Mathematical and Computer Modelling 54 (2011) 1671–1675

Contents lists available at ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

An analysis of the recurrence–progression process in bladder carcinoma by means of joint frailty models

C. Santamaría [a,*], B. García-Mora [a], G. Rubio [a], S. Luján [b]

[a] Instituto Universitario de Matemática Multidisciplinar, Universidad Politécnica de Valencia, Spain

[b] Departamento de Urología, Hospital Universitario La Fe de Valencia, Spain

ARTICLE INFO

Article history:
Received 15 October 2010
Accepted 3 November 2010

Keywords:
Joint frailty model
Survival analysis
Event dependence
Bladder carcinoma

ABSTRACT

Multiple sequential recurrences are one of the most important characteristics of superficial transitional cell carcinoma of the bladder, more than 50% of the patients will have *recurrences* (reappearance of a new superficial tumor). When in the same subject recurrent events are considered and these observed events are clustered into groups, independence between the clustered survival times cannot be assumed. A natural way to model the dependence of clustered event times is through the introduction of a cluster-specific random effect: the *frailty term*. On the other hand, between 10% and 30% of patients diagnosed with bladder carcinoma will present a muscle invasive *progression*, so the observation process of *recurrences* could be interrupted by a major failure event (*progression*). In this regard, a joint modelling of the two processes could make the study of a joint evolution over time possible, giving unbiased and efficient parameters. We jointly analyze recurrences and progression processes by means of the *joint frailty model*.

1. Introduction

In recent years, there has been a growing interest in studying processes which generate events repeatedly over time such as recurrent infections in AIDS patients [1] and the process recurrence–progression in bladder carcinoma [2] among others. In this regard, several modeling approaches have been proposed to analyze this type of data [3] where the interest centers on the effect of covariates on the failure risk.

The general framework is the survival analysis area. Survival analysis is a set of statistical tools to analyze data related to times from an origin time until the occurrence of some or other event. The period of time from the start point to the event is the survival time, represented by a nonnegative random variable T for each individual. Generally two functions are of central interest: the survival function and the hazard function. The survival function $S(t)$ is the probability that survival time T be greater than t, $S(t) = P(T > t)$, and the hazard function is the instantaneous probability of event at time t per unit time, given survival up to time t. On the other hand, if the event of interest is not observed during the follow-up period the survival time is censored. It is usually required that survival time be independent of the censoring mechanism, that is to say, the follow-up of the individual is interrupted by causes independent of the event. Otherwise, the censoring is named informative.

When the event of interest occurs repeatedly in the same subject, a correlation between the recurrent relapse times may exist due to either *heterogeneity* among individuals or *event dependence*. *Heterogeneity* is produced because some subjects have a higher (or lower) event rate than the other ones, due to unknown or unmeasurable effects such as, for example

* Corresponding author. Tel.: +34 963879144.
E-mail address: crisanna@imm.upv.es (C. Santamaría).

0895-7177/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2010.11.004

a genetic susceptibility to developing some disease and consequently relapsing more quickly than the rest of population. *Event dependence* can be produced when the occurrence of a given event may make further relapses more or less likely. This event dependence may be produced by a learning process or by biologically weakening/strengthening the body and implies that the occurrence of a relapse itself may raise (or lower) the subsequent event possibility. Any correlation among events can be produced by heterogeneity, event dependence or jointly.

Furthermore, this type of repeated events process, such as recurrent tumors or infections, could be terminated by a *terminating event* such as death. In this context, the major failure event could be correlated with recurrent events and the usual assumption of noninformative censoring of the recurrent event process by *terminating event* can be violated. This dependence should be accounted for in the joint modelling of *recurrent events* and *terminating event* [4] assuming the *terminating event* as a censoring event.

Recently joint modelling for two survival processes has received special attention because it makes possible to deal with informative censoring for recurrent event data and it also allows us to study the joint evolution over time of two processes: *recurrent events* and *terminating event*. In this regard, we use a general *joint frailty model* [5] where a non-parametric penalized likelihood method for estimating hazard functions is used and this gives unbiased and efficient parameters.

The analysis we develop in this paper is motivated by a study of patients diagnosed with bladder carcinoma. The course of this disease is characterized by a primary tumor which is managed with a transurethral resection (*TUR*), followed by superficial *recurrent tumors* (*recurrences*) in the bladder or a muscle invasive *progression* (*terminating event*). As there is an association between these, we jointly study them by means of the *joint frailty model*, which will allow us simultaneously to incorporate the effects of covariates, the impact of accumulating event occurrences on the subjects, the effect of latent or unobserved variables which, for each patient, endow correlation among the inter-event times. We will obtain the hazard and survival functions associated with both *recurrence* and *progression* that would allow us to establish accurate predictions of this carcinoma.

The paper is organized as follows: in Section 2, we present the models to study both the *recurrence* and *progression* processes. In Section 3, we jointly analyze them in an application to bladder carcinoma. Finally, the most relevant concluding remarks are summarized in the Section 4.

2. Modelling the recurrence–progression processes

After the *TUR*, each patient can undergo one or more *recurrences* that can be interrupted by the occurrence of the *progression* or suffering the *progression* after the *TUR*. Each patient is monitored regularly by routine-visits and the presence of a new *recurrence* or a *progression* is notified at each visit. Those patients who have experienced neither recurrence nor progression are *censored*.

2.1. Modelling the recurrence process

We consider the first event as the *first recurrence*, the second event as the *second recurrence*, and so on. The follow-up period of a patient is completed by a *progression* or by a censoring time. We use the Andersen and Gill model (referred to as AG model) [6] where the risk at time t for the kth recurrence of the ith patient conditionally on observed covariates X_{ik} is given by the expression

$$\lambda_k(t; X_{ik}) = \mathrm{e}^{\beta' X_{ik}(t)} \lambda_0(t) \tag{1}$$

where $\lambda_0(\cdot)$ is an unknown baseline hazard rate function and β is a vector of parameters associated with covariate effects that can be time dependent. The AG model is ideally suited to the situation of mutual independence of the observations within a subject. However, if we assume that after experiencing the first recurrence, the risk of the next recurrence may increase, this would suggest the introduction of a time-dependent covariate to capture the dependence structure among the recurrence times. So we will introduce in (1) the variable *prior number* of recurrences to capture this effect.

Another way to take into account the presence of correlation between *recurrences* is through of the introduction of a cluster-specific random effect: *the frailty term*. In this situation, the hazard function of the ith subject for the kth event is defined by

$$\lambda_{ik}(t; X_{ik}, Z_i) = Z_i \mathrm{e}^{\beta' X_{ik}(t)} \lambda_0(t) \tag{2}$$

where the correlation between *recurrences* is modelled with the unobserved frailties Z_i which capture the *heterogeneity* among subjects. Z_i follows a gamma distribution with mean 1 and unknown variance ξ, and the estimated value of ξ determines the degree of association among *recurrences*.

2.2. Modelling the recurrence–progression process

We analyze a joint treatment of two processes over time and use a non-parametric penalized likelihood method for estimating hazard functions in a general *joint frailty model* for *recurrent events* and the *terminating event*: in our case