THE ANALYSIS OF MULTIVARIATE INCOMPLETE FAILURE TIME DATA

MD. BAKI BILLAH

The Analysis of Multivariate Incomplete Failure Time Data

By

©Md. Baki Billah

A practicum submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of
Master's of Applied Statistics
Faculty of Science
Memorial University of Newfoundland
July, 1995

St. John's Newfoundland Canada

Abstract

In survival studies the individual study subjects may experience multiple failures. These failures may be repetitions of the same kind of event or may be events of different natures. Most of the multivariate incomplete failure time problems discussed so far in the literature are of the former type. The multivariate incomplete failure time problem of the second type has not been adequately addressed in the literature. In this practicum, we concentrate on the multivariate incomplete failure time problem of the first kind. Many authors, for example, Wei, Lin, and Weissfeld (1989, JASA) have analysed this type of multivariate incomplete failure time problem by using the univariate partial likelihood approach. The application of the univariate partial likelihood approach to such correlated failure time data may not, however, reveal the actual effect of the treatment. To overcome this problem, we propose an ad hoc modification to this type of multivariate incomplete failure time data, in order to make the failure times (recorded at different stages) independent, and then apply the univariate partial likelihood approach to obtain (estimate) the treatment effects. Further, under the assumption that the treatment effect remains the same all throughout the study group (failure group), we estimate the combined treatment effects for multivariate incomplete failure time data, by using the restricted partial likelihood estimation (RPLE) method. The univariate partial likelihood method applied to the modified data appears to provide more appealing inferences about the treatment effects than when this method is applied to the original data. Also the restricted partial likelihood estimation (RPLE) method appears to provide more precise estimates for the combined treatment effects as compared to the linear estimation (LE) method used in the literature, for example, by Wei et al. (1989).

Acknowledgement

I am very grateful to my supervisor, Professor B.C. Sutradhar, for leading me to the research field of this practicum , and for his many helpful and thoughtful comments, discussions and suggestions throughout the preparation of this practicum. Also I wish to thank him for his support, patience and kindness, which made my study here very pleasant and fruitful. Without his guidence, this practicum would have been impossible.

I would like to thank Professor Bruce Shawyer for his invaluable assistance during the typing of this practicum. He was always open to discussions and a great source of inspiration.

The Department of Mathematics and Statistics deserves my special thanks for providing me with the facilities that I needed and a friendly working environment.

Finally, I am grateful to the School of Graduate Studies and the Department of Mathematics and Statistics for supporting me with a Memorial University Fellowship and a Teaching Assistantship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background of Multivariate Incomplete Failure Time Study

Multivariate failure time data arise when each subject experiences several types of event or when there are clustering of observational units such that failure times within the same cluster are correlated. More specifically, one encounters in practice multivariate failure time problems in the following two ways:

(i) **Repetitions of the same event:** In this case each individual may experience two or more distinct failures and these failures may be due to the repetitions of the same kind of event.

(ii) **Failure of distinct event:** Here each subject may experience two or more distinct failures and these failures may be events of different natures.

Most of the multivariate failure problems considered so far in the literature belong to the first group. Frequently this type of data can be found in biomedical sciences. For example, Makuch and Parks (1988) considered the following multivariate incom-

plete failure time problem. In their study, in order to evaluate the effectiveness of the drug ribavrin, patients with acquired immune deficiency syndrome (AIDS) were randomly assigned to one of the three groups: placebo, low-dose ribavrin, and high-dose ribavrin. Blood samples for each patient were collected at weeks 4, 8 and 12 and for each serum sample, measurements of p24 antigen levels. (which are important markers of HIV-1 infection) were recorded. The "viral load" in each serum was evaluated by measuring the number of days when virus positivity was detected, that is, when the p24 level was greater than 100 picograms per milliliter. Therefore, potentially each patient in the study should have three such event times (number of days) corresponding to the samples taken at weeks 4. 8, and 12, which are repetitions of the same event for three times. This problem may be treated as a 3- dimensional multivariate failure time problem. Note, however, that in this 3-dimensional multivariate study, some observations were missing, because patients did not make the scheduled visits or because serum specimens were inadequate for laboratory analysis. In addition, censored observations occurred when the culture required a longer period of time to register as virus positive than was achievable in the laboratory, or when the serum sample was contaminated before positivity was detected. Based on these virological data, one would like to know, for example, whether the drug ribavrin effectively prolonged the time to virus positivity and how the drug effects changed over time. This 3-dimensional multivariate problem, because of the occurance of censored observations, now becomes a multivariate incomplete failure time based regression problem.

A similar multivariate failure time problem, in the context of bladder cancer, was studied by the Veterans Administration Cooperative Urological Research Group (Byar 1980). In this study, all patients had superficial bladder tumors when they entered

the trial. These tumors were removed transurethrally and patients were randomly assigned to one of three treatments: placebo, thiotepa and pyridoxine. Many patients had multiple recurrences of tumors during the study and new tumors were removed at each visit. Here each recurrence time of a patient was measured from the beginning of his/her treatment. Since each individual visited the clinic multiple numbers of times, the experiment belongs to a multivariate failure time problem. As indicated in Byar (1980), one of the analyses to evaluate the effectiveness of thiotepa should be based on the tumor recurrence times computed from patients based on patient's visit. This tumor recurrence data for bladder cancer patients was later analysed by Wei, Lin and Weissfeld (1989).

Recently, Guo and Lin (1994) has referred to another multivariate failure study conducted by D. H. Uttal of Northwestern University in the context of a psychological problem. In this psychological experiment, the main objective was to study the childrens' abilities to locate hidden objects. For this a sample of 83 children was considered and each of the 83 children was asked to search for objects hidden in 10 different locations. For each location, the child was given three chances to find the object. The experiments might differ in terms of whether a map was taken while searching for the object and whether the map was rotated. These two factors were expected to affect how quickly the child could find the object. The child's age (which was categorized into groups, 4-5 years vs. 6-7 years old) was also thought to be predictive. Since each child was using similar technique to detect hidden objects in 10 different locations, this problem may be considered as a multivariate failure time problem of the first kind, where every finding is a repetition of the first finding.

Huster et al. (1989) and Liang et al. (1993) used an example on Diabetic Retinopathy Study (DRS) which was begun in 1971 to study the effectiveness of

laser photocoagulation in delaying the onset of blindness in patients with diabetic retinopathy. Diabetic retinopathy is a complication associated with diabetes mellitus consisting of abnormalities in the microvasculature within the retina of the eye. It is the leading cause of new cases of blindness in patients under 60 years of age in the United States and is the major cause of visual loss elsewhere in many industrialized countries (Murphy and Patz, 1978). Patients with diabetic retinopathy in both eyes and visual acuity of 20/100 or better in both eyes were eligible for the study. One eye of each patient was randomly selected for treatment and the other eye was observed without treatment. The total study size was 1,742 patients followed over several years. A sample of size N=197 of the high-risk patients as defined by DRS criteria was considered for the analysis. The end-point used to assess the treatment effect was the occurence of visual acuity less than 5/200 at two consecutively completed 4-month follow-ups. Since each patient in the study should have two event (same kind) times, this problem also may be considered as a two-dimensional multivariate failure time problem of the first kind.

The primary question of the DRS study was to assess the effectiveness of the laser photocoagulation treatment. Secondary questions were whether the survival times for the eyes of a patient were related and whether the treatment and type of diabetes were related.

Another example can be found in the experiment conducted by Thompson et al. (1978) on the development of mammary cancer. In this experiment seventy-six rats were injected with a carcinogen for mammary cancer at day zero, and then all animals were given retinal acetate to prevent cancer for sixty days. After 60 days, the 48 animals which remained tumor-free were randomly assigned to continued retinoid prophylaxis (Treatment Group 1) or control (Treatment Group 2). Rats were palpated

for tumors twice weekly, and observation ended 182 days after the initial carcinogen injection. The times to development of mammary cancer were measured from the day of injection. Since the same kind of event occurred repeatedly to each animal, the data may be considered as multivariate failure time data of the first kind.

The multivariate failure time problem of the second type, that is, failure of distinct event, has not been addressed adequately in the biostatistical literature. This problem can, however, occur in many practical situations. For example, AIDS researchers are often interested in time to the drop of the CD4-lymphocyte count below a threshold, in time to the first detectable level of HIV antigen, as well as in times to prominent changes of other biological markers for an HIV-infected person. Here, this is a 3 or more dimensional multivariate failure problem of the second kind. The scientific interest of a multivariate survival study of this kind typically lies in the effects of covariates on the risks for failures.

In this practicum, we will concentrate on the multivariate incomplete failure time problem of the first kind. We will mainly be concerned with the methodological development in estimating as well as testing the treatment effect in a multivariate incomplete failure time problem. For this purpose, we review and examine the effect of the univariate partial likelihood approach applied by Wei et al. (1989) to a bladder cancer data set as mentioned before. In this regard, it is to be noted that application of such a partial likelihood approach to a correlated data set of first kind may not reveal the actual effect of the treatment. As a remedy, a modification to the bladder cancer data is proposed to make the multivariate data groupwise independent, and then we apply the univariate partial likelihood approach to analyse such data. In contrast to Wei et al.'s (1989) linear estimation method we provide a likelihood based estimation approach to estimate the combined cancer treatment effects and to study

the inference about them.

The specific outline of the practicum is given in the next section.

1.2 Outline of Practicum

1.2.1 Review of Univariate Incomplete Failure Time Regression Model

A review of analysis of incomplete failure time data is given in chapter 2. In section 2.1, attention is given to the partial likelihood method of estimation under Cox's (1975) proportional hazards models. This method was used by Wei et al. (1989) in estimating the treatment effects from bladder cancer data as mentioned above. The estimation method and test for regression parameters are described in section 2.2.

1.2.2 Multivariate Incomplete Failure Time Data Analysis

Chapter 3 concerns the analysis of multivariate incomplete failure time data. The regression analysis of such multivariate incomplete (correlated) data is not adequately studied in the literature. In this case, the main difficulty is to develop a suitable methodology for regression effects, taking the correlation of the failure times into account. Prentice et al. (1981), Wei et al.(1989), and Wei and Lin (1989) have, however, used the univariate partial likelihood approach to analyse such multivariate incomplete failure time data. Regression estimates in such cases may not be reliable. To overcome this problem, modification to the data is proposed in order to make the failure times as independent as possible, and the univariate partial likelihood method is then applied.

In section 3.1, we outline the univariate partial likelihood method to analyse the multivariate incomplete failure time data. The estimation method and test for regression parameters are outlined in section 3.2. In section 3.3, in contrast to Wei et al. (1989), we discuss a method to estimate the combined treatment effects for multivariate incomplete failure time data by using the restricted partial likelihood estimation (RPLE) method. In section 3.4, a method is given to modify the multivariate failure time data in order to make all failure groups independent of each other.

1.2.3 Analysis of Modified Bladder Cancer Data

The analysis of modified bladder cancer data is discussed in chapter 4. In section 4.1, three different approaches are discussed to analyse the multivariate incomplete failure time data, and these three approaches are compared for each recurrence group. Estimation and test for treatment effects in different recurrence groups for the modified as well as the original blood cancer data is discussed in section 4.2. The application of the univariate partial likelihood estimation method to the modified data seems to provide more appealing interpretations than when this method is applied to the original data. In contrast to Wei et al's linear estimation (LE) method a restricted partial likelihood estimation (RPLE) method is discussed to estimate the combined treatment effects in section 4.3. and it is found that the restricted partial likelihood estimation (RPLE) method is more precise than the method used by Wei et al. (1989). In section 4.4, the likelihood ratio and Wald type tests are performed for simultaneous treatment effects.

Chapter 2

Review of Univariate Incomplete Failure Time Regression Model

2.1 Introduction

Over the last two decades there has been a great deal of interest in the analysis of censored data, particularly in the context of survival analysis in medical trials where patients often survive beyond the end of the trial period or are lost to follow-up for various reasons. For the univariate failure time regression study with censored observations, many parametric and non-paramertric methods are available in the literature (cf. Lawless 1982). Here, attention is focused on the methods of estimation under proportional hazards models, suggested by Cox (1972). The proportional hazards model is non-parametric in the sense that it involves an unspecified function in the form of an arbitrary base-line hazard function. Therefore, this model is flexible, but different approaches are required for estimation and testing under different conditions.

Let n be the number of study subjects and T_i be the failure time of the ith individual ($i = 1, 2, \cdots, n$). Let C_i be the corresponding censoring time. Instead of

observing failure times $T_1, T_2, \cdots, T_n$, we observe

$$X_i = \min(T_i, C_i).$$

Define the indicator variable

$$\begin{aligned} \delta_i &= 1, \quad \text{if} \quad T_i \leq C_i \quad \text{(uncensored)} \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Whenever T_i is missing, we consider $C_i = 0$. Because T_i is always nonnegative, this implies that $\delta_i = 0$, in such missing cases. Now, let $Z_i(t) = (Z_{i1}(t), \cdots, Z_{ip}(t))'$ denote a $p \times 1$ vector of covariates for the ith subject at time $t \geq 0$. We assume that $(T_i, C_i, Z_i(.)), i = 1, \cdots, n$, are independently identically distributed random quantities. In such studies our main interest is to know the effect of the covariates on the failure time of the individuals. For the estimation purpose, Cox's proportional hazard model is widely used. The Cox model for censored survival data specifies that the hazard function for the ith $(i = 1, \cdots, n)$ individual is given by $\lambda_i(t, z_i) = \lim_{h \to 0} \frac{1}{h} Pr\left[T \leq t + h | T > t; z_i\right], t \geq 0$, where z_i corresponds to $z_i(t)$. For simplicity, all throughout $\lambda_i(t, z_i)$ will be denoted by $\lambda_i(t)$ which is given by

$$\lambda_i(t) = \lambda_0(t) \exp(\beta' Z_i(t)), \quad \text{if} \quad t \geq 0 \tag{2.1}$$

where $\lambda_0(t)$ is an arbitrary and unspecified base-line hazard function and $\beta = (\beta_1, \cdots, \beta_p)'$ is a vector of regression coefficients, indicating the effects of the covariates.

Here one may also want to estimate $\lambda_0(t)$, from censored data. One approach is to maximize the likelihood function for the observed data simultaneously with respect to β and $\lambda_0(t)$. A more attractive approach is one based on the concept of partial likelihood as presented by Cox (1975). In Cox's approach, the likelihood function for

β does not depend upon $\lambda_0(t)$, which then can be maximized to give an estimate of β and to provide tests for β in the absence of knowledge of $\lambda_0(t)$. Once β has been estimated, $\lambda_0(t)$ can be estimated by maximizing the full likelihood function for $\lambda_0(t)$, assuming that β is equal to the maximum partial likelihood estimate (m.p.l.e.) of β. Partial likelihood applied to censored data problems has been discussed by Cox (1975), Efron (1977), Kalbfleisch and Mackay (1978), Kalbfleisch and Prentice (1980), and others. For the sake of completeness, we now outline the partial likelihood as well as the estimation method and inferences for β based on the partial likelihood.

2.2 Partial Likelihood

Assume that in a random sample of n study subjects we have a sample of r distinct observed failure times and $n - r$ censoring times. Let the r individuals be observed to fail at $T_1, \cdots, T_r$. The order statistics corresponding to these failure times ordered from smallest to largest are denoted by $T_{(1)} <, \cdots, < T_{(r)}$. Let $R(T_{(i')})$ denote the risk set at $T_{(i')} - 0$, that is, the set of individuals alive and uncensored just prior to $T_{(i')}$, for $i' = 1, \cdots, r$. Now the partial likelihood function suggested by Cox (1972,1975) for estimating β in the absence of knowledge of $\lambda_0(t)$ is

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\beta' Z_i(X_i))}{\sum_{l \in R(X_{(i)})} \exp(\beta' Z_l(X_i))} \right]^{\delta_i} \tag{2.2}$$

where $Z_i(X_i)$ is the covariate vector associated with the *ith* individual at time $X_i = \min(T_i, C_i)$.

The above likelihood function does not depend on $\lambda_0(t)$ and is traditionally maximized to estimate the regression parameter vector β.

In practice ties frequently occur in data on continous variables because of rounding off or grouping. If there are only a few ties, the partial likelihood can be obtained

based on one of the modifications suggested by Peto (1972a) as

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\beta' S_i)}{\left(\sum_{l \in R(X_{(i)})} \exp(\beta' Z_l(X_i)) \right)^{d_i}} \right]^{\delta_i} \tag{2.3}$$

where d_i is the number of lifetimes equal to $X_{(i)}$ and S_i is the sum of the covariate vectors Z for these d_i individuals. In such tie situations, the regression parameters are estimated by maximizing (2.3) instead of (2.2).

It is important to note that the partial likelihood method is distribution free and certain properties of the procedure do not depend on the underlying lifetime distribution, more specifically, on the baseline hazard function $\lambda_0(t)$. This is actually true when there is no censoring, but with many types of censoring the dependence on $\lambda_0(t)$ is small. If the distribution form of $\lambda_0(t)$ is known. that is , the data come from a particular known baseline hazard function, there will be some loss of efficiency in using the non-parametric approach instead of the correct parametric model. In some cases, however, this loss of efficiency is slight (cf. Lawless 1882). It should also be mentioned that the regression parameter vector β can be estimated by direct maximum likelihood rather than through the partial likelihood. Maximum likelihood is, however, less convenient than the partial likelihood method because it requires simultaneous consideration of β and $\lambda_0(t)$.

The estimation of β based on Cox's partial likelihood and the inference about β are now summarized below.

2.3 Estimation and Test for β Based on Partial Likelihood Approach

The log likelihood arising from (2.2) is given by

$$\log L(\beta) = \sum_{i=1}^{n} \delta_i \beta' Z_i(X_i) - \sum_{i=1}^{n} \delta_i \log \left(\sum_{l \in R(X_{(i)})} \exp(\beta' Z_l(X_i)) \right) \tag{2.4}$$

and the first derivatives of $\log L$ with respect to the elements of β are

$$\begin{aligned} U(\beta_u) &= \frac{\partial \log L}{\partial \beta_u} \\ &= \sum_{i=1}^{n} \delta_i Z_{iu}(X_i) - \sum_{i=1}^{n} \delta_i \left[\frac{\sum_{l \in R(X_{(i)})} Z_{lu}(X_i) exp(\beta' Z_l(X_i))}{\sum_{l \in R(X_{(i)})} exp(\beta' Z_l(X_i))} \right] \\ &= \sum_{i=1}^{n} \delta_i Z_{iu}(X_i) - \sum_{i=1}^{n} \delta_i \left[\frac{\sum_{l=1}^{n} Y_l(X_i) Z_{lu}(X_i) exp(\beta' Z_l(X_i))}{\sum_{l=1}^{n} Y_l(X_i) exp(\beta' Z_l(X_i))} \right] \end{aligned} \tag{2.5}$$

for $u = 1, \cdots, p$, where $Y_l(t) = I(X_l \geq t)$ with

$$\begin{aligned} I &= 1, \quad \text{if} \quad X_l \geq t \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Substituting $g_l(\beta) = Y_l(X_i) exp(\beta' Z_l(X_i))$ in (2.5) the score function reduces to

$$U(\beta_u) = \sum_{i=1}^{n} \delta_i Z_{iu}(X_i) - \sum_{i=1}^{n} \delta_i \left[\frac{\sum_{l=1}^{n} Z_{lu}(X_i) g_l(\beta)}{\sum_{l=1}^{n} g_l(\beta)} \right]. \tag{2.6}$$

The second partial derivatives of (2.4) are

$$\begin{aligned} I_{uw}(\beta) &= \frac{\partial^2 \log L}{\partial \beta_u \partial \beta_w} \\ &= -\sum_{i=1}^{n} \delta_i \left[\frac{(\sum_{l=1}^{n} Y_l(X_i) Z_{lu}(X_i) Z_{lw}(X_i) \exp(\beta' Z_l(X_i))) (\sum_{l=1}^{n} Y_l(X_i) \exp(\beta' Z_l(X_i)))}{(\sum_{l=1}^{n} Y_l(X_i) \exp(\beta' Z_l(X_i)))^2} \right. \\ &\quad \left. - \frac{(\sum_{l=1}^{n} Y_l(X_i) Z_{lu}(X_i) \exp(\beta' Z_l(X_i))) (\sum_{l=1}^{n} Y_l(X_i) Z_{lw}(X_i) \exp(\beta' Z_l(X_i)))}{(\sum_{l=1}^{n} Y_l(X_i) \exp(\beta' Z_l(X_i)))^2} \right] \\ &= -\sum_{i=1}^{n} \delta_i \left[\frac{(\sum_{l=1}^{n} Z_{lu}(X_i) Z_{lw}(X_i) g_l(\beta)) (\sum_{l=1}^{n} g_l(\beta))}{(\sum_{l=1}^{n} g_l(\beta))^2} \right. \\ &\quad \left. - \frac{(\sum_{l=1}^{n} Z_{lu}(X_i) g_i(\beta)) (\sum_{l=1}^{n} Z_{lw}(X_i) g_l(\beta))}{(\sum_{l=1}^{n} g_l(\beta))^2} \right]. \end{aligned} \tag{2.7}$$

The maximum partial likelihood estimator $\hat{\beta}$ for β is defined as the solution to the likelihood equation

$$\frac{\partial \log L}{\partial \beta} = 0$$

which can be solved easily by the Newton-Raphson iterative method utilizing (2.6) and (2.7) as

$$\hat{\beta}^{h+1} = \hat{\beta}^{h} - D_h^{-1} U(\hat{\beta}^{h})$$

where $\hat{\beta}_h$ denotes the value of β at the hth iteration. $U(\hat{\beta}^{h})$ is the score vector $U(\beta) = (U(\beta_1), \cdots . U(\beta_p))'$ evaluated at $\beta = \hat{\beta}^{h}$, and

$$D_h = \left(\frac{\partial^2 \log L}{\partial \beta_u \partial \beta_w} \right)_{\hat{\beta}^{h}}$$

is the $p \times p$ matrix with general (u, w)th element given by (2.7) evaluated at $\hat{\beta}^{h}$.

It appears that maximum partial likelihood estimates (m.p.l.e.'s) obtained by maximizing $L(\beta)$ possess the usual asymptotic properties of ordinary maximum likelihood estimates under quite broad conditions. Cox (1975) and Kalbfleisch and MacKay (1978) give heuristic treatments that attempt to place only mild conditions on the censoring and lifetime processes. Alen (1978) gives some relevant results as part of a general treatment of counting processes. Tsiatis (1978a) and Liu and Crowley (1978) demonstrate under models involving random independent censoring mechanisms that the m.p.l.e. is consistent and asymptotically normal and the likelihood ratio tests based on $L(\beta)$ are valid.

Now to make inferences about β one can use a very simple approach by treating $\hat{\beta}$ as being approximately normal with mean β and covariance matrix $I(\hat{\beta})^{-1}$, where

$$I(\hat{\beta}) = \left(-\frac{\partial^2 \log L}{\partial \beta_u \partial \beta_w} \right)_{\hat{\beta}} \quad \text{with} \quad u = 1, \cdots, p; \quad w = 1, \cdots, p$$

which implies that under $H_0 : \beta = \beta_0$

$$(\hat{\beta} - \beta_0)' I^{-1}(\hat{\beta})(\hat{\beta} - \beta_0) \tag{2.8}$$

has asymptotically a χ^2 distribution with p degrees of freedom.

Inferences can also be based on likelihood ratio methods, in which case,

$$\lambda = -2\log\frac{L(\beta_0)}{L(\hat{\beta})} \tag{2.9}$$

has an asymptotic χ^2 distribution with p degrees of freedom.

Further, for testing the hypothesis $H_0 : \beta = \beta_0$, one can construct the well-known score test statistic

$$U(\beta_0) I^{-1}(\beta_0) U(\beta_0) \tag{2.10}$$

which again has an asymptotic χ^2 distribution with p degrees of freedom. Here

$$U(\beta_0) = \frac{\partial \log L}{\partial \beta_u} \Big|_{\beta_0}$$

is the score function which has an asymptotic normal distribution with mean zero and covariance matrix $I^{-1}(\beta_0)$.

In some survival studies, it is, seen that one patient may experience two or more distinct failures. Suppose that each patient experiences K (≥ 2) such failures. These K failures may be due to repetitions of the same kind of event or they may be events of different nature. In either case, these K failure times for a patient will be correlated. The regression analysis of such multivariate failure time data is not adequately discussed in the literature. While, in general, it is not easy to take the correlation of the data into account to analyse such multivariate failure regression data, some authors (cf. Prentice et al. (1981) Wei et al. (1989), Wei and Lin (1989)) have used the univariate partial likelihood approach to analyse the multivariate failure time data

which arise due to the repetitions of the same kind of event. Regression estimates in such cases are usually consistent but will not be fully efficient. This is because, in this approach, estimates are obtained based on the assumption that the K groups of failure time data are independent, when in practice they are correlated. Nevertheless, in the next chapter, we outline the univariate partial likelihood approach used by Wei et al. (1989), among others, to analyse the multivariate incomplete failure time data. In the same chapter, we suggest a modification to the multivariate failure time data discussed by Wei et al. (1989), in order to make the K groups of failure times almost independent of each other. The application of the univariate partial likelihood method to such modified data will naturally produce regression estimates similar to the estimates obtained by certain suitable methods using the correct correlation matrix of the failure time data. The modification of the multivariate failure time data is discussed in the context of cancer data considered by Wei et al. (1989).

Chapter 3

Multivariate Incomplete Failure Time Data Analysis

3.1 Introduction

In many survival studies we record the times of two or more distinct failures on each subject. These failures may be events of different natures or may be repetitions of the same kind of event, as mentioned in the last chapter. Several regression methods have been proposed in the literature to deal with situations where individuals experience repeated failures such as multiple tumor recurrences. These methods impose specific structures of dependence among the recurrences on each subject and can be thought of as generalizations of survival data techniques in which the hazard function modeling is continued beyond a subject's first failure to the second and subsequent failures. For example, Lawless (1987) presented a class of parametric and semiparametric procedures based on nonhomogeneous Poisson process models with proportional intensity assumptions. The counting process formulation of Andersen and Gill (1982) can be regarded as a special case of the Cox proportional intensity

model. The approach of Prentice, Williams, and Paterson (1981) differs from that of Andersen and Gill (1982) in two aspects: (*a*) the risk sets for the $(k+1)$th recurrences are restricted to the individuals who have experienced the first k recurrences; and (*b*) the underlying intensity functions and regression parameters are allowed to vary among distinct recurrences. The method of Gail, Santner, and Brown (1980) is a two-sample special case of Prentice et al. (1981). Wei et al. (1989) proposed semiparametric methods to analyze general multivariate failure time data. Here, the multivariate failure time data correspond to a $K \times 1$ dimensional random vector, K being the number of recurrence groups. Wei et al. (1989) model such multivariate data by using the marginal distribution of each of the K groups. No particular structure of dependence among distinct failure times on each subject was imposed here. The regression parameters were estimated by maximizing the failure-specific partial likelihoods, and the resulting estimators across all types of failures were shown to be asymptotically jointly normal with a covariance matrix that can be easily estimated from the data. In this chapter we outline the regression analysis of multivariate failure time observations following Wei et al. (1989), among others by using the univariate partial likelihood method and we suggest a modified multivariate incomplete failure time data in the context of tumor recurrence data used by Wei et al. (1989). Also, unlike Wei et al. (1989), we discuss a method to estimate the combined treatment effects for the multivariate data by using the restricted partial likelihood estimation (RPLE) method.

3.2 Univariate Partial Likelihood Approach

Let there be n individuals in a study and let T_{ki}, $(i = 1, \cdots, n;\ k = 1, \cdots, K)$ be the kth failure time of the ith subject. Here, instead of T_{ki} one observes the bivariate vector (X_{ki}, δ_{ki}), where

$$X_{ki} = \min(T_{ki}, C_{ki}),$$

C_{ki} is the censoring time and

$$\delta_{ki} = \begin{cases} 1, & \text{if } T_{ki} \leq C_{ki} \text{ (uncensored)} \\ 0, & \text{otherwise.} \end{cases}$$

If T_{ki} is missing, we assume that $C_{ki} = 0$. Since T_{ki} is always positive, this implies that $X_{ki} = 0$ and $\delta_{ki} = 0$. Now, let us assume that $Z_{ki}(t) = (Z_{ki1}(t), \cdots, Z_{kip}(t))'$ denotes a $p \times 1$ vector of covariates for the ith patient at time $t \geq 0$ with respect to the kth type of failure. Conditional on Z_{ki}, the failure vector $T_i = (T_{1i}, \cdots, T_{Ki})'$ and censoring vector $C_i = (C_{1i}, \cdots, C_{Ki})'$ $(i = 1, \cdots, n)$ are assumed to be independent. Furthermore, one can assume that $(X_i, \delta_i, Z_i(.))$ $(i = 1, \cdots, n)$, where $Z_i = (Z_{1i}', \cdots, Z_{Ki}')$, are independently identically distributed (iid) random vectors with bounded covariates $Z_i(.)$. In such multivariate failure time data to estimate the effects of covariates, the widely used Cox's proportional hazard function $\lambda_{ki}(t)$, for the kth type of failure of the ith subject has the form

$$\lambda_{ki}(t) = \lambda_{k0}(t)\exp(\beta_k' Z_{ki}(t)), \quad t \geq 0 \tag{3.1}$$

where $\lambda_{k0}(t)$ is an arbitrary and unspecified baseline hazard function and $\beta_k = (\beta_{k1}, \cdots, \beta_{kp})'$ is the failure-specific covariate parameter vector.

As in the univariate case to estimate the regression parameter β one can use the Cox's (1972, 1975) partial likelihood approach. Let us assume that in the above

random sample of n individuals there are r distinct observed failure times and $n-r$ censoring times with respect to kth failure group. Let $R_k(t) = \{l : X_{kl} \geq t\}$, that is, the set of subjects at risk just prior to time t with respect to the kth type of failure. Then similar to (2.2), Cox's (1972,1975) partial likelihood for the kth group is

$$L(\beta_k) = \prod_{i=1}^{n} \left[\frac{\exp(\beta_k' Z_{ki}(X_{ki}))}{\sum_{l \in R_k(X_{(ki)})} \exp(\beta_k' Z_{kl}(X_{ki}))} \right]^{\delta_{ki}} \tag{3.2}$$

where Z_{ki} is the covariate vector corresponding to the individual observed to die at $X_{(ki)}$.

Then the maximum partial likelihood estimator $\hat{\beta}_k$ for β_k is the value that maximizes the partial likelihood function (3.2).

In case of a few ties the likelihood function (3.2) is replaced by

$$L(\beta_k) = \prod_{i=1}^{n} \left[\frac{\exp(\beta_k' S_{ki}(X_{ki}))}{\left(\sum_{l \in R_k(X_{(ki)})} \exp(\beta_k' Z_{kl}(X_{ki}))\right)^{d_{ki}}} \right]^{\delta_{ki}} \tag{3.3}$$

where d_{ki} is the number of failure times equal to X_{ki} and S_{ki} is the sum of the covariate vectors Z_{ki} for these d_{ki} subjects.

Therefore, in tie situations to estimate the regression parameters one should use the likelihood function (3.3). But the likelihood function (3.3) is more difficult for computation than that of (3.2). However, as recommanded in the literature (cf. Lawless, 1982) that (3.3) can be approximated by (3.2). We discuss the estimation and test based on (3.2) as follows.

3.3 Estimation and Tests For β:

The log likelihood arising from (3.2) is given by

$$\log L(\beta_k) = \sum_{i=1}^{n} \delta_{ki} \beta_k' Z_{ki}(X_{ki}) - \sum_{i=1}^{n} \delta_{ki} \log \left(\sum_{l \in R(X_{ki})} \exp(\beta_k' Z_{kl}(X_{ki})) \right) \tag{3.4}$$

and the first derivatives of $\log L$ are

$$\begin{aligned} U(\beta_{ku}) &= \frac{\partial \log L(\beta_k)}{\partial \beta_{ku}} \\ &= \sum_{i=1}^{n} \delta_{ki} Z_{kiu}(X_{ki}) - \sum_{i=1}^{n} \delta_{ki} \left[\frac{\sum_{l \in R(X_{ki})} Z_{klu}(X_{ki}) \exp(\beta_k' Z_{kl}(X_{ki}))}{\sum_{l \in R(X_{ki})} \exp(\beta_k' Z_{kl}(X_{ki}))} \right] \\ &= \sum_{i=1}^{n} \delta_{ki} Z_{kiu}(X_{ki}) - \sum_{i=1}^{n} \delta_{ki} \left[\frac{\sum_{l=1}^{n} Y_{kl}(X_{ki}) Z_{klu}(X_{ki}) \exp(\beta_k' Z_{kl}(X_{ki}))}{\sum_{l=1}^{n} Y_{kl}(X_{ki}) \exp(\beta_k' Z_{kl}(X_{ki}))} \right] \end{aligned} \quad (3.5)$$

where $Y_{kl}(t) = I(X_{kl} \geq t)$ and more specifically, $Y_{kl}(X_{ki}) = 1$, if $X_{kl} \geq X_{ki}$ and 0, otherwise.

Substituting $G_{kl}(\beta_k) = Y_{kl}(X_{ki}) \exp(\beta_k' Z_{kl}(X_{ki}))$ in (3.5) the score function reduces to

$$U(\beta_{ku}) = \sum_{i=1}^{n} \delta_{ki} Z_{kiu}(X_{ki}) - \sum_{i=1}^{n} \delta_{ki} \left[\frac{\sum_{l=1}^{n} Z_{klu}(X_{ki}) G_{kl}(\beta_k)}{\sum_{l=1}^{n} G_{kl}(\beta_k)} \right] \quad (3.6)$$

Wei et al. (1989) obtained the maximum partial likelihood estimates $\hat{\beta}_k$ of β_k as the solution of the likelihood equation

$$U(\beta_{ku}) = \frac{\partial \log L(\beta_k)}{\partial \beta_{ku}} = 0$$

The estimator $\hat{\beta}_k$ is consistent for β_k if the model (3.1) is correctly specified.

The estimator $\hat{\beta}_k$'s are generally correlated. As shown by Wei et al. (1989), for large n, $(\hat{\beta}_1', \cdots, \hat{\beta}_K')'$ is approximately normal with mean $(\beta_1', \cdots, \beta_K')'$ and covariance matrix Q, say. For large n, the covariance matrix of $(\hat{\beta}_1', \cdots, \hat{\beta}_K')'$ can be estimated by

$$\hat{Q} = n^{-1} \begin{bmatrix} \hat{D}_{11}(\hat{\beta}_1, \hat{\beta}_1) & \cdots & \hat{D}_{1K}(\hat{\beta}_1, \hat{\beta}_K) \\ \vdots & \ddots & \vdots \\ \hat{D}_{K1}(\hat{\beta}_K, \hat{\beta}_1) & \cdots & \hat{D}_{KK}(\hat{\beta}_K, \hat{\beta}_K) \end{bmatrix} \quad (3.7)$$

where $\hat{D}_{kl}(\hat{\beta}_k, \hat{\beta}_l)$, $(k, l = 1, \cdots, K)$ is the $p \times p$ estimated asymptotic covariance matrix between $n^{1/2}(\hat{\beta}_k - \beta_k)$ and $n^{1/2}(\hat{\beta}_l - \beta_l)$ which is given by

$$\hat{D}_{kl}(\hat{\beta}_k, \hat{\beta}_l) = \hat{A}_k^{-1}(\hat{\beta}_k) \hat{B}_{kl}(\hat{\beta}_k, \hat{\beta}_l) \hat{A}_l^{-1}(\hat{\beta}_l)$$

For additional details see Wei et al. (1989). By theorem 4.2 of Anderson and Gill (1982), the matrix $\hat{A}_k(\hat{\beta}_k)$ is the consistent estimate of $A_k(\beta_k)$ and is given by

$$\hat{A}_k(\hat{\beta}_k) = n^{-1}\sum_{j=1}^{n}\delta_{kj}\left[\frac{\sum_{i=1}^{n}Y_{ki}(X_{kj})Z_{ki}(X_{kj})^{\otimes 2}\exp(\hat{\beta}_k'Z_{ki}(X_{kj}))}{\sum_{i=1}^{n}Y_{ki}(X_{kj})\exp(\hat{\beta}_k'Z_{ki}(X_{kj}))}\right.$$
$$\left.-\left(\frac{\sum_{i=1}^{n}Y_{ki}(X_{kj})Z_{ki}(X_{kj})\exp(\hat{\beta}_k'Z_{ki}(X_{kj}))}{\sum_{i=1}^{n}Y_{ki}(X_{kj})\exp(\hat{\beta}_k'Z_{ki}(X_{kj}))}\right)^{\otimes 2}\right]. \qquad (3.8)$$

Here, $a^{\otimes 2}$ denotes the matrix aa' for a colum vector a.

In (3.7),

$$\hat{B}_{kl}(\hat{\beta}_k,\hat{\beta}_l) = n^{-1}\sum_{j=1}^{n}W_{kj}(\hat{\beta}_k)W_{lj}'(\hat{\beta}_l)$$

where

$$W_{kj}(\beta_k) = \delta_{kj}\left[Z_{kj}(X_{kj}) - \frac{S_k^{(1)}(\beta_k;X_{kj})}{S_k^{(0)}(\beta_k;X_{kj})}\right]$$
$$-\sum_{m=1}^{n}\frac{\delta_{km}Y_{kj}(X_{km})\exp(\beta_k'Z_{kj}(X_{km}))}{nS_k^{(1)}(\beta_k;X_{km})}$$
$$\times\left[Z_{kj}(X_{km}) - \frac{S_k^{(1)}(\beta_k;X_{km})}{S_k^{(0)}(\beta_k;X_{km})}\right]$$

and

$$S_k^{(1)}(\beta_k;t) = n^{-1}\sum_{i=1}^{n}Y_{ki}(t)Z_{ki}(t)\exp(\beta_k'Z_{ki}(t))$$
$$S_k^{0}(\beta_k;t) = n^{-1}\sum_{i=1}^{n}Y_{ki}(t)\exp(\beta_k'Z_{ki}(t)).$$

The matrix $\hat{Q}$ provides a basis for simultaneous inferences about the β_k's. For example, suppose that one is interested in the effects of a particular type of covariate on the K event times. More specifically, one may be interested to test jointly the hypothesis

$$H_0 : \beta_{k1} \geq 0, \quad k = 1,\cdots,K \qquad (3.9)$$
$$H_A : \beta_{k1} < 0$$

where β_{k1} is the treatment effect for the kth group. To test this hypothesis Wei et al. (1989) used the Wald type test statistic

$$W = (\hat{\beta}_{11}, \cdots, \hat{\beta}_{k1})\hat{\psi}^{-1}_{K\times K}(\hat{\beta}_{11}, \cdots, \hat{\beta}_{k1})' \tag{3.10}$$

which has approximately χ^2 distribution with K degrees of freedom. Here in (3.10), $\hat{\psi}_{K\times K}$ is $K \times K$ estimated covariance matrix of $\hat{\beta}_{k1}$, obtained by partitioning the $pK \times pK$ estimated covariance matrix $\hat{Q}$, of $\hat{\beta}$. Further, based on the assumption that treatment effects of all groups are same, that is, $\beta_{11} = \cdots = \beta_{k1} = \beta^*_{.1}$, Wei et al. (1989) estimate $\beta^*_{.1}$ by using a linear combination of the $\hat{\beta}_{k1}$'s, that is, $\sum_{k=1}^{K} W_k\hat{\beta}_{k1}$ with $\sum_{k=1}^{K} W_k = 1$. By Wei and Johnson (1985), the estimator $\hat{\beta}_1 = (\hat{\beta}_{11}, \cdots, \hat{\beta}_{K1})'$ with weight

$$W = (W_1, \cdots, W_K)' = (e'\hat{\psi}^{-1}_{K\times K}e)^{-1}\hat{\psi}^{-1}_{K\times K}e$$

where $e = (1, \cdots, 1)'$, has the smallest asymptotic variance among all of the linear estimators. We use this linear estimation approach to estimate $\beta^*_{.1}$ for modified data as well as for the original data.

We remark here that in contrast to Wei et al. (1989), one may test the hypothesis (3.9) by using the likelihood ratio method, and the combined estimate of $\beta^*_{.1}$ can be obtained by using the restricted maximum likelihood approach. We follow these procedures in the next chapter for the modified data set as well as for the original blood cancer data.

3.4 Combined Treatment Effects: Restricted Partial Likelihood Estimation (RPLE) Method

In this section, unlike Wei et al. (1989), we discuss a method to estimate the combined treatment effects for multivariate incomplete failure time data by using the restricted partial likelihood estimation (RPLE) method. First, we estimate the combined treatment effect considering all other covariates as nuisance and second under the presence of treatment only (other covariates are ignored).

3.4.1 When Other Covariates are Nuisance

For this case we impose the restrictions

$$\beta_{11} = \beta_{21} = \cdots = \beta_{K1} = \beta^*_{\cdot 1} \tag{3.11}$$

to the partial likelihood function (3.2) and estimate the combined treatment effects $\hat{\beta}^*_{\cdot 1}$. Under the restriction the likelihood function (3.2) now reduces to

$$L(\beta^*_{\cdot 1}, \beta^*_1, \cdots, \beta^*_K) = \prod_{k=1}^{K} \prod_{i=1}^{n} \left[\frac{\exp(\beta^*_{\cdot 1} Z_{ki1}(X_{ki}) + \beta^{*\prime}_k Z^*_{ki}(X_{ki}))}{\sum_{l \in R_k(X_{ki})} \exp(\beta^*_{\cdot 1} Z_{kil} + \beta^{*\prime}_k Z^*_{kl}(X_{ki}))} \right]^{\delta_{ki}} \tag{3.12}$$

where $Z^*_{ki} = (Z_{ki2}, \cdots, Z_{kip})'$ is the covariate vector corresponding to the individual observed to die at X_{ki} and $\beta^*_k = (\beta_{k2}, \cdots, \beta_{kp})'$.

Letting $\beta^* = (\beta^*_{\cdot 1}, \beta^{*\prime}_1, \cdots, \beta^{*\prime}_K)'$ the log likelihood arising from (3.12) is given by

$$\begin{aligned} \log L(\beta^*) &= \sum_{k=1}^{K} \sum_{i=1}^{n} \delta_{ki} (\beta^*_{\cdot 1} Z_{ki1}(X_{ki}) + \beta^{*\prime}_k Z^*_{ki}(X_{ki})) \\ &\quad - \sum_{k=1}^{K} \sum_{i=1}^{n} \delta_{ki} \log \left(\sum_{l \in R(X_{ki})} \exp(\beta^*_{\cdot 1} Z_{ki1}(X_{ki}) + \beta^{*\prime}_k Z^*_{kl}(X_{ki})) \right) \end{aligned}$$

and the score functions for estimating $\beta^*_{\cdot 1}$ and $\beta_{k2}, \cdots, \beta_{kp}$ are given by,

$$U_1(\beta^*) = \frac{\partial \log L(\beta^*)}{\partial \beta^*_{\cdot 1}}$$

$$= \sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}Z_{ki1}(X_{ki}) - \sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}\left[\frac{\sum_{l=1}^{n}Z_{kl1}(X_{ki})G_{kl}^{*}(\beta^{*})}{\sum_{l=1}^{n}G_{kl}^{*}(\beta^{*})}\right] \quad (3.13)$$

and

$$\begin{aligned} U_2(\beta^{*}) &= \frac{\partial \log L(\beta^{*})}{\partial \beta_{ku}} \\ &= \sum_{i=1}^{n}\delta_{ki}Z_{kiu}(X_{ki}) - \sum_{i=1}^{n}\delta_{ki}\left[\frac{\sum_{l=1}^{n}Z_{klu}(X_{ki})G_{kl}^{*}(\beta^{*})}{\sum_{l=1}^{n}G_{kl}^{*}(\beta*)}\right] \end{aligned} \quad (3.14)$$

respectively. where, $u = 2, \cdots, p$, and $G_{kl}^{*}(\beta^{*}) = Y_{kl}(X_{ki})\exp(\beta_{.1}^{*}Z_{ki1}(X_{ki}) + \beta_{k}^{*\prime}Z_{kl}^{*}(X_{ki}))$.

Now the maximum partial likelihood estimators $\hat{\beta}_{.1}^{*}$ and $\hat{\beta}_{k}^{*} = (\hat{\beta}_{k2}, \cdots, \hat{\beta}_{kp})'$ are obtained by using Newton Raphson method as the solution of the equations

$$\frac{\partial \log L(\beta^{*})}{\partial \beta_{.1}^{*}} = 0, \quad \text{and} \quad \frac{\partial \log L(\beta^{*})}{\partial \beta_{ku}} = 0$$

It then follows from Wei et al.(1989) that, for large n, $(\hat{\beta}_{.1}^{*}, \hat{\beta}_{1}^{*\prime}, \cdots, \hat{\beta}_{K}^{*\prime})'$, where, $\hat{\beta}_{k}^{*} = (\hat{\beta}_{k2}, \cdots, \hat{\beta}_{kp})'$ is approximately normal with mean $(\beta_{.1}^{*}, \beta_{1}', \cdots, \beta_{K}')'$ and covariance matrix Q^{*}, (say), where Q^{*} can be estimated following the procedure that was used to compute $\hat{Q}$ in (3.7). Now the new estimated variance covariance matrix $\hat{Q}^{*}$ of Q^{*} is given by

$$\hat{Q}^{*} = n^{-1}\begin{bmatrix} \hat{D}_{..}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{.1}^{*}) & \hat{D}_{.1}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{1}^{*}) & \cdots & \hat{D}_{.K}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{K}^{*}) \\ \hat{D}_{1.}(\hat{\beta}_{1}^{*}, \hat{\beta}_{.1}^{*}) & \hat{D}_{11}(\hat{\beta}_{1}^{*}, \hat{\beta}_{1}^{*}) & \cdots & \hat{D}_{1K}(\hat{\beta}_{1}^{*}, \hat{\beta}_{K}^{*}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{D}_{K.}(\hat{\beta}_{K}^{*}, \hat{\beta}_{.1}^{*}) & \hat{D}_{K1}(\hat{\beta}_{K}^{*}, \hat{\beta}_{1}^{*}) & \cdots & \hat{D}_{KK}(\hat{\beta}_{K}^{*}, \hat{\beta}_{K}^{*}) \end{bmatrix}$$

where $\hat{D}_{.k}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{k}^{*})$, $(k = 1, \cdots, K)$ is the $1 \times (p-1)$ estimated asymptotic covariance vector between $n^{1/2}(\hat{\beta}_{.1}^{*} - \beta_{.1}^{*})$ and $n^{1/2}(\hat{\beta}_{k}^{*} - \beta_{k}^{*})$ which is given by

$$\hat{D}_{.k}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{k}^{*}) = \hat{A}_{.}^{-1}(\hat{\beta}_{.1}^{*})\hat{B}_{.k}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{k}^{*})\hat{A}_{k}^{-1}(\hat{\beta}_{k}^{*}).$$

Similarly, $\hat{D}_{kk'}(\hat{\beta}^*_k, \hat{\beta}^*_{k'})$, $(k, k' = 1, \cdots, K)$ is the $(p-1)\times(p-1)$ estimated asymptotic covariance matrix between $n^{1/2}(\hat{\beta}^*_k - \beta^*_k)$ and $n^{1/2}(\hat{\beta}^*_{k'} - \beta^*_{k'})$ which is given by

$$\hat{D}_{kk'}(\hat{\beta}^*_k, \hat{\beta}^*_{k'}) = \hat{A}^{-1}_k(\hat{\beta}^*_k)\hat{B}_{kk'}(\hat{\beta}^*_k, \hat{\beta}^*_{k'})\hat{A}^{-1}_k(\hat{\beta}^*_{k'}).$$

For additional details see Wei et al. (1989). The variance of the combined estimate $\hat{\beta}^*_{.1}$ is obtained as

$$\begin{aligned}\mathrm{Var}(\hat{\beta}^*_{.1}) &= n^{-1}\hat{D}_{..}(\hat{\beta}^*_{.1}, \hat{\beta}^*_{.1}) \\ &= n^{-1}\hat{A}^{-1}_{.}(\hat{\beta}^*_{.1})\hat{B}_{..}(\hat{\beta}^*_{.1}, \hat{\beta}^*_{.1})\hat{A}^{-1}_{.}(\hat{\beta}^*_{.1})\end{aligned}$$

where,

$$\begin{aligned}\hat{A}_{.}(\hat{\beta}^*_{.1}) &= n^{-1}\sum_{k=1}^{K}\sum_{j=1}^{n}\delta_{kj}\left[\frac{\sum_{i=1}^{n} Z_{ki1}(X_{kj})^2 G^*_{ki}(\beta^*)}{\sum_{i=1}^{n} G^*_{ki}(\beta^*)}\right. \\ &\quad \left. - \left(\frac{\sum_{i=1}^{n} Z_{ki1}(X_{kj}) G^*_{ki}(\beta^*)}{\sum_{i=1}^{n} G^*_{ki}(\beta^*)}\right)^{\otimes 2}\right]\end{aligned}$$

with $G^*_{ki}(\beta^*) = Y_{ki}(X_{kj})\exp(\beta^*_{.1} Z_{ki1}(X_{kj}) + \beta^{*\prime}_k Z^*_{ki}(X_{kj}))$ and

$$\hat{B}_{..}(\hat{\beta}^*_{.1}, \hat{\beta}^*_{.1}) = n^{-1}\sum_{k=1}^{K}\sum_{j=1}^{n} W^{\otimes 2}_{.j1}(\hat{\beta}^*_{.1})$$

where

$$\begin{aligned}W_{.j1}(\beta^*_{.1}) &= \delta_{kj}\left[Z_{kj1}(X_{kj}) - \frac{S^{*(1)}_k(\beta^*_{.1}; X_{kj})}{S^{*(0)}_k(\beta^*_{.1}; X_{kj})}\right] \\ &\quad - \sum_{m=1}^{n}\frac{\delta_{km} Y_{kj}(X_{km})\exp(\beta^*_{.1} Z_{kj1}(X_{km}) + \beta^{*\prime}_k Z^*_{kj}(X_{km}))}{n S^{*(1)}_k(\beta^*_{.1}; X_{km})} \\ &\quad \times\left[Z_{kj1}(X_{km}) - \frac{S^{*(1)}_k(\beta^*_{.1}; X_{km})}{S^{*(0)}_k(\beta^*_{.1}; X_{km})}\right]\end{aligned}$$

and

$$\begin{aligned}S^{*(1)}_k(\beta^*_{.1}; t) &= n^{-1}\sum_{i=1}^{n} Y_{ki}(t) Z_{ki1}(t)\exp\left(\beta^*_{.1} Z_{ki1}(t) + \beta^{*\prime}_k Z^*_{ki}(t)\right) \\ S^{*(0)}_k(\beta^*_{.1}; t) &= n^{-1}\sum_{i=1}^{n} Y_{ki}(t)\exp\left(\beta^*_{.1} Z_{ki1}(t) + \beta^{*\prime}_k Z^*_{ki}(t)\right).\end{aligned}$$

3.4.2 When Other Covariates are Ignored

As in the above section applying the same restriction (3.11), the partial likelihood function (3.2) reduces to

$$L(\beta_{.1}^{*}) = \prod_{k=1}^{K}\prod_{i=1}^{n}\left[\frac{\exp(\beta_{.1}^{*}Z_{ki1}(X_{ki})}{\sum_{l\in R_k(X_{(ki)})}\exp(\beta_{.1}^{*}Z_{kl1}(X_{ki})}\right]^{\delta_{ki}} \tag{3.15}$$

The log likelihood function from (3.15) is given by

$$\begin{aligned}\log L(\beta_{.1}^{*}) &= \sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}(\beta_{.1}^{*}Z_{ki1}(X_{ki})) \\ &\quad -\sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}\log\left(\sum_{l\in R(X_{ki})}\exp(\beta_{.1}^{*}Z_{kl1}(X_{ki}))\right)\end{aligned}$$

and the score function for estimating $\beta_{.1}^{*}$ is given by,

$$\begin{aligned}U(\beta_{.1}^{*}) &= \frac{\partial \log L(\beta_{.1}^{*})}{\partial \beta_{.1}^{*}} \\ &= \sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}Z_{ki1}(X_{ki}) - \sum_{k=1}^{K}\sum_{i=1}^{n}\delta_{ki}\left[\frac{\sum_{l=1}^{n}Z_{kl1}(X_{ki})G_{kl}^{**}(\beta_{.1}^{*})}{\sum_{l=1}^{n}G_{kl}^{**}(\beta_{.1}^{*})}\right]\end{aligned} \tag{3.16}$$

where, $G_{kl}^{**}(\beta_{.1}^{*}) = Y_{kl}(X_{ki})\exp(\beta_{.1}^{*}Z_{kl1}(X_{ki}))$

The maximum partial likelihood estimator $\hat{\beta}_{.1}^{*}$ is obtained as the solution of the equation

$$\frac{\partial \log L(\beta_{.1}^{*})}{\partial \beta_{.1}^{*}} = 0.$$

Now following (3.7) the estimated variance covariance matrix $\hat{Q}^{**}$ of Q^{**} is given by

$$\hat{Q}^{**} = n^{-1}\hat{D}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{.1}^{*})$$

and the variance of the combined estimate $\hat{\beta}_{.1}^{*}$ is obtained as

$$\begin{aligned}\mathrm{Var}(\hat{\beta}_{.1}^{*}) &= n^{-1}\hat{D}_{..}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{.1}^{*}) \\ &= n^{-1}\hat{A}_{..}^{-1}(\hat{\beta}_{.1}^{*})\hat{B}_{..}(\hat{\beta}_{.1}^{*}, \hat{\beta}_{.1}^{*})\hat{A}_{..}^{-1}(\hat{\beta}_{.1}^{*})\end{aligned}$$

where,

$$\hat{A}_{\cdot}(\hat{\beta}^{*}_{\cdot 1}) = n^{-1}\sum_{k=1}^{K}\sum_{j=1}^{n}\delta_{kj}\left[\frac{\sum_{i=1}^{n} Z_{ki1}(X_{kj})^2 G^{**}_{ki}(\beta^{*}_{\cdot 1})}{\sum_{i=1}^{n} G^{**}_{ki}(\beta^{*}_{\cdot 1})} - \left(\frac{\sum_{i=1}^{n} Z_{ki1}(X_{kj}) G^{**}_{ki}(\beta^{*}_{\cdot 1})}{\sum_{i=1}^{n} G^{**}_{ki}(\beta.1^{*})}\right)^{\otimes 2}\right]$$

with $G^{**}_{ki}(\beta^{*}_{\cdot 1}) = Y_{ki}(X_{kj})\exp(\beta^{*}_{\cdot 1} Z_{ki1}(X_{kj}))$ and

$$\hat{B}_{\cdot\cdot}(\hat{\beta}^{*}_{\cdot 1}, \hat{\beta}^{*}_{\cdot 1}) = n^{-1}\sum_{k=1}^{K}\sum_{j=1}^{n} W^{2}_{\cdot j1}(\hat{\beta}^{*}_{\cdot 1})$$

where

$$W_{\cdot j1}(\beta^{*}_{\cdot 1}) = \delta_{kj}\left[Z_{kj1}(X_{kj}) - \frac{S^{*(1)}_{k}(\beta^{*}_{\cdot 1}; X_{kj})}{S^{**(0)}_{k}(\beta^{*}_{\cdot 1}; X_{kj})}\right] - \sum_{m=1}^{n}\frac{\delta_{km} Y_{kj}(X_{km})\exp(\beta^{*}_{\cdot 1} Z_{kj1}(X_{km}))}{n S^{**(1)}_{k}(\beta^{*}_{\cdot 1}; X_{km})} \times \left[Z_{kj1}(X_{km}) - \frac{S^{**(1)}_{k}(\beta^{*}_{\cdot 1}; X_{km})}{S^{**(0)}_{k}(\beta^{*}_{\cdot 1}; X_{km})}\right]$$

and

$$S^{**(1)}_{k}(\beta^{*}_{\cdot 1}; t) = n^{-1}\sum_{i=1}^{n} Y_{ki}(t) Z_{ki1}(t)\exp(\beta^{*}_{\cdot 1} Z_{ki1}(t))$$

$$S^{**(0)}_{k}(\beta^{*}_{\cdot 1}; t) = n^{-1}\sum_{i=1}^{n} Y_{ki}(t)\exp(\beta^{*}_{\cdot 1} Z_{ki1}(t)).$$

3.5 Ad Hoc Modification of Bladder Cancer Data

The multivariate failure time for each blood cancer patient discussed in Wei et al. (1989) are not independent for the K ($K = 4$) different groups. Wei et al. (1989), however, used the univariate (marginal) partial likelihood approach to obtain the regression estimates as well as the estimate of their covariance matrix. These regression estimates may be treated as be interpretable like the least square regression estimates

obtained from a correlated linear regression model, and the estimate of the covariance matrix of the regression estimates may be treated as interpretable as the estimate of the correct covariance matrix of the least square regression estimates. This type of regression estimate naturally will not be optimal. With this in view, we now try to obtain a group-wise independent data set as follows, which will be analysed in the next chapter by using the univariate partial likelihood approach.

The original blood cancer data consist of initial tumor number, initial tumor size, a followup time and four different recurrence times for each of the 86 patients. These recurrence times were recorded based on the repeated visits of the patients, and consequently they are correlated. The analysis of such multivariate correlated data is quite complex. Since in every visit the recurred tumors are removed and the patient is kept under the therapy (treatment) until the next visit, it seems appropriate to consider the time gap between two consecutive recurrences as an independent recurrence time for the patient involved, provided the followup time is adjusted (modified) accordingly. This we do as in the following.

Let r^*_{ki}, for $k = 1, \cdots, K$, be the original tumor recurrence time of the ith patient corresponding to his (her) kth visit. Suppose r_{ki} be the kth modified recurrence (visit) time of the ith patient. We now define r_{ki} as

$$\begin{aligned} r_{1i} &= r^*_{1i} \\ \text{and} \quad r_{ki} &= r^*_{ki} - r^*_{(k-1)i}, \quad \text{for} \quad k \geq 2. \end{aligned}$$

Further, let t_i $(i = 1, \cdots, n)$ denote the original followup time for the ith patient. We now consider K different followup times for K distinct visits. The followup times for the first two visits are given by

$$f_{1i} = t_i$$

$$\text{and} \quad f_{2i} = t_i - r_{1i}.$$

Next, the followup time with respect to the kth visit for $k \geq 3$ of the ith patient is defined as follows:

$$f_{ki} = f_{(k-1)i} - r_{(k-1)i} \quad \text{for} \quad k \geq 3.$$

To demonstrate how these modified tumor recurrence and followup times are computed, we consider here. for example, the 15th patient in Table 7.1. From the Table 7.1, we have $t_{15} = 24$, $r^*_{1,15} = 7$, $r^*_{2,15} = 10$, $r^*_{3,15} = 16$ and $r^*_{4,15} = 24$. Now the modified recurrence times for different groups are

$$\begin{aligned} r_{1,15} &= r^*_{1,15} = 7 \\ r_{2,15} &= r^*_{2,15} - r^*_{1.15} = 3 \\ r_{3,15} &= r^*_{3,15} - r^*_{2,15} = 6 \\ r_{4,15} &= r^*_{4.15} - r^*_{3.15} = 8. \end{aligned}$$

Also the modified followup times for different failure groups are obtained as follows:

$$\begin{aligned} f_{1,15} &= t_{1,15} = 24 \\ f_{2,15} &= t_{1,15} - r_{1,15} = 17 \\ f_{3,15} &= f_{2,15} - r_{2,15} = 14 \\ f_{4,15} &= f_{3,15} - r_{3,15} = 8. \end{aligned}$$

This data for modified followup and recurrence times is shown in Table 3.1. Also for the sake of completeness, we exhibit the original blood cancer data in the appendix.

Table 3.1: Modified Tumor Recurrence Data for Patients With Bladder Cancer

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time				Recurrence time			
			f_{1i}	f_{2i}	f_{3i}	f_{4i}	r_{1i}	r_{2i}	r_{3i}	r_{4i}
1	1	1	0	0	0	0	0	0	0	0
1	1	3	1	1	1	1	0	0	0	0
1	2	1	4	4	4	4	0	0	0	0
1	1	1	7	7	7	7	0	0	0	0
1	5	1	10	10	10	10	0	0	0	0
1	4	1	10	4	4	4	6	0	0	0
1	1	1	14	14	14	14	0	0	0	0
1	1	1	18	18	18	18	0	0	0	0
1	1	3	18	13	13	13	5	0	0	0
1	1	1	18	6	2	2	12	4	0	0
1	3	3	23	23	23	23	0	0	0	0
1	1	3	23	13	13	13	10	5	0	0
1	1	1	23	20	7	0	3	13	7	0
1	3	1	23	20	14	2	3	6	12	0
1	2	3	24	17	14	8	7	3	6	8
1	1	1	25	22	10	0	3	12	10	0

table 3.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time				Recurrence time			
			f_{1i}	f_{2i}	f_{3i}	f_{4i}	r_{1i}	r_{2i}	r_{3i}	r_{4i}
1	1	2	26	26	26	26	0	0	0	0
1	8	1	26	25	25	25	1	0	0	0
1	1	4	26	24	0	0	2	24	0	0
1	1	2	28	3	3	3	25	0	0	0
1	1	4	29	29	29	29	0	0	0	0
1	1	2	29	29	29	29	0	0	0	0
1	4	1	29	29	29	29	0	0	0	0
1	1	6	30	2	0	0	28	2	0	0
1	1	5	30	28	13	8	2	15	5	0
1	2	1	30	27	24	22	3	3	2	4
1	1	3	31	16	7	12	12	3	9	0
1	1	2	32	32	32	32	0	0	0	0
1	2	1	34	34	34	34	0	0	0	0
1	2	1	36	36	36	36	0	0	0	0
1	3	1	36	7	7	7	29	0	0	0
1	1	2	37	37	37	17	0	0	0	0
1	4	1	40	31	23	18	9	8	5	2
1	5	1	40	24	21	17	16	3	4	6

table 3.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time				Recurrence time			
			f_{1i}	f_{2i}	f_{3i}	f_{4i}	r_{1i}	r_{2i}	r_{3i}	r_{4i}
1	1	2	41	41	41	41	0	0	0	0
1	1	1	43	40	40	40	3	0	0	0
1	2	6	43	37	37	37	6	0	0	0
1	2	1	44	41	38	35	3	3	3	0
1	1	1	45	36	34	25	9	11	20	26
1	1	1	48	30	30	30	18	0	0	0
1	1	3	49	49	49	0	0	0	0	0
1	3	1	51	16	16	16	35	0	0	0
1	1	7	53	36	36	36	17	0	0	0
1	3	1	53	50	37	7	3	12	31	5
1	1	1	59	59	59	59	0	0	0	0
1	3	2	61	59	46	37	2	13	9	6
1	1	3	64	59	50	45	5	9	5	8
1	2	3	64	62	54	50	2	6	4	1
2	1	3	1	1	1	1	0	0	0	0
2	1	1	1	1	1	1	0	0	0	0
2	8	1	5	0	0	0	5	0	0	0
2	1	2	9	9	9	9	0	0	0	0

table 3.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time				Recurrence time			
			f_{1i}	f_{2i}	f_{3i}	f_{4i}	r_{1i}	r_{2i}	r_{3i}	r_{4i}
2	1	1	10	10	10	10	0	0	0	0
2	1	1	13	13	13	13	0	0	0	0
2	2	6	14	11	11	11	3	0	0	0
2	5	3	17	16	14	12	1	2	2	2
2	5	1	18	18	18	18	0	0	0	0
2	1	3	18	1	1	1	17	0	0	0
2	5	1	19	17	17	17	2	0	0	0
2	1	1	21	4	2	2	17	2	0	0
2	1	1	22	22	22	22	0	0	0	0
2	1	3	25	25	25	25	0	0	0	0
2	1	5	25	25	25	25	0	0	0	0
2	1	1	25	25	25	25	0	0	0	0
2	1	1	26	20	12	13	6	6	1	0
2	1	1	27	21	21	21	6	0	0	0
2	2	1	29	27	27	27	2	0	0	0
2	8	3	36	10	1	1	26	9	0	0
2	1	1	38	38	38	38	0	0	0	0
2	1	1	39	17	16	12	22	1	4	5

table 3.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time				Recurrence time			
			f_{1i}	f_{2i}	f_{3i}	f_{4i}	r_{1i}	r_{2i}	r_{3i}	r_{4i}
2	6	1	39	35	23	16	4	12	7	4
2	3	1	40	14	16	11	24	2	3	11
2	3	2	41	41	41	41	0	0	0	0
2	1	1	41	41	41	41	0	0	0	0
2	1	1	43	42	16	16	1	26	0	0
2	1	1	44	44	44	44	0	0	0	0
2	6	1	44	42	24	21	2	18	3	4
2	1	2	45	45	45	45	0	0	0	0
2	1	4	46	44	44	44	2	0	0	0
2	1	4	46	46	46	46	0	0	0	0
2	3	3	49	49	49	49	0	0	0	0
2	1	1	50	50	50	50	0	0	0	0
2	4	1	50	46	26	3	4	20	21	0
2	3	4	54	54	54	54	0	0	0	0
2	2	1	54	16	16	16	38	0	0	0
2	1	3	59	59	59	59	0	0	0	0

Note:

- Treatment group: 1, placebo: 2, thitepa, Follow-up times (f_{ki}) and recurrence times (r_{ki}) are measured in months. Initial tumor size is measured in centimeters. Initial tumor number of eight denotes eight or more initial tumors.

Chapter 4

Analysis of Modified Bladder Cancer Data

4.1 Risk Comparison

In this chapter we will analyse the modified as well as the original bladder cancer data by using the univariate partial likelihood method described in chapter 3, but the risk set for this likelihood computation will be generated based on the following three approaches. These approaches will differ from each other due to the difference in definitions of how censored observations are counted under each of the K (≥ 2) groups.

Approach One (A(I)): Wei et al. (1989) used this approach to detect the number of risk cases in each group. Here all the subjects (patients) censored and uncensored are considered in each recurrence group, that is, the total number of individuals in any recurrence group remains the same. The risk set for the kth recurrence group is defined as follows:

Let $R_k(t)$ denotes the risk set corresponding to the recurrence time t. The lth

$(l = 1, \cdots, n)$ individual will belong to the risk set $R_k(t)$ if

$$X_{kl} \geq t, \quad \text{or equivalently} \quad f_{kl} \geq t \quad \text{and} \quad \delta_{kl} = 0,$$

where f_{kl} and X_{kl} are the follow-up and recurrence time for the lth patient in the kth group respectively, and δ_{kl} is the censoring indicator for the lth individual. Further, let $Y_k(t)$ be the number of individuals in the risk set $R_k(t)$ corresponding to the recurrence time t. It then follows that

$$Y_k(t) = \#\{l : X_{kl} \geq t, \quad \text{or equivalently} \quad f_{kl} \geq t \quad \text{and} \quad \delta_{kl} = 0\}$$

Approach Two (A(II)): Following Prentice et al. (1981), in this approach the patients who were censored in the kth failure group are excluded from the $(k+1)$th failure group. Here the risk set is defined as:

$$X_{1l} \geq t, \quad \text{or equivalently} \quad f_{1l} \geq t \quad \text{and} \quad \delta_{1l} = 0, \quad \text{for} \quad k = 1$$

and for $k = 2, \cdots, K$

$$X_{kl} \geq t, \quad \text{or equivelently} \quad f_{kl} \geq t \quad \text{and} \quad \delta_{(k-1)l} = 0$$

yielding

$$Y_{1l}(t) = \#\{l : X_{1l} \geq t \quad \text{or equivalently} \quad f_{1l} \geq t \quad \text{and} \quad \delta_{1l} = 0, \quad \text{for} \quad k = 1$$

and

$$Y_{kl}(t) = \#\{l : X_{kl} \geq t, \quad \text{or equivalently} \quad f_{kl} \geq t \quad \text{and} \quad \delta_{(k-1)l} = 0\}$$

for $k = 2, \cdots, K$.

In the first recurrence group the risk sets are same for $A(I)$ and $A(II)$. For the other groups the risk sets in this approach are less than the risk sets in $A(I)$, that is, the number of patients who are at risk here is less than those in $A(I)$.

Approach Three (A(III)): This approach provides a further reduction in the risk sets. Here we do not consider at all the censored individuals in the risk sets. For the kth failure group the risk sets are given by:

$$X_{kl} \geq t$$

and in this case

$$Y_{kl}(t) = \#\{l : X_{kl} \geq t\}$$

Now using the above three approaches we have computed $Y_{kl}(t)$, the number of individuals under risk for all of the recurrence groups. The risk sets constructed under the above approaches for original and modified data sets are shown in figures 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, and 4.8, for four different recurrence groups respectively. Two figures for each of the four groups, for examples, 4.1 and 4.2 for the first group, show the risk set for the original and modified blood cancer data.

Since in the first recurrence group the risk sets are equal for the modified and original blood cancer data, there is no difference between figure 4.1 and 4.2. When approaches are compared to each other for any data set, original or modified, the risk sets are always smaller in $A(II)$ as compared to $A(I)$, except for the first recurrence group, where they are same. The risk sets of $A(III)$ are, however, always smaller than those of $A(II)$.

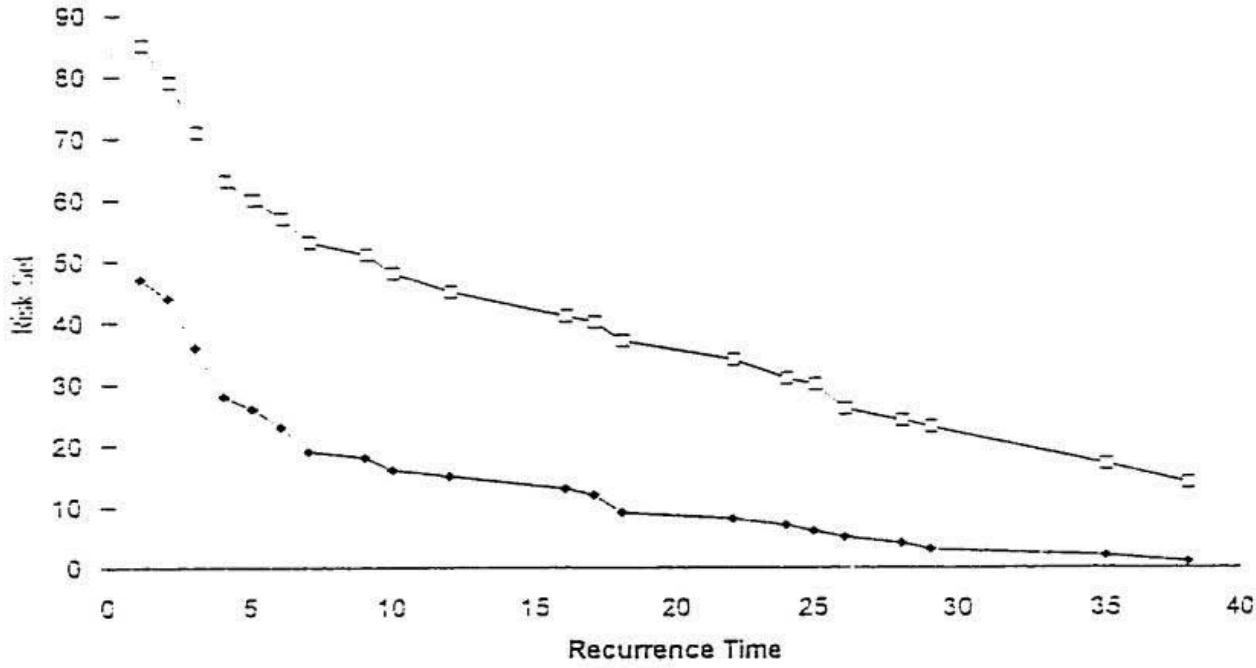

Figure 4.1: Risk Comparision of Original Tumor Recurrence Data for the First Recurrence Group Based on $A(I)(\square)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

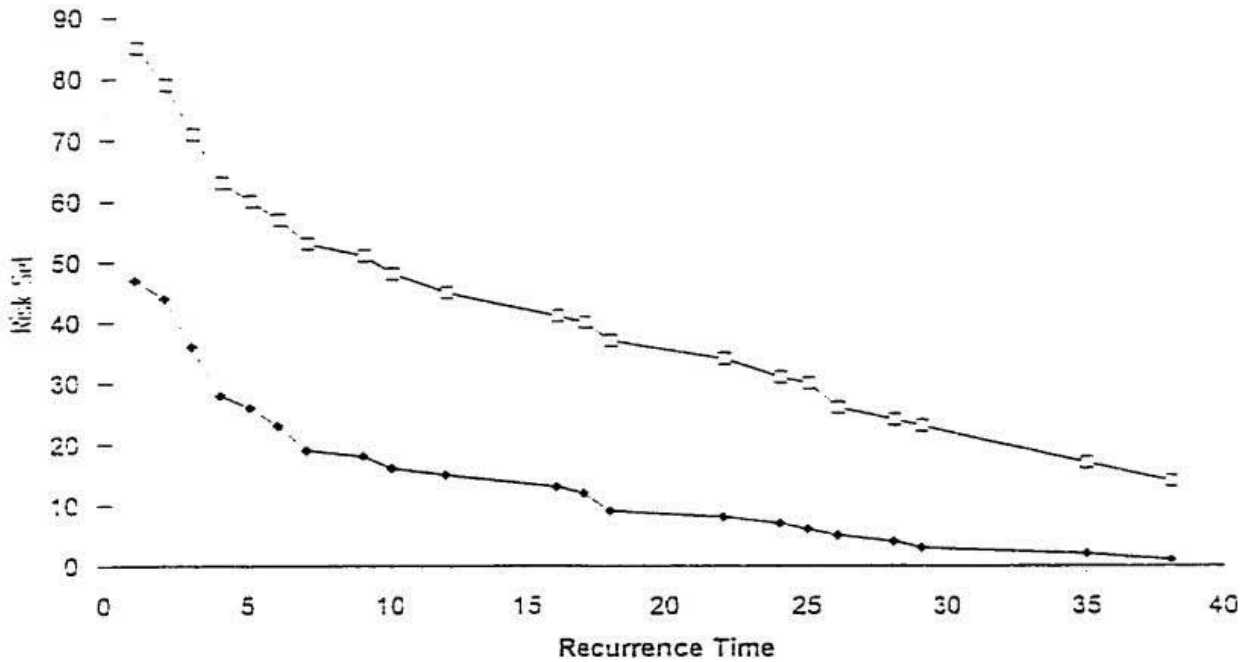

Figure 4.2: Risk Comparision of Modified Tumor Recurrence Data for the First Recurrence Group Based on $A(I)(\square)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

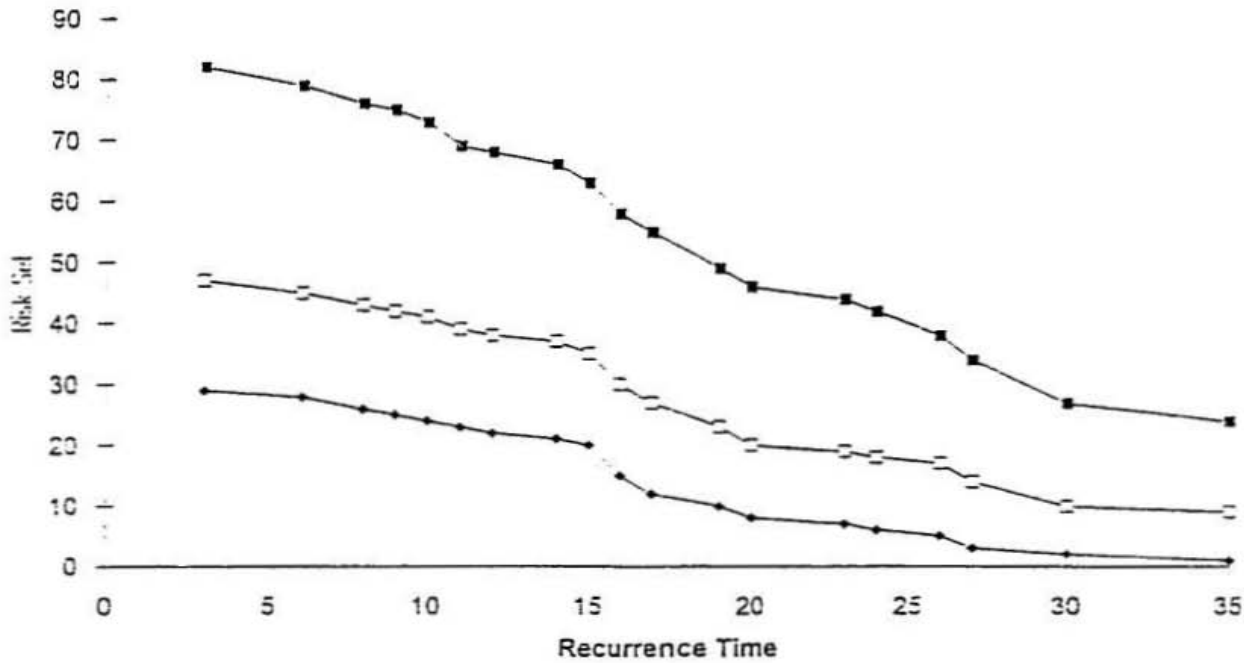

Figure 4.3: Risk Comparision of Original Tumor Recurrence Data for the Second Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

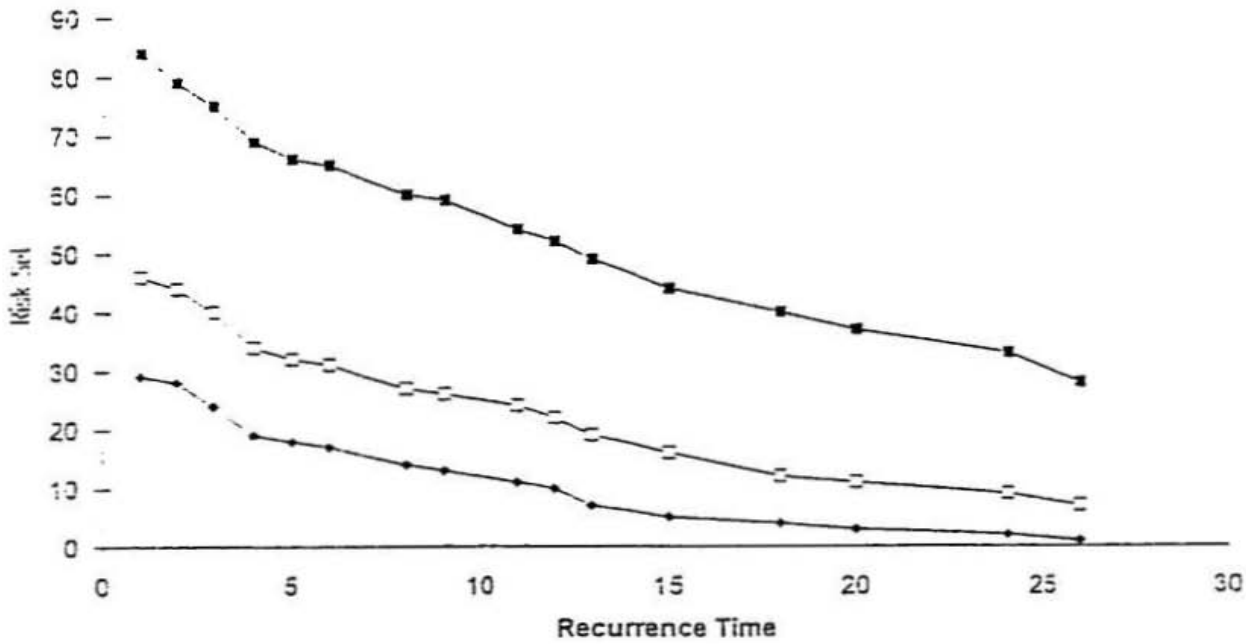

Figure 4.4: Risk Comparision of Modified Tumor Recurrence Data for the Second Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

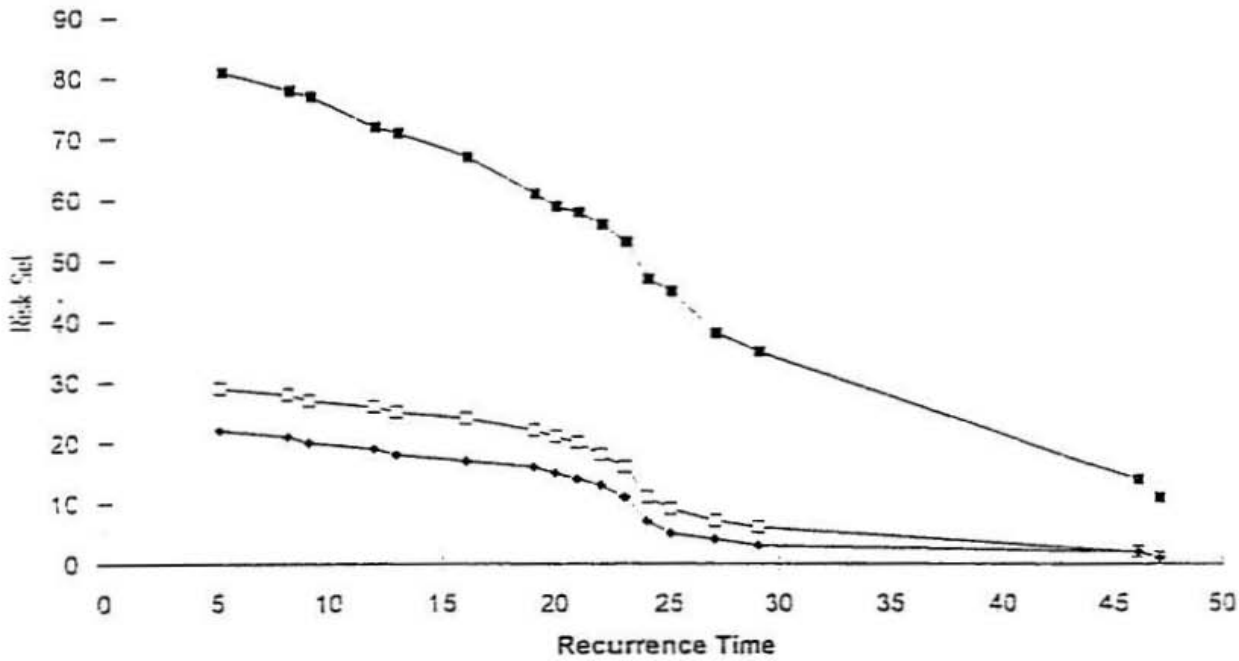

Figure 4.5: Risk Comparision of Original Tumor Recurrence Data for the Third Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

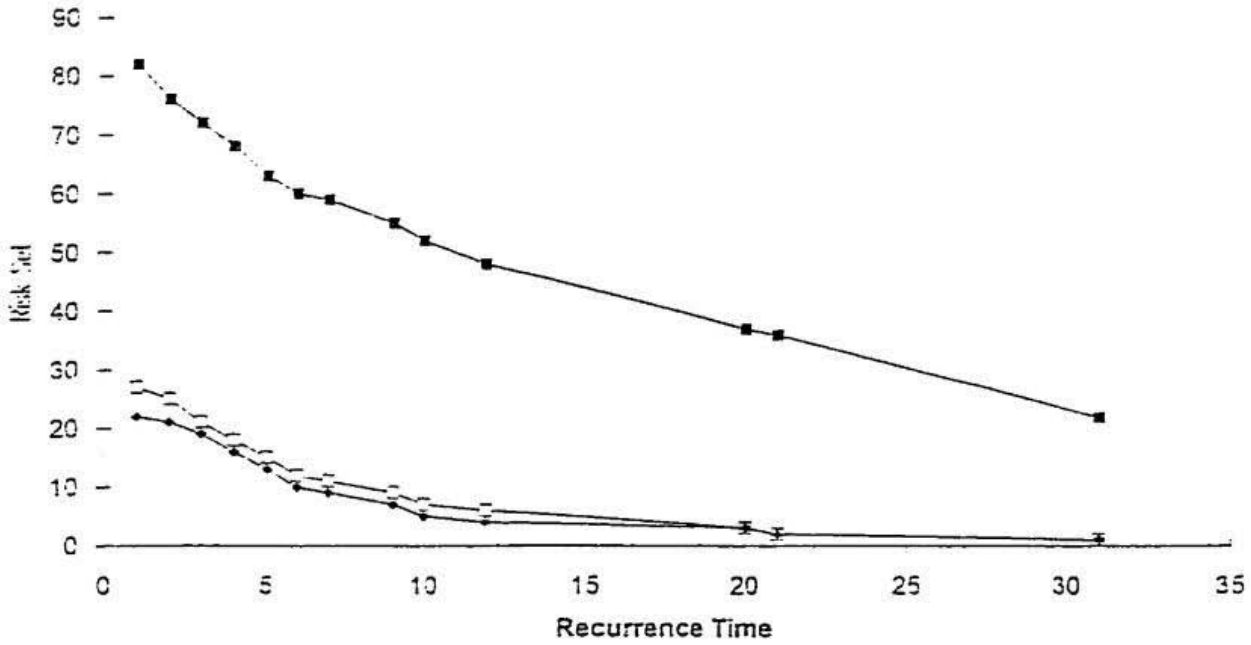

Figure 4.6: Risk Comparision of Modified Tumor Recurrence Data for the Third Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

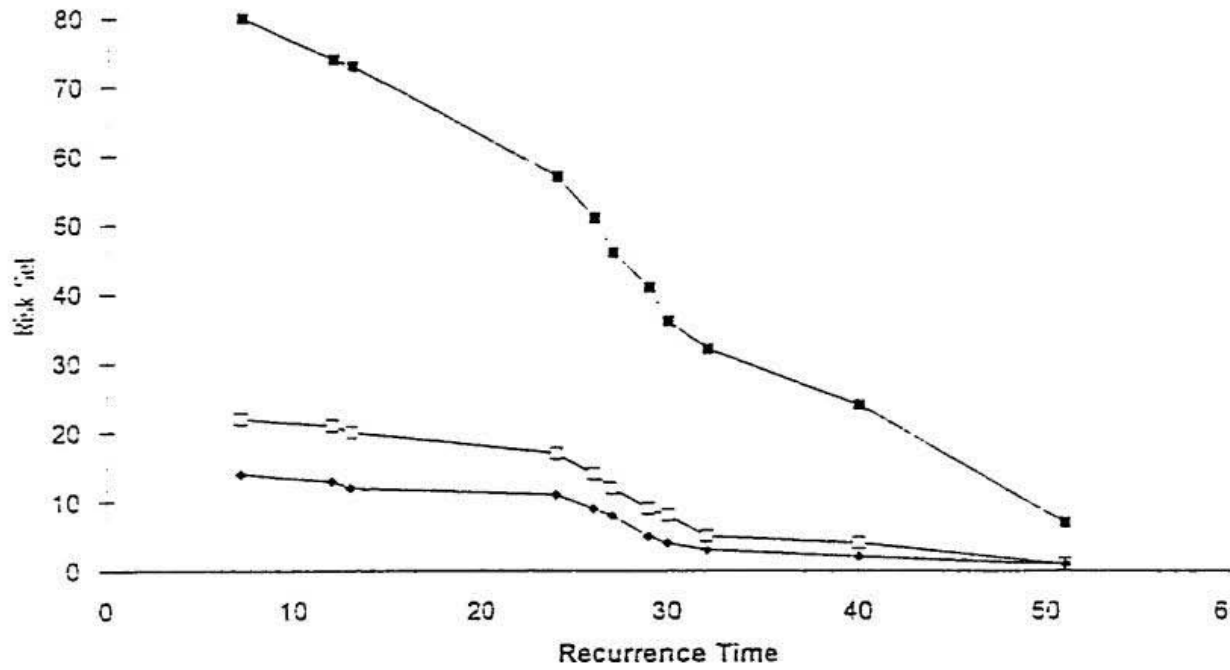

Figure 4.7: Risk Comparision of Original Tumor Recurrence Data for the Fourth Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\boxminus)$, and $A(III)(\blacklozenge)$ Approaches.

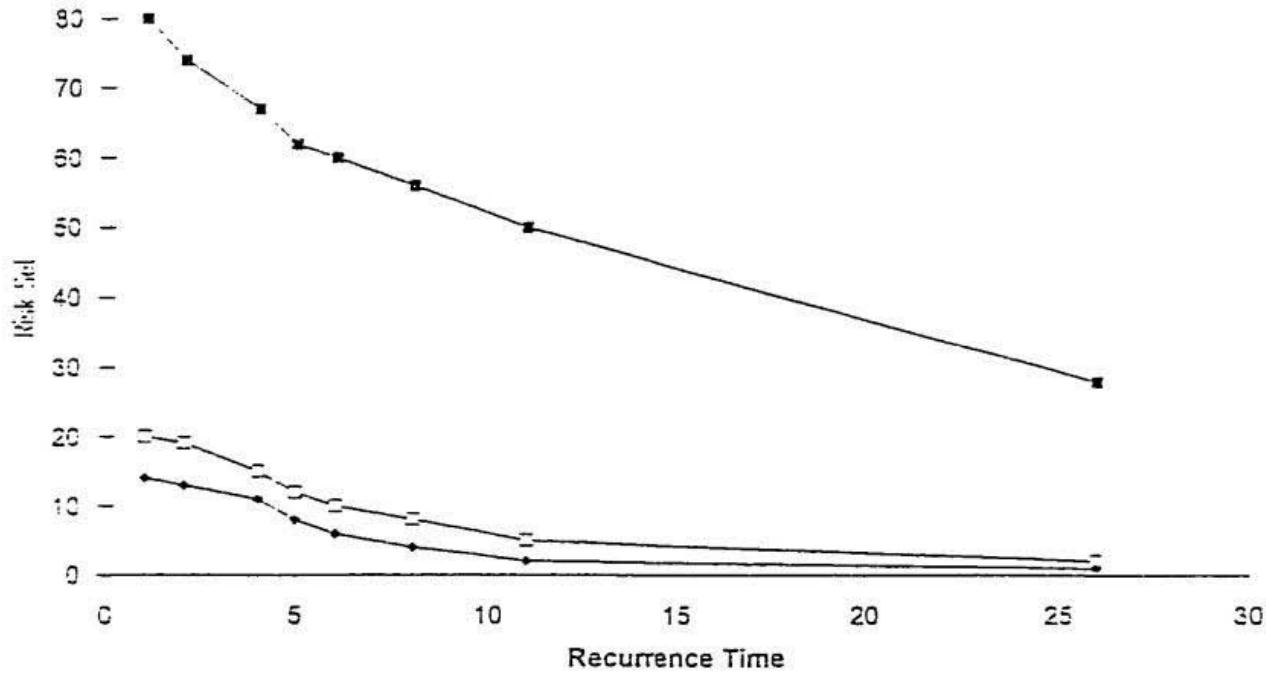

Figure 4.8: Risk Comparision of Modified Tumor Recurrence Data for the Fourth Recurrence Group Based on $A(I)(\blacksquare)$, $A(II)(\square)$, and $A(III)(\blacklozenge)$ Approaches.

4.2 Estimation of Treatment Effects in Different Recurrence Groups: Univariate Partial Likelihood Approach

In this section similar to Wei et al. (1989), we apply the univariate (marginal) partial likelihood estimation approach to the modified data and estimate the treatment effects for four different recurrence groups. The application of the univariate (marginal) partial likelihood estimation method to the modified data appears to be much more meaningful than to applying this method to the original data. This is because the modification to the original data is done in such a way that the four recurrence time groups would be moderately independent of each other. More specifically, the current recurrence time is computed in such a way that it is independent of the previous time(s) for the individual concerned. We report the regression estimate of the effect of the treatment covariate for the modified data with the standard error in columns 2 to 5 of table 4.1. Our results are presented for three different approaches $A(I)$, $A(II)$, and $A(III)$. where the approaches are defined in the last section. We also apply the univariate partial likelihood method to the original data and estimate the treatment effects for four different groups under all of the three approaches (Table 4.2). Observe that the regression estimates and their standard errors under approach $A(I)$ in Table 4.2 are same as the estimate obtained by Wei et al. (1989, Table 5, p. 1070). It is clear from Table 4.1 that when $H_0 : \beta_{k1} \geq 0$ $(k = 1, \cdots, 4)$ is tested against $H_a : \beta_k < 0$, the treatment effect appears to be insignificant except in group 1 under both $A(I)$ and $A(II)$ approaches. Here $\beta_{k1} \geq 0$ indicates that there is no treatment effect or the treatment may detoriate patient's condition, and $\beta_k < 0$ indicates that treatment is effective. Under approach $A(III)$, the treatment effects appear to be insignificant

in all the recurrence groups. The significant pattern of the treatment effects for the original data (cf. Table 4.2), in general, however, appears to be quite different. For the original data, the treatment effects under approach $A(I)$, as also mentioned by Wei et al. (cf. Table 5), are found to be significant for the first three recurence groups, whereas treatment was found to be insignificant in the last group. In approach $A(II)$, the treatment effect was found to be significant only for the first recurrence group. We remark here that the censoring mechanism in $A(II)$ is similar to that of Prentice et al. (1981). Wei et al. (1989) also obtained the treatment effect under approach $A(II)$, but by using the model (2) due to Prentice et al. (1981), and it was found that the treatment was significant in the recurrence groups 1 and 3, but not in group 2. Wei et al. (1989) interpreted this difficulty as an effect of smaller risk sets in approach $A(II)$. But, as it was mentioned above, this difficulty does not arise when treatment effects are computed under approach $A(II)$ by using the univariate partial likelihood method. Thus, this problem does not appear to be due to the smaller risk set, rather it may be attributed to the selection of the estimation method. When the treatment effects for the modified data in Table 4.1 are compared to the treatment effects for the original data in Table 4.2 the treatment appears to be significant on more occasions for the original data, which may be due to the application of the univariate partial method to the correlated data. To be more specific, one of the main reasons that treatment effects appear to be significant in more occasions for the original data is that the standard errors of $\hat{\beta}_{k1}$ (regression estimate corresponding to the treatment effect) appear to be smaller, in general, than those for the modified data. This behaviour of the estimates of the standard errors of $\hat{\beta}_{k1}$ does not guarantee inference accuracy of the treatment effects for the original data. This is because, it has been shown empirically in the context of cluster regression study by some authors

(cf. Sutradhar and Qu (1995)) that when correlation among observations in a cluster increases, the standard errors of the regression estimates get smaller. Consequently, although the estimated standard errors of the treatment effects for the modified data are found to be generally larger, they appear to be more correct standard errors as they were obtained by the marginal partial likelihood approach applied to the right type of data, i.e. to the independent data.

We also analyse the modified and original data using only treatment as a covariate. The results for modified and original data are presented in Tables 4.3 and 4.4 respectively. From Table 4.3, the treatment effects appear to be insignificant for all of the recurrences under approaches $A(I)$, $A(II)$ and $A(III)$, which may be due to ignoring the other two covariates, initial tumor number and initial tumor size. Similar results are found from Table 4.4 for the original data except for group-2 under $A(III)$.

Table 4.1: Regression Analyses of Treatment Effects for the Modified Data Based on Partial Likelihood Method for all Groups; and Linear and Restricted Partial Likelihood (RPL) Estimation Methods for the Combined Group.

App.	Recurrence number				Combined Est. $\hat{\beta}^{\bullet}_{\cdot 1}$	
	1:$\hat{\beta}_{11}$	2:$\hat{\beta}_{21}$	3:$\hat{\beta}_{31}$	4:$\hat{\beta}_{41}$	LE	RPLE
A(I)	-0.514	-0.526	-0.525	-0.327	-0.508	-0.494
	(0.308)	(0.347)	(0.458)	(0.566)	(0.291)	(0.193)
	[-1.66]	[-1.52]	[-1.15]	[-0.58]	[-1.75]	[-2.55]
A(II)	-0.514	-0.246	0.243	-0.027	-0.252	-0.276
	(0.308)	(0.508)	(0.821)	(1.557)	(0.244)	(0.201)
	[-1.66]	[-0.48]	[0.29]	[-0.017]	[-1.03]	[-1.37]
A(III)	-0.066	-0.419	0.802	0.045	-0.058	-0.006
	(0.331)	(0.642)	(0.727)	(0.705)	(0.247)	(0.209)
	[-0.199]	[-0.653]	[1.103]	[0.064]	[-0.235]	[-0.029]

Note:

- Estimated standard errors are in parenthesis ().
- Z-scores ($Z = \frac{\hat{\beta}_{k1}}{s.e.(\hat{\beta}_{k1})}$) are in square braket [].
- Critical value for $Z = -1.645$ for 1-tailed test with $\alpha = 0.05$.

Table 4.2: Regression Analyses of Treatment Effects for the Original Data Based on Partial Likelihood Method for all Groups; and Linear and Restricted Partial Likelihood (RPL) Estimation Methods for the Combined Group.

App.	Recurrence number				Combined Est. $\hat{\beta}_{\cdot 1}^{*}$	
	1:$\hat{\beta}_{11}$	2:$\hat{\beta}_{21}$	3:$\hat{\beta}_{31}$	4:$\hat{\beta}_{41}$	LE	RPLE
A(I)	-0.514	-0.619	-0.697	-0.650	-0.547	-0.580
	(0.308)	(0.364)	(0.415)	(0.488)	(0.286)	(0.212)
	[-1.668]	[-1.701]	[-1.680]	[-1.339]	[-1.910]	[-2.740]
A(II)	-0.514	-0.431	0.146	0.754	-0.400	-0.515
	(0.308)	(0.410)	(0.495)	(0.916)	(0.239)	(0.194)
	[-1.668]	[-1.051]	[0.295]	[0.823]	[-1.670]	[-2.650]
A(III)	-0.066	-1.386	-0.495	0.183	-0.349	-0.429
	(0.331)	(0.460)	(0.537)	(0.615)	(0.273)	(0.196)
	[-0.199]	[-3.013]	[-0.922]	[0.298]	[-1.270]	[-2.180]

Note:

- Estimated standard errors are in parenthesis ().
- Z-scores ($Z = \frac{\hat{\beta}_{k1}}{s.e.(\hat{\beta}_{k1})}$) are in square braket [].
- Critical value for $Z = -1.645$ (1-sided test at $\alpha = 0.05$).

Table 4.3: Regression Analyses of Treatment Effects (ignoring other covariates) for the Modified Data Based on Partial Likelihood Method for all Groups; and Linear and Restricted Partial Likelihood (RPL) Estimation Methods for the Combined Group.

App.	Recurrence number				Combined Est. $\hat{\beta}_{\cdot 1}^{*}$	
	1:$\hat{\beta}_{11}$	2:$\hat{\beta}_{21}$	3:$\hat{\beta}_{31}$	4:$\hat{\beta}_{41}$	LE	RPLE
A(I)	-0.362	-0.472	-0.546	-0.414	-0.368	-0.433
	(0.298)	(0.386)	(0.498)	(0.551)	(0.329)	(0.196)
	[-1.210]	[-1.220]	[-1.100]	[-0.750]	[-1.120]	[-2.210]
A(II)	-0.362	-0.163	0.274	0.387	-0.119	-0.116
	(0.298)	(0.380)	(0.498)	(0.533)	(0.197)	(0.198)
	[-1.210]	[-0.420]	[0.550]	[0.720]	[-0.600]	[-0.590]
A(III)	-0.023	-0.276	0.223	0.341	-0.039	-0.043
	(0.302)	(0.431)	(0.535)	(0.546)	(0.231)	(0.206)
	[-0.100]	[-0.640]	[0.420]	[0.620]	[-0.170]	[-0.210]

Note:

- Estimated standard errors are in parenthesis ().
- Z-scores ($Z = \frac{\hat{\beta}_{k1}}{s.e.(\hat{\beta}_{k1})}$) are in square braket [].
- Critical value for $Z = -1.645$ (1-sided test at $\alpha = 0.05$).

Table 4.4: Regression Analyses of Treatment Effects (ignoring other covariates) for the Original Data Based on Partial Likelihood Method for all Groups; and Linear and Restricted Partial Likelihood (RPL) Estimation Methods for the Combined Group.

App.	Recurrence number				Combined Est. $\hat{\beta}^{*}_{\cdot 1}$	
	1:$\hat{\beta}_{11}$	2:$\hat{\beta}_{21}$	3:$\hat{\beta}_{31}$	4:$\hat{\beta}_{41}$	LE	RPLE
A(I)	-0.362	-0.472	-0.547	-0.414	-0.390	-0.470
	(0.298)	(0.372)	(0.442)	(0.529)	(0.291)	(0.192)
	[-1.210]	[-1.270]	[-1.240]	[-0.780]	[-1.340]	[-2.450]
A(II)	-0.362	-0.323	-0.664	-0.031	-0.330	-0.368
	(0.298)	(0.353)	(0.471)	(0.515)	(0.215)	(0.190)
	[-120]	[-0.920]	[-1.410]	[-0.060]	[-1.530]	[-1.940]
A(III)	-0.023	-0.788	-0.624	-0.243	-0.276	-0.361
	(0.301)	(0.396)	(0.483)	(0.506)	(0.257)	(0.198)
	[-0.080]	[-1.990]	[-1.290]	[-0.480]	[-1.070]	[-1.820]

Note:

- Estimated standard errors are in parenthesis ().
- Z-scores ($Z = \frac{\hat{\beta}_{k1}}{s.e.(\hat{\beta}_{k1})}$) are in square braket [].
- Critical value for $Z = -1.645$ (1-sided test at $\alpha = 0.05$).

4.3 Combined Estimate of Treatment Effect: Linear Estimation (LE) Versus Restricted Partial Likelihood Estimation (RPLE)

In this section unlike Wei et al. (1989), we estimate the combined effect of treatment for modified as well as original data by using the restricted partial likelihood estimation (RPLE) method under all three approaches. In this RPLE method we use the restriction $\beta_{k1} = \beta^*_{.1}$ for $k = 1, \cdots, K$ and then exploit the partial likelihood function (3.2) to derive $\hat{\beta}^*_{.1}$ and $\hat{\beta}_{ku}$ for $u = 2, \cdots, p$ and $k = 1, \cdots, K$. To obtain the standard errors of $\hat{\beta}^*_{.1}$ we partition the $(K(p-1)+1) \times (K(p-1)+1)$ covariance matrix Q^* of $\beta^* = (\beta^*_{.1}, \beta^*_1, \cdots, \beta^*_K)'$, where, $\beta^*_k = (\beta_{k2}, \cdots, \beta_{Kp})'$. We also compute the combined treatment effect for the modified data by using the linear estimation (LE) method used by Wei et al. (1989). The standard error of the estimate is calculated by computing $\sum_{k=1}^{K} W^{*2}_k V(\hat{\beta}_{k1})$, where $V(\hat{\beta}_{k1})$ is obtained from the $K \times K$ covariance matrix $\hat{\psi}$ of $\hat{\beta}_{k1}$ (see Wei et al. (1989), 3.2) and W^*_k is the kth element of the weight vector W^*. Further, we present the results on the combined effect of treatment for the original data based on the LE method. These estimates and their standard errors are shown in the last two columns of Tables 4.1, 4.2, 4.3 and 4.4. Consider the results shown in Table 4.1 and 4.2. It is interesting to observe that the RPLE method always yields smaller standard errors than the LE method. This is true for both the original and modified data. This is not surprising because unlike the linear estimate used by Wei at el. the RPL estimate exploits the likelihood function. Observe that for the modified and original data, both methods generally appear to give similar conclusions about the significance of the combined treatment effects. More specifically, both methods yield significant treatment effects for the original data under approaches $A(I)$ and

$A(II)$, and they yield similar significant effects under approach $A(I)$ for the modified data. But, the combined treatment effects yielded by the RPLE method are found to be significant at a lower level of significance (1%) than those yielded by the LE method (significant at 5% level). Since the inference made by the RPLE method is likelihood based, in contrast to the inference made by the LE method, the results produced by the RPLE method about the treatment effects are preferable to the LE method.

Further, when a particular method (LE or RPLE) is examined for the combined linear effects for modified and original data, it is found that the standard errors of the estimates are generally smaller for the original data as compared to the modified data. Consequently, the treatment effects appear to be significant on more occasions for the original data as compared to the modified data. These conclusions about the combined treatment effects are quite similar to the conclusions about the group treatment effects for the modified and original data, reported in the last section.

In Table 4.3. where treatment is the only covariate, the combined treatment effects are found to be insignificant under all of the approaches except for the RPLE method under $A(I)$. For the original data, the combined treatment effects appear to be insignificant for the LE method and significant for the RPLE method under all of the approaches. Thus we observe again that the significance pattern for treatment effect for the original data differs from that of the modified data, the treatment effects being significant in more occasions for the original data. It is also observed again that under all of the approaches in Table 4.3 and 4.4 the RPLE method gives smaller standard errors of estimates than the LE method.

We now combine all four groups to form a similar group with respect to all three covariates: treatment, initial tumor number, and initial tumor size for the modified

as well as the original data and study the significance pattern of treatment. For both data sets, the RPLE method is used to obtain the combined estimates of treatment effects and combined effects of each of the other two covariates. The results presented in Table 4.5 show that for the modified data the treatment effects are significant only under approach $A(I)$, whereas, for the original data the treatment effects are significant under all three approaches. Thus the conclusions derived from Table 4.5 about the combined treatment effect is similar to the conclusions already obtained above. Observe, again that the standared errors of the RPL estimates for the modified data are larger than those for the original data under all of the approaches, which may be because of the correlation in the original data.

Table 4.5: Regression Analyses of All Covariates (treatment, tumor size, and tumor number) for Modified and Original Data Based on Restricted Partial Likelihood Estimation (RPLE) Method for the Combined Group.

App.	Modified Data			Original Data		
	$1{:}\hat{\beta}^{*}_{11}$	$2{:}\hat{\beta}^{*}_{12}$	$3{:}\hat{\beta}^{*}_{13}$	$1{:}\hat{\beta}^{*}_{11}$	$2{:}\hat{\beta}^{*}_{12}$	$3{:}\hat{\beta}^{*}_{13}$
A(I)	-0.501	0.225	-0.011	-0.579	0.209	-0.051
	(0.199)	(0.047)	(0.070)	(0.201)	(0.047)	(0.069)
	[-2.51]	[4.78]	[-0.15]	[-2.88]	[4.44]	[-0.74]
A(II)	-0.244	0.161	0.013	-0.489	0.110	-0.038
	(0.209)	(0.052)	(0.069)	(0.209)	(0.051)	(0.068)
	[-1.17]	[29.8]	[-0.188]	[-2.34]	[2.20]	[-0.55]
A(III)	-0.005	0.055	0.047	-0.384	-0.005	0.029
	(0.236)	(0.057)	(0.069)	(0.227)	(0.056)	(0.069)
	[-0.021]	[0.96]	[0.68]	[-1.69]	[-0.089]	[0.420]

Note:

- Estimated standard errors are in parenthesis ().
- Z-scores ($Z = \frac{\hat{\beta}^{*}_{iu}}{s.e.(\hat{\beta}^{*}_{iu})}$) are in square braket [].
- Critical value for $Z = -1.645$ (1-sided test at $\alpha = 0.05$).

4.4 Likelihood Ratio and Wald Type Tests for Simultaneous Treatment Effects

In this section we apply the Likelihood Ratio and Wald type tests to test the treatment effects of all recurrence groups for the modified as well as for the original data under all three approaches to handling censoring. More specifically, we test jointly the hypothesis $H_k : \beta_{k1} = 0 \quad (k = 1,2,3.4)$, first for the case where other covariates (initial tumor number and initial tumor size) are nuisance and second under the assumption that the other covariates are absent (ignored). Finally, we test the same hypothesis about the treatment effect for the combined group for the first case. To test the above hypothesis, for the first case, the Wald type test statistic as mentioned in chapter 3 is given by

$$W_1 = (\hat{\beta}_{11}, \cdots, \hat{\beta}_{K1})\hat{\psi}^{-1}_{KK}(\hat{\beta}_{11}, \cdots, \hat{\beta}_{K1})'. \tag{4.1}$$

This was computed for both modified and original data. Also we performed the likelihood ratio test, where the test statistic λ^* for the first case is given by

$$\lambda_1 = -2\log\left[\frac{L_0(\beta_{k1} = 0; \hat{\beta}_{ku'})}{L_1(\tilde{\beta}_{ku})}\right] \tag{4.2}$$

with $u' = 2,3, \quad u = 1,2,3, \quad$ and $\quad k = 1, \cdots, K$. The Wald and likelihood ratio test statistics W_1 (4.1) and λ_1 (4.2) have an approximately χ^2 distribution with K degrees of freedom. In (4.2), $\hat{\beta}_{ku'}$ denotes the restricted regression estimate of $\beta_{ku'}$ and $\tilde{\beta}_{ku}$ denotes the unrestricted regression estimate of β_{ku}. The values of W_1 and λ_1 are shown in Table 4.6.

From Table 4.6 observed values of W_1 and λ_1 for testing the hypotheses $H_k : \beta_{k1} = 0 \quad (k = 1,2,3,4)$ jointly, each with 4 df appears to be insignificant (at $\alpha = 0.05$) for the modified data under all of the approaches. The Wald and likelihood ratio test

Table 4.6: The Values of the Wald type and Likelihood Ratio Test Statistics for Testing Treatment Effects (other covariates being nuisance)

Approach	Modified Data		Original Data	
	W_1	λ_1	W_1	λ_1
A(I)	3.44 ($p > 0.1$)	6.45 ($p > 0.1$)	3.93 ($p > 0.1$)	5.05 ($p > 0.1$)
A(II)	7.71 ($p > 0.1$)	7.41 ($p > 0.1$)	5.08 ($p > 0.1$)	30.91 ($p < .005$)
A(III)	3.32 ($p > 0.1$)	2.30 ($p > 0.1$)	10.71 ($.025 < p < .05$)	8.69 ($p > 0.1$)

appears to give conflicting inferences for the original data under approaches $A(II)$ and $A(III)$.

To test the treatment effect for the second case we computed the Wald type test statistic W_1 given in (4.1), but the β's and ψ's are computed under the condition that we do not have available information on the other two covariates. We denote the Wald statistic in this case by W_2. The likelihood ratio test statistic here reduces to

$$\lambda_2 = -2\log\left[\frac{L_0(\beta_{k1} = 0)}{L_1(\tilde{\beta}_{k1})}\right]$$

where $\tilde{\beta}_{k1}$ is the regression estimate of β_{k1}. The values of W_2 and λ_2 are given in Table 4.7. The observed values of W_2 and λ_2 each with 4 d.f. to test the above hypothesis appears to be insignificant (at $\alpha = 0.05$) both for the original and the modified data under all three approaches.

Table 4.7: The Values of the Wald type and Likelihood Ratio Test Statistics for Testing Treatment Effects (other covariates being ignoreed)

Approach	Modified Data		Original Data	
	W_2	λ_2	W_2	λ_2
A(I)	0.86 ($p > 0.1$)	5.08 ($p > 0.1$)	1.52 ($p > 0.1$)	6.11 ($p > 0.1$)
A(II)	1.50 ($p > 0.1$)	2.41 ($p > 0.1$)	1.37 ($p > 0.1$)	4.15 ($p > 0.1$)
A(III)	0.04 ($p > 0.1$)	1.89 ($p > 0.1$)	0.04 ($p > 0.1$)	5.71 ($p > 0.1$)

Finally, for the third case we also compute the Wald type and likelihood ratio test statistics W_3 and λ_3 respectively to test the treatment effects of the combined group, that is , to test the hypothesis $H_0 : \beta^*_{.1} = 0$. Here the Wald test statistic is given by

$$W_3 = \hat{\beta}^*_{.1}\hat{\psi}^{-1}_{11}\hat{\beta}^*_{.1}$$

where $\hat{\beta}^*_{.1}$ is the combined treatment effect and $\hat{\psi}_{11}$ is the variance covariance matrix corresponding to treatment which is marginally pickedup from the full variance covariance matrix $\hat{\psi}_{33}$. And the likelihood ratio test statistic is as follows

$$\lambda_3 = -2\log\left[\frac{L_0(\beta^*_{.1} = 0; \hat{\beta}^*_{.u'})}{L_1(\tilde{\beta}^*_{.u})}\right]$$

where $u' = 2, 3$; $u = 1, 2, 3$ and $\hat{\beta}^*_{.u'}$ is the estimated combined effects of the other two covariates under the null hypothesis $H_0 : \beta^*_{.1} = 0$ and $\tilde{\beta}^*_{.u}$ is the regression estimate under the assumption: $\beta_{ku} = \beta_{.u}$, for $k = 1, 2, 3, 4$ and $u = 1, 2, 3$. Now the values of W_3 and λ_3 each with 1 d.f. presented in Table 4.8 show that the treatment effect

Table 4.8: The Values of the Likelihood Ratio Test Statistic for Testing Treatment Effects (combining the groups for all covariates)

Approach	Modified Data		Original Data	
	W_3	λ_3	W_3	λ_3
A(I)	0.95 ($p > 0.1$)	6.64 ($p = 0.01$)	1.11 ($p > 0.1$)	8.93 ($p < 0.005$)
A(II)	0.066 ($p > 0.1$)	1.48 ($p > 0.1$)	0.67 ($p > 0.1$)	6.12 ($.01 < p < .025$)
A(III)	0.009 ($p > 0.1$)	0.001 ($p > 0.1$)	0.63 ($p > 0.1$)	3.51 ($.05 < p < .1$)

appears to be insignificant (at $\alpha = 0.05$) for the modified data under all three approaches except for the likelihood ratio test under approach $A(I)$. For the original data the Wald and the likelihood ratio tests give conflicting inferences for the first two approaches.

Chapter 5

Summary and Some Topics For Further Research

5.1 Summary

It was found by Wei et al. (1989) that theotepa (treatment) was significantly effective for the first three recurrence groups under approach $A(I)$. When the data were modified to make the recurrence groups as independent as possible and a partial likelihood approach similar to Wei et al. (1989) was used for estimation, it was found that the treatment effects are significant for the first recurrence group under approach $A(I)$ only. This discrepancy may be attributed due to the fact that Wei et al. applied the univariate partial likelihood method to correlated data.

Again when we estimated the combined treatment effects for both data sets, it was found that the treatment effects appeared to be significant in more occasions for the original data as compared to the modified data. Thus we get the same conclusion about the treatment effects from the combined and group estimates.

It was found that although in some cases both the methods (RPLE and LE)

lead to similar conclusions about the significance of the treatment effects, the RPLE method yields the regression estimates with smaller standard errors. Thus for this blood cancer data set, there seems to be a difference in the outcomes depending on the method chosen.

5.2 Topics for Further Research

In this practicum, the univariate partial likelihood was applied to the original as well as modified data, where the modified data was obtained by a suitable differencing technique. The purpose of the modification was to make the failure times recorded at different stages, as independent as possible. It is interesting to observe that the standard errors of the regression estimates for the modified data were generally larger than those from the original data. This finding is in agreement with the simulation results shown by Sutradhar and Qu (1995), in the context of a cluster regression problem, that as the cluster correlation decreases the standard errors of the regression estimates increases. Note, however, that a theoretical derivation to justify the above finding is not easy in general. Further study in this direction is necessary.

In contrast to the construction of independent failure time data, an alternative way to take the correlation of the data into account, is to apply the generalized estimation equation approach suggested by Liang and Zeger (1986). This requires the modelling of the correlations among the repeated failure times for each individual, which also does not appear to be easy.

Recently, Cai and Prentice (1995) have considered a specific correlation structure to analyse such correlated data. Since their work seems highly relevant to this work, a comparative study can be made in future.

Chapter 6

References

Alen, O. (1978). Nonparametric Inference for a Family of Counting Processes. *Ann. of Stat.*, 6, 701-726.

Andersen, P. K., and Gill, R. D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann. of Stat.* 10, 1100-1120.

Byar, D. P. (1980). The Veterans Administration Study of Chemoprophylaxis for Recurrent Stage I Bladder Tumors: Comparisons of Placebo, Pyridoxine, and Topical Thiotepa, *in Bladder Tumors and Other Topics in Urological Oncology*, eds. M. Pavone-Macaluso, P.H. Smith, and F. Edsmyn. New York: Plenum, 363-370.

Cai, J., and Prentice, R.L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika*, 82, 1, 151-164.

Cox, D. R. (1972). Regression Models and Life Tables (with discussion). *J. R.*

Statist. Soc. B, 34, 187-202.

Cox, D. R. (1975). Partial Likelihood. *Biometrika*, 62, 2, 269-276.

Efron, B (1977). The Efficiency of Cox's Likelihood Function for Censored Data. *Am. Stat. Assoc.* 72, 555-565.

Gail, M.H., Santer, T.J., and Brown, C.C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics*, 36, 255-266.

Guo, S. W., and Lin, D. Y. (1994). Regression Analysis of Multivariate Group Survival Data. *Biometrics*, 50, 632-639.

Huster W S., Brookmeyer, R. and Self, S.G. (1989). Modeling Paired Survival Data with Covariates. *Biometrics*. 45, 145-156.

Kalbfleisch, J. D., and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Kalbfleisch, J. D., and R. J. MacKay (1978). Censoring and the Immutable Likelihood. University of Waterloo, Department of Statistics. *Technical Report* 78-09.

Liang, K.-Y. and Zeger, S.L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, 72, 13-22.

Liang, K.-Y., Steven, G. S. and Chang, Y.-C. (1993). Modelling Marginal Hazards in Multivariate Failure Time Data. *J.R. Statist. Soc.* B, 55, 2, 441-453.

Lawless, J. F. (1982). *Statistical Models and Methods for Life Time Data.* New York: John Wiley.

Lawless, J. F. (1987). Regression Methods for Poisson Process Data. *Journal of the American Statistical Association.* 82, 808-815.

Liu, P. Y., and J. Crowley (1978). Large Sample Properties of the Maximum Partial Likelihood Estimate Based on Cox's Regression Model for Censored Survival Data. University of Wisconsin, Department of Biostatistics, Wisconsin Clinical Career Center, *Technical Report* No. 1.

Mukuch, R. W., and Parks, W. P. (1988). Statistical Methods for the Analysis of HIV-1 Core Polypeptide Antigen Data in Clinical Studies. *AIDS Research and Human Retroviruses.* 4, 305-316.

Murphy, R.P. and Patz, A. (1978). New Concept in Management of Retinal Vascular Disorder. *In Ophthalmology Update, Internctional Congress Series*, No. 508, K. Mizuno and Y. Mitsui (eds), 111-125. Amsterdam: Excerpta Medica.

Peto, R. (1972a). Discussion of Paper by D.R. Cox. *J.R. Stat. Soc. B*, 34, 205-207.

Prentice R.L., B.J. Williams and A.V. Peterson (1981). On the Regression Analysis of Multivariate Failure Time Data. *Biometrika*, 68, 373-79.

Sutradhar, B.C. and Qu, Z. (1995). On Approximate Likelihood Inference in Poisson Mixed Model. *Submitted for Publication.*

Thompson, H.F., Grubbs, C.J., Moon, R.C. and Sporn, M.B. (1978). Continual Requirement of Retinoid for Maintenance of Mammary Cancer Inhibition. *Proceedings of the Annual Meeting of the American Association for Cancer Research* 19, 74.

Tsiatis, A. A. (1978a). A Large Sample Study of the Estimate for the Integrated Hazard Function in Cox's Regression Model for Survival Data. University of Wisconsin, Department of Statistics, *Technical Report* No. 526.

Wei, L.J., Lin, D.Y. and Weissfeld, L. (1989). Regression Analysis of Multivariate Incomplete Failure Time Data by Modelling Marginal Distributions. *JASA*, 84, 408.

Wei, L.J., and Johnson, W.E. (1985). Combining dependent tests with incomplete repeated measurements. *Biometrika*, 72, 359-364.

Wei, L.J. and Lin, D.Y. (1989). The Robust Inference for Cox Proportional Hazards Models. *JASSA*, 84, 408.

Chapter 7

Appendix

Table 7.1: Original Tumor Recurrence Data for Patients With Bladder Cancer

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time t_i	Recurrence time			
				r_{1i}^*	r_{2i}^*	r_{3i}^*	r_{4i}^*
1	1	1	0	0	0	0	0
1	1	3	1	0	0	0	0
1	2	1	4	0	0	0	0
1	1	1	7	0	0	0	0
1	5	1	10	0	0	0	0
1	4	1	10	6	0	0	0
1	1	1	14	0	0	0	0
1	1	1	18	0	0	0	0
1	1	3	18	5	0	0	0
1	1	1	18	12	16	0	0
1	3	3	23	0	0	0	0

table 7.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time t_i	Recurrence time			
				r_{1i}^{-}	r_{2i}^{-}	r_{3i}^{-}	r_{4i}^{-}
1	1	3	23	10	15	0	0
1	1	1	23	3	16	23	0
1	3	1	23	3	9	21	0
1	2	3	24	7	10	16	24
1	1	1	25	3	15	25	0
1	1	2	26	0	0	0	0
1	8	1	26	1	0	0	0
1	1	4	26	2	26	0	0
1	1	2	28	25	0	0	0
1	1	4	29	0	0	0	0
1	1	2	29	0	0	0	0
1	4	1	29	0	0	0	0
1	1	6	30	28	30	0	0
1	1	5	30	2	17	22	0
1	2	1	30	3	6	8	12
1	1	3	31	12	15	24	0
1	1	2	32	0	0	0	0
1	2	1	34	0	0	0	0
1	2	1	36	0	0	0	0

table 7.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time t_i	Recurrence time			
				r^*_{1i}	r^*_{2i}	r^*_{3i}	r^*_{4i}
1	3	1	36	29	0	0	0
1	1	2	37	0	0	0	0
1	4	1	40	9	17	22	24
1	5	1	40	16	19	23	29
1	1	2	41	0	0	0	0
1	1	1	43	3	0	0	0
1	2	6	43	6	0	0	0
1	2	1	44	3	6	9	0
1	1	1	45	9	11	20	26
1	1	1	48	18	0	0	0
1	1	3	49	0	0	0	0
1	3	1	51	35	0	0	0
1	1	7	53	17	0	0	0
1	3	1	53	3	15	46	51
1	1	1	59	0	0	0	0
1	3	2	61	2	15	24	30
1	1	3	64	5	14	19	27
1	2	3	64	2	8	12	13
2	1	3	1	0	0	0	0

table 7.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time t_i	Recurrence time			
				r^*_{1i}	r^*_{2i}	r^*_{3i}	r^*_{4i}
2	1	1	1	0	0	0	0
2	8	1	5	5	0	0	0
2	1	2	9	0	0	0	0
2	1	1	10	0	0	0	0
2	1	1	13	0	0	0	0
2	2	6	14	3	0	0	0
2	5	3	17	1	3	5	7
2	5	1	18	0	0	0	0
2	1	3	18	17	0	0	0
2	5	1	19	2	0	0	0
2	1	1	21	17	19	0	0
2	1	1	22	0	0	0	0
2	1	3	25	0	0	0	0
2	1	5	25	0	0	0	0
2	1	1	25	0	0	0	0
2	1	1	26	6	12	13	0
2	1	1	27	6	0	0	0
2	2	1	29	2	0	0	0
2	8	3	36	26	35	0	0
2	1	1	38	0	0	0	0

table 7.1 contd.

Treatment group	Ini. Tum. size	Ini. Tum. number	Follow-up time t_i	Recurrence time			
				r^*_{1i}	r^*_{2i}	r^*_{3i}	r^*_{4i}
2	1	1	39	22	23	27	32
2	6	1	39	4	16	23	27
2	3	1	40	24	26	29	40
2	3	2	41	0	0	0	0
2	1	1	41	0	0	0	0
2	1	1	43	1	27	0	0
2	1	1	44	0	0	0	0
2	6	1	44	2	20	23	27
2	1	2	45	0	0	0	0
2	1	4	46	2	0	0	0
2	1	4	46	0	0	0	0
2	3	3	49	0	0	0	0
2	1	1	50	0	0	0	0
2	4	1	50	4	24	27	0
2	3	4	54	0	0	0	0
2	2	1	54	38	0	0	0
2	1	3	59	0	0	0	0

- Note: Treatment group: 1, placebo; 2, thiotepa. Follow-up time and recurrence time are measured in months. Initial tumor size is measured in centimeters. Initial tumor number of 8 denotes eight or more initial tumors.