascular disease may well be different between smoking and non-smoking patients and informative censoring is introduced, thus the survival curves obtained will produce a biased estimate. In a population in which competing risks are not operating, the "potential" failure time would be observed, and thus is not entirely hypothetical.

Considering the original expression of both the estimators previously introduced, all the quantities involved are consistently estimable in the presence of independent censoring. The plug-in of the corresponding estimators, allows, in the simplest way, to account for independent censoring.

The second problem, encountered in efforts to render the proposed modified estimator more appealing, is related to estimating association. A recommended solution, perhaps the only feasible one, in a "classical" competing risks setting is to assign two possible values for the association parameter, in order to draw estimation bounds based on a priori knowledge of the experimental setting. It has to be noted, however, that clinicians – the researchers most involved – are usually unfamiliar with association measures, and it is therefore quite hard for them to give a subjective judgment.

In the Klein and Moeschberger estimator, parameter θ can be estimated resorting to its relation with Kendall's tau. In a longitudinal data framework, it is possible to take into consideration partial information on the event history. Complete information is available when two events are both observed (when DM occurred before death), or when only the terminal event is censored, otherwise the information available is only partial (incomparable pairs). This makes it necessary to resort to the proposed Kendall's tau estimator, in the presence of censoring. A brief review and a recent proposal can be found in (12). One of the best performing estimators was the one proposed by Brown et al. (1979) and also presented in (12): these authors proposed assigning, only to incomparable pairs, a proper score which utilis-

es the Kaplan-Meier estimates of marginal survival distribution.

In the Slud and Rubinstein formula, the question of dependence might be addressed using the "hazard ratio": a proportional hazard model could be fitted using a time-dependent covariate for the time of occurrence of the competing event; however, this can be rather difficult for practitioners to interpret. A valid alternative and possible solution is to resort to a "self-consistent procedure". The idea is to identify a constant value in time as the initial estimate for $\rho(\cdot)$, and to obtain a proximal estimate for $\hat{S}_{T_r}^{\rho}(t)$, subsequently reiterating the process, using the formula [9] until the estimated marginal survival function reaches convergence.

Semi-competing risks

In this paragraph, we look at some early and important results of the analysis of *semi-competing risks*, without, however, extending our discussion to details of more recent, upcoming research.

A straightforward example is that of breast cancer dynamics. Intrinsic ordering is imposed for events like DMs in vital organs and death. DMs occurring secondarily to other tumours are of little interest; they should be disregarded as an events of interest and the observation should be stopped upon the occurrence of OPs.

The authors assume a dependence structure, the same gamma frailty model shown in [7], on the upper wedge, $T_r \leq T_R$, where data are observable, leaving the model on the lower wedge completely unspecified. In the new framework, not only the overall survival function $S_{T_f}(t)$, but also the marginal survival distribution of the terminal event $S_{T_R}(t_R)$ can be correctly estimated on the basis of observed data, as it is subjected only to independent censoring.

Resorting to the overall survival relation, marginal survival distribution of the non-terminal event may simply be obtained by difference.

It comes out as:

$$\hat{S}_{T_r}(t) = \left\{ \left[\hat{S}_{T_f}(t) \right]^{1-\hat{\theta}} - \left[\hat{S}_{T_R}(t) \right]^{1-\hat{\theta}} - 1 \right\}^{1/(\hat{\theta}-1)} \quad [8]$$

It may well be that the proposed estimator results in a non-monotone function, notably on the tail, due to the action of independent censoring. It was modified in order to accept $\hat{S}_{T_r}(s)$ if this satisfies the monotonic constraint, otherwise it carries forward the smallest value of $\hat{S}_{T_r}(s)$ for each interval in which $s \leq t$.

Here, the parameter θ no longer shows the same relation with Kendall's tau and this forced the authors to define a closed-form estimator for it (1). In accordance with the spirit of the model, they propose estimating it like a concordance estimator, relying only on comparable pairs; potentially it can be adjusted by means of an appropriate weight function, so that censoring observations can contribute and also so as not to attribute excessive weight to pairs whose risk set is too small. A good choice may be to weight each pair exactly by the inverse of its risk set.

The authors give the expression for the variance estimate of the association parameter, as well as the asymptotic properties of the survival function estimator.

A real case study example

The dataset

The present study includes 2233 consecutive early breast cancer patients hospitalised at the National Cancer Institute in Milan between October 1970 and December 1987 and treated with conservative surgery. All patients underwent quadrantectomy – removal of a substantial portion of the quadrant containing the tumour by means of a radial incision that includes a portion of the skin – with axillary lymph node dissection followed by radiotherapy (QUART). For a detailed description of the dataset, including the demographic, clinical and pathological characteristics of the patients, see (13). The patients were followed up regularly for about 12/13 years, neoplastic events (date and type) were recorded, as was date of death. A median follow-up time of about 8.5 years was calculated.

In all, 744 "first" neoplastic events were observed: 151 IBTRs, 414 DMs, 179 other primary tumours (this category including CCs and OPs). A total of 39 patients died without any evidence of breast cancer recurrence. Also 73 "second" events were recorded: 57 DMs (following an IBTR) and 16 OPs. Finally, 384 patients died following one or more adverse events.

The main objective of the present analysis was to monitor the possible occurrence of DMs in any site. In order to evaluate more efficiently the marginal survival distribution, we decided to follow the "commonsense approach" suggested in (14), and to ignore the occurrence of an IBTR when it preceded the event of interest, as an IBTR does not preclude the observation and evaluation of subsequent DMs. This approach makes it possible to account for an increased number of events and to investigate the dependence measure properly. Many authors have criticised this approach, for a discussion see (15, 16), whose authors discussed, in fact, two sources of potential bias:

• Administration of additional therapy;

• Reduced vigilance towards the event of interest.

In our opinion, the second aspect, in particular, does not correspond to usual clinical practice. A local recurrence may be a sign of a tumour's aggressiveness and thus of its capacity to recur in any site. Accordingly, a patient experiencing an IBTR, being supposed to be at higher risk of further adverse events, would rarely receive less attention in their follow up. Competing events were defined by the occurrence of death as well as the occurrence of a second CC or of an OP. These events are, on their own, of little interest; they imply informative dropout, making subsequent DMs no longer interpretable as treatment failures. The long follow-up data available make it possible to calculate the semi-competing risks estimator.

Results

Competing risks analysis

The disease-free survival estimate, considering the first event (among DM, CC, OP and death), provides a suitable estimate of Peterson's lower bound, while 1– CCI provides Peterson's upper bound, see Fig. 2. These bounds currently correspond to the extremes, minimal and maximal, of the dependence structure between failure times: if the first event occurring implies that other events will follow very shortly after, the lower bound is reached; on the contrary, if the first event means that the others will never occur, then the upper bound is reached. It is thus possible to see how one might be misled, at worst, into erroneously assuming independence. As expected, original bounding on estimable quantities provides very weak information in relation to the true marginal survival distribution. From this example, even though the event of interest has a dominant incidence (471 vs 228) over the other events, making the CCI for the event of concern much lower than the overall incidence, the bounding emerges as quite wide. It has to be pointed out that these bounds do not depend on the sample size: they are not confidence bounds. Each curve falling into the bounds is, in some sense, a good estimate of the marginal distribution, as the biased Kaplan-Meier is, and would have its own confidence limit, and the union of those confidence limits would produce even wider bounds.

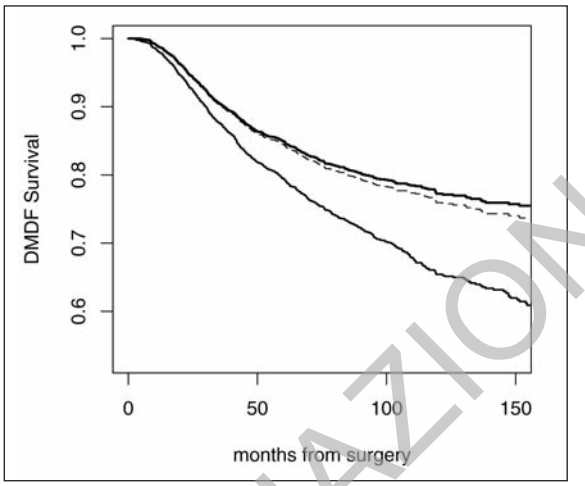

Figure 2. Peterson bounding for distant metastases disease-free survival - Solid lines -; dashed line: Kaplan-Meier biased estimate.

The available information on the partially observed event history allows us to gain a proxy knowledge of the dependence between the failure times. The estimation procedure for Kendall's tau under censoring, between time to DM and time to death, provides a rough value $\hat{\tau}_K = 0.25$. Thus, in estimating marginal survival distribution using the estimator proposed in [15], we use the corresponding value of the predictive hazard ratio for the Clayton Oakes model, which was $\hat{\theta} = 0.6685$. The same value was used as the first estimate for the dynamic dependence measure defined by Slud and Rubinstein, for the modified estimator in [12]. Both estimates for marginal survival distribution, shown in Fig. 3, overlap the biased Ka-

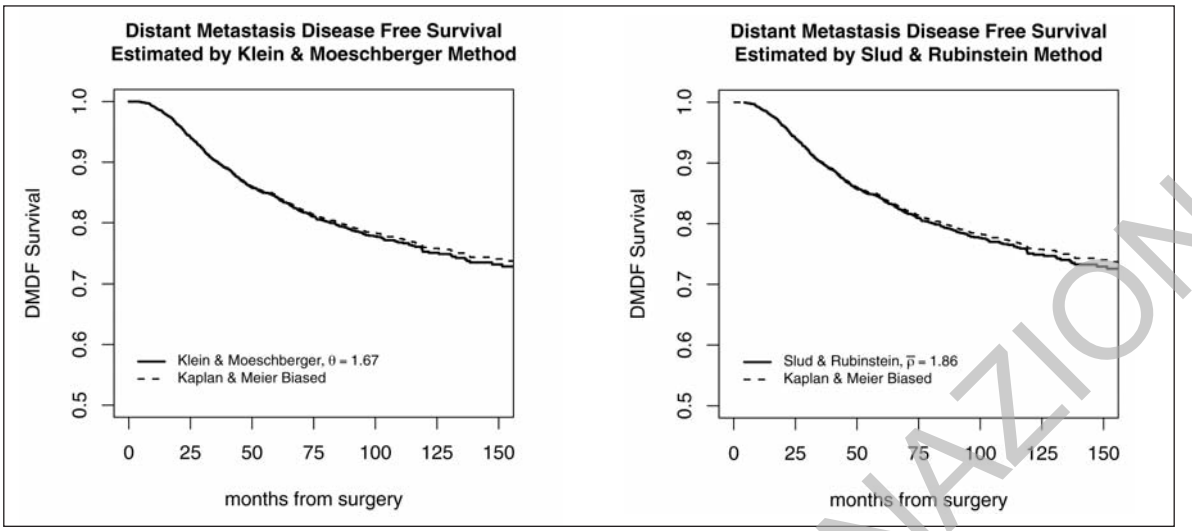

Figure 3. Estimates of distant metastasis disease-free survival. Klein-Moeschberger (left panel), Slud-Rubinstein (right panel) in comparison with the "naive" Kaplan-Meier estimate

plan-Meier: after 12 years of follow up the estimates result in respectively ${}^{KM}\hat{S}_{T_{DM}}(t) = 0.7344$ and ${}^{B}\hat{S}_{T_{DM}}(t) = 0.7430$, suggesting very weak association between these event times. This may indicate that using the Kaplan-Meier biased estimate produces a small bias in this case.

However, it should be noted that in this study the Kendall's tau coefficient is not a proper dependence measure because of the constraint $T_{DM} \leq T_{Other}$ in defining failure times; more suitable concordance indexes for the association measure would need to be considered.

Semi-competing risks analysis

In attempting to account explicitly for the information available on the event history and the natural ordering of the events, instead of relying only on the first failure, this application makes it possible to construct the semi-competing risks estimates for the marginal survival distribution of DM.

The weighted concordance index results in $\hat{\theta}_W = 13.5$; also the unweighted concordance index was calculated, and gave a quite similar value of $\hat{\theta}_{UW} = 13.5$: this is not surprising given the large number of patients still at risk.

Figure 4 shows the estimate obtained by using Fine's method. The improvement that can be achieved in the estimate, when a suitable structure is specified and a proper association measure is used, is evident. The corresponding curve suggests a much stronger dependence between failure times: after only two years of follow up, the semi-competing risks estimate starts to decline with respect to the biased one and the discrepancy between them becomes much more appreciable: within five years (60 months) the estimates result in ${}^{SCR}\hat{S}_{T_{DM}}(t) = 0.8173$ and ${}^{B}\hat{S}_{T_{DM}}(t) = 0.8452$; later, at 12 years, the difference increases to ${}^{SCR}\hat{S}_{T_{DM}}(t) = 0.6508$ vs ${}^{B}\hat{S}_{T_{DM}}(t) = 0.7430$. This result reinforces the idea that occurrence of DM is always a serious, life-threatening event, and may be highly predictive of death.

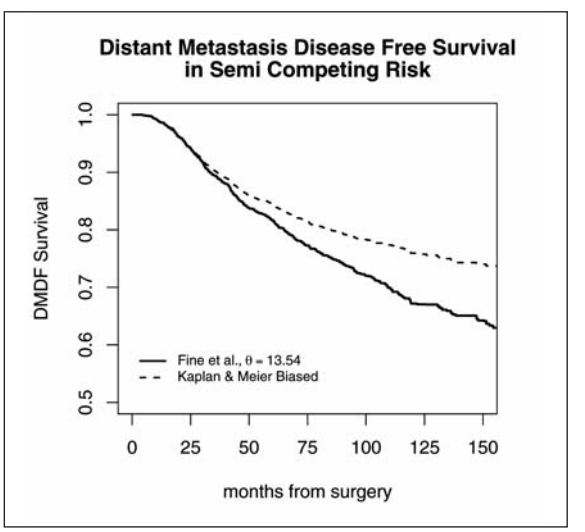

Figure 4. Semi-competing risks estimate of time to distant metastasis distribution.

Conclusion and discussion

To deal with multiple dependent failure times, different statistical methodologies have been developed, each one addressing a specific endpoint of interest. Other than in the univariate survival setting, there is no standardised approach, and often there is much confusion and some lack of agreement as to the proper statistical procedure to adopt. A careful definition of the clinical endpoint according to the study aims is the essential starting point, and this is not a simple matter. When different events are not of equal importance, the outcome measure of concern may be the time of occurrence of a specific event.

This issue is bound up with the estimate of the marginal, or net, distribution of treatment failure, in the presence of correlated events or competing risks. Its interpretation is not clear and is currently debated; indeed, because of this many statisticians have advocated the use only of observable crude functions in analysis of competing risks.

In the present work, the authors probed, in depth, the issue of the definition of the clinical endpoint of interest and the most appropriate corresponding methodology. They then went on to tackle the problem of estimating marginal survival distribution in the presence of competing events. Many proposals have been advanced in the literature. We investigated the ones that seem most suited to a clinical setting, given their basic assumption, and we proposed a generalisation to make them actually usable.

The performances of the proposed estimators were studied in preliminary simulation studies. Dependent failure times were generated from the copula model discussed in the methods, and also from a different Archimedean copula. These studies were mainly intended to evaluate the behaviour of the estimators under different amounts of both dependent, i.e. competing events, and independent censoring, by means of the association parameter, allowing it to take on different values, and by varying the shape and the scale parameter of the marginal distributions involved. Different percentages of independent censoring were obtained by an exponential distribution with different parameters. When the correct model was specified, the estimators showed a very good performance: the main difficulties in reproducing the true marginal distribution were due to the increased probability of the competing events. Under misspecification of the model, it appears to be much more difficult to fit the true marginal distribution, especially in the very short term.

Further studies may be conducted to compare the modified estimator with the non-parametric method originally proposed, managing any mixture of dependent and independent censoring in a similar way. Additional work in different areas is needed, and further research is also encouraged. At the moment, the main unsolved question and open issues concern the assessment of the overall adequacy of the model, and of the dependence structure, with reference in particular to the use of copula. From this perspective, the more recent field of semi-competing risks seems promising: the use of Archimedean copulas is appealing on account of their nice interpretation in terms of frailty models – albeit not always justified – and their mathematical tractability. On the basis of their experience, the authors recommend using all the information available on the event history and always resorting to the better approach of semi-competing risks in the presence of an ordered sequence of events. In the results presented, the advantage of the last approach was evident. A fully parametric approach, too, would be feasible and could be taken into account, even though it is not without its drawbacks. An alternative and surely more advantageous approach – given that fewer untestable assumptions are required – is that of the multi-state model, which in recent years has been investigated much more extensively thanks to recent software developments.

Acknowledgments

The work was partially supported by EU NoE Biopattern (FP6-2002-IST-1 No. 508803)

References

1. Fine JP, Jiang H, Chappell R. On semi-competing risks data. Biometrika 2001; 88: 907-920.
2. Prentice RL, Kalbfleisch JD, Peterson AV Jr, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics 1978; 34: 541-554.
3. Jiang H, Chappell R, Fine JP. Estimating the distribution of nonterminal event time in the presence of mortality or informative dropout. Control Clin Trials 2003; 24: 135-146.
4. Fusaro ER, Bacchetti P, Jewell NP. A competing risks

analysis of presenting AIDS diagnoses trends. Biometrics 1996; 52: 211-225.
5. Tsiatis A. A nonidentifiability aspect of the problem of competing risks. Proceeding of the National Academy Science USA 1975; 72: 20-22.
6. Slud EV, Rubinstein LV. Dependent competing risks and summary survival curves. Biometrika 1983; 70: 643-649.
7. Klein JP, Moeschberger ML. Bounds on net survival probabilities for dependent competing risks. Biometrics 1988; 44: 529-538.
8. Zheng M, Klein JP. A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula. Communication in Statististics Theory and Methods 1994; 23: 2299-2311.
9. Veronesi U, Saccozzi R, Del Vecchio M, Banfi A, Clemente C, De Lena M, Gallus G, Greco M, Luini A, Marubini E, Muscolino G, Rilke F, Salvadori B, A Zecchini, Zucali R. Comparing radical mastectomy with quadrantectomy, axillary dissection and radiotherapy in patients with small cancers of the breast. N Engl J Med 1981; 305: 6-11.
10. Peterson AV. Bounds for a joint distribution function with fixed sub-distribution functions: application to competing risks. Proceeding of the National Academy Science USA 1976; 73: 1113-1115.
11. Dignam JJ, Weissfeld LA, Anderson SJ. Methods for bounding the marginal survival distribution. Stat Med 1995; 14: 1985-1998.
12. Wang W, Wells MT. Estimation of Kendall's tau under censoring. Statistica Sinica 2000; 10: 1199-1215.
13. Veronesi U, Marubini E, Del Vecchio M, Manzari A, Andreola S, Greco M, Luini A, Merson M, Saccozzi R, Rilke F, Salvadori B. Local recurrences and distant metastasis after conservative breast cancer treatments: partly independent events. Journal of the National Cancer Institute 1995; 87: 19-27.
14. Koscielny S, Thames HD. Biased Methods for estimating local and distant failure rates in breast carcinoma and a "commonsense" approach. Cancer 2001; 92: 2220-2227.
15. Gelman R, Gelber R, Henderson IC, Coleman CN, Harris JR. Improved methodology for analyzing local and distant recurrence. J Clin Oncol 1990; 8: 548-555.
16. Alberti C, Métivier F, Landais P, Thervet E, Legendre C, Chevret S. Improving estimates of event incidence over time in populations exposed to other events. Applications to Three Large databases. Journal of Clinical Epidemiology 2003; 56: 536-545.
17. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, second edition. New York: John Wiley & Sons, 2002.
18. Marubini E, Valsecchi MG. Analysing survival data from clinical trials and observational studies. Chichester: John Wiley & Sons, 1995.
19. Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? Stat Med 1993; 12: 737-51.
20. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med 1999; 18: 695-706.

Appendix – More Details on the Slud-Rubinstein and Klein-Moeschberger Estimators

The dynamic dependence function reported in [6] can alternatively be written as:

$$\rho(t)=\left\{\left[f_{T_r}(t)/f_r(t)\right]-1\right\}\left\{\left[S_{T_r}(t)/S_{T_f}(t)\right]-1\right\}^{-1} \quad [9]$$

Furthermore, $\rho(t) = a$ if independence is assumed.
Slud and Rubinstein derived their estimator by solving the differential equation with respect to $S_{Tr}(t)$, and imposing a boundary condition with $S_{Tr}(0) = 1$: a closed-form expression is thus found:

$$S_{T_r}(t)=\exp\left\{-\int_0^t f_r(s)\rho(s)\left[S_{T_f}(s)\right]^{-1}ds\right\}* \quad [10]$$
$$*\left[1+\int_0^t f_r(s)\left(\rho(s)-1\right)\times\exp\left\{\int_0^s f_r(u)\rho(u)\left[S_{T_f}(u)\right]^{-1}du\right\}ds\right]$$

The final estimator then results in:

$$\overset{SR}{\hat{S}}{}^{\rho}_{T_r}(t)=N^{-1}\left\{n(t)+\sum_{k=0}^{d_r(t)-1}c_k\prod_{j=k+1}^{d_r(t)}\frac{n_j-1}{n_j+\rho_i-1}\right\} \quad [11]$$

where c_k is the number of subjects failed for competing events observed between t_k and t_{k+1}, $n(t)$ is the number of subjects at risk in t and $d_r(t)$ is the number of events of interest up to time t.
The authors of the present work obtained the following formulas for their adjusted Slud-Rubinstein estimator directly from equation 10 by having recourse to a plug-in of the corresponding estimators of involved quantities:

$$\hat{S}^{\rho}_{T_r}(t)=\exp\left\{-\sum_{s\le t}\hat{f}_r(s)\hat{\rho}(s)\left[\hat{S}_{T_f}(s)\right]^{-1}\right\}* \quad [12]$$
$$*\left[1+\sum_{s\le t}\hat{f}_r(s)\left(\hat{\rho}(s)-1\right)\exp\left\{\sum_{u\le s}\hat{f}_r(u)\hat{\rho}(u)\left[\hat{S}_{T_f}(u)\right]^{-1}\right\}\right]$$

The Klein-Moeschberger estimate for the marginal survival is given as:

$$S^{\theta}_{T_r}(t)=\left[1+(\theta-1)\int_0^t\frac{f_r(s)}{\left[S_{T_f}(s)\right]^{\theta}}ds\right]^{-1/(\theta-1)} \quad [13]$$

An "original" expression for the estimator, in the absence of independent censoring is thus:

$$\overset{KM}{\hat{S}}{}^{\theta}_{T_r}(t)=\left[1+(\theta-1)N^{(\theta-1)}\sum_{j:t_j\le t}\frac{d_{rj}}{n_j^{\theta}}\right]^{-1/(\theta-1)} \quad [14]$$

An alternative formulation for Klein-Moeschberger, which the current authors proposed in order to take into account independent censoring, is the following:

$$\hat{S}^{\theta}_{T_r}(t)=\left[1+\left(\hat{\theta}-1\right)\sum_{s\le t}\frac{\hat{f}_r(s)}{\left[\hat{S}_{T_f}(s)\right]^{\hat{\theta}}}ds\right]^{-1/\left(\hat{\theta}-1\right)} \quad [15]$$

where $\hat{f}_r(\cdot)$ is, clearly, obtained by differentiating the corresponding crude cumulative incidence function, reported in [2], as: $\hat{f}_r(t)=\hat{F}_r(t)-\hat{F}_r(t^-)$.