4.1 Introduction

In survival data, the outcome events may occur more than once over the follow-up time for a subject. It’s known as ‘recurrent events’. It is usually possible to study just the time until the first event, it may be useful to incorporate subsequent events to increase the information available. Modeling this kind of data can be carried out using Cox-PH model with the re-constructed data layout, so that each subject has a line of data corresponding to each recurrent event (Anderson, 1992). Variation of this approach uses a stratified Cox-PH model, which stratifies on the order in which the recurrent events occur. It considers adjusting the variances of estimated model coefficients for the likely correlation among recurrent events within the same subject. This adjusted variances estimates are known has ‘robust variance estimates’ (Kleinbaum, 1996).

The standard survival models for the additional adjustment for the recurrent event data is correlated between the original events for the same patients, leading both marginal models (Prentice et al., 1981) and for the frailty models. The counting process, marginal, conditional A (Gap time) and conditional B models can be represented by parametric and semi parametric context (Clayton, 1994). In these models, we assumed that all censoring is non-informative and independent. Suppose that censoring is non-informative, which means that knowledge of a censoring time for a subject provides no further information about the subject’s likelihood of survival at a future time. The gap time model requires the same assumptions as the Cox proportional
hazards model, but they allow the baseline hazard to vary from recurrence to recurrence (Kelly and Lim, 2000). Gap time is defined the time between two successive failures experienced by the same subject. In this chapter, we illustrated and compared the counting process and gap time hazards models for analysis of recurrent event durations. The choice between the counting process and gap time models will typically be an empirical matter. Due to the frequency of recurrent event duration data in clinical and epidemiologic studies, these two models are widely used Clayton (1994). The two modeling approaches have sound biological bases and gives complementary information about the association between risk factors. Hence, two hazards models give different information and it seems desirable to use them. It is not as alternatives to each other, but as complementary methods. Anderson and Gill (1982) proposed use of modeling under a Markov assumption. Wei et al., (1989) and Lee et al., (1992) explored the use of the marginal approach. Prentice et al., (1981) proposed use of a semi-parametric model when multivariate failure times are conditionally independent, given the covariates. Others used the random effect frailty model or the conditional frailty model for such recurrent event data analysis by Box-Steffensmeier et al., (2006). The popularity of these multiplicative models derives not only from their utility the semi parametric additive hazards model proposed by Lin and Ying, (1997) and is more closely connected analogue of the multiplicative Cox hazards model. Within the framework of the multiplicative or additive hazards regression models, a variety of models have been proposed and utilized in real applications. Among the rich selection of different models, the gap time model is an extension of the multiplicative Cox proportional hazards model and the additive model (L-Y model) received the greatest attention due to easy interpretation of the covariate effects. These two models assume unspecified baseline hazards and constant covariate effects.
The different ways of representing for the recurrent event data based on the time scale that used for the analysis. The time at risk for an event in total time representation starts at time zero when the subject enters the study until the particular event is experienced. Because of censoring or non-censoring for an event, the problem arises for what to do with the time when the subject was not at risk. This time scale may be subtracted but then the total time does no longer present the actual time since randomization (Rondeau, 2007). Furthermore, the total time representation has little appealing, because a patient is at risk for their events at the start of the study even when it is known that particular events can happened after the previous events occurred. Because these reasons we may not consider the total time in the remainder in case of the representation for the gap time, the risk at 0th time, but the length of the time at risk corresponding to the time since the end of the previous event occurred until the time the particular event occurred. The formulation for the counting process time, the length of the time at risk period is the same, but the start of the risk time is not reset at zero but the actual time since the study starts.

4.2 Counting Process

The counting process approach (Anderson et al., 1993) is used when recurrent events are treated has iid’s and the order of the events not important. The model typically used to carry out the counting process approach is the standard Cox PH model; the PH assumption needs to be evaluated for any time-independent variable. If recurrent events different disease categories or event order important, then stratified Cox model approach is used. The extended Cox model can be used if one or more time-independent variables did not satisfy the PH assumption. An extended Cox model would be required if inherently time-dependent variables were considered.
A counting process \( N(t) \), \( t \geq 0 \), as a stochastic process with \( N(0) \) is zero; \( N(t) \sim \alpha \), with probability one; and the sample of \( N(t) \) are right-continuous and piecewise constant with the size +1. With a right-censored sample, the processes,

\[
N_i(t) = \mathbb{I}[T_i \leq t, \delta_i = 1],
\]

which are ‘0’ until individual event \( i \) and then to one, are counting processes. The counting process can also be defined as

\[
N(t) = \sum_{i=1}^{n} N_i(t) = \sum_{i=1}^{\infty} \delta_i.
\]

It simply counts the number of events in the sample at or prior to time \( t \). The counting process indicates about when the events are occurring. In addition to this information, this process will give additional information on the study subjects at a time \( t \). For right censored data, this information at time \( t \) includes knowledge about who has been censored prior to time \( t \) and who met with the events at or prior to time \( t \). In some problems, our information may include values for a set of fixed time covariates, such as age, sex, treatment at time zero and possibly the values of time-dependent covariates, at all times prior to \( t \).

### 4.3 Gap Time (Conditional A) Model

The recurrent events may not be independent always for a same subject and the order of the events are important. Here the strata are the time interval numbers. The gap time or conditional A approach uses the exact same (start, stop) data layout format used for the counting process approach, except that for conditional A, an stratified Cox model is used rather than a standard (unstratified) PH model. The strata variable here is interval in this listing. Conditional A, the time until the first event affects the composition of the risk set for later events. The conditional A approach is preferred if the study goal is to use time of occurrence of each recurrent event from entry into the
study to assess a subject’s risk for an event of a specific order (i.e., as defined by a stratum #) to occur. The model is given by

\[ h_k(t, z, \beta_k) = h_{0k}(t) \exp(z' \beta_k), \]  

(4.3)

4.4 Conditional B Model

The conditional B approach also uses a (start, stop) data layout, but the start value is always 0 and the stop value is the time interval length since the previous event. The model here is also a stratified Cox model. The time until the first event does not influence the composition of the risk set for a second or later event. In other words, the clock for determining who is at risk gets reset to 0 after each event occurs.

The conditional B approach would be preferred if the time interval of interest is the time (reset from 0) from the previous event to the next recurrent event rather than time from study entry until each recurrent event. The model is

\[ h_k(t, z, \beta_k) = h_{0k}(t - t_{k-1}) \exp(z' \beta_k), \]  

(4.4)

4.5 Marginal Model

The marginal approach uses the standard (non-recurrent event) data layout instead of the (start, stop) layout. The marginal approach, each subject is considered to be at risk for all failures that might occur, regardless of the number of events a subject actually experienced. The basic idea behind the marginal approach is that it allows each failure to be considered as a separate process. The marginal approach not only allows the investigator to consider the ordering of failures as separate events (i.e., strata) of interest, but also allows the different failures to represent different types of events that may occur on the same subject. It considers all subject in this study contribute follow-up times to all possible recurrent events, whether they experienced that particular recurrence or not. This approach is in contrast to each conditional approach, focuses on
total survival time from study entry until the occurrence of a specific event. This approach is suggested when recurrent events are viewed to be of different types. It is as an alternative model for multivariate survival data, it has been suggested by (Lee et al., 1992). The proportional hazards model is assumed marginally for each individual, that is, for the \( j^{th} \) individual in the \( i^{th} \) group, the marginal hazard rate given an individual’s covariates \( Z_{ij} \) is given by

\[
h_{ij}(t | Z_{ij}) = h_0(t) \exp(\beta'Z_{ij}), \quad j=1,2,\ldots,n_i; \quad i=1,2,\ldots,G \quad (4.5)
\]

For estimating the estimate \( \beta \) coefficients, it is proceeding with an independence working model for the data. We assume that all observations are independent of each other and construct the partial likelihood function for a sample of \( \sum_{i=1}^{G} n_i \) observations.

Using this partial likelihood function, \( \hat{b} \), the estimator of \( \beta \) is found. Lee et al., (1992) showed that this estimator is consistent show that this estimator is consistent for \( \beta \), provided the marginal model is correctly specified. However, the information matrix obtained from this likelihood does not provide a valid estimator of the variance-covariance matrix of \( \hat{b} \). To estimate the variance of \( \hat{b} \), a “sandwich” estimator is used. This estimator adjusts the usual covariance matrix for the possible association between the event times within groups. To construct the covariance matrix, let \( \hat{V} \) be the usual \( p \times p \) covariance matrix for \( \hat{b} \), based on the independence working model. To estimate the variance correction matrix, let \( (T_{ij}, \delta_{ij}, Z_{ij}) \) be the on study time, event indicator, and covariate vector for the \( j^{th} \) individual in the \( i^{th} \) group. Let \( Y_{ij}(t) \) indicate if the \( j^{th} \) individual in the \( i^{th} \) group is at risk at time \( t \). It is given by

\[
S_{ij}(t) = \sum_{i=1}^{G} \sum_{j=1}^{n_i} Y_{ij}(t) \exp(b'Z_{ij}) \quad (4.6)
\]

and
\[
S_{ik}(t) = \sum_{i=1}^{G} \sum_{j=1}^{n_i} Y_{ij}(t) \exp(b'Z_{ij}), \quad k=1,2,\ldots,p \quad \text{for } k=1,2,\ldots,p \tag{4.7}
\]

compute

\[
W_{ijk} = \delta_{ij} \left[ Z_{ijk} - \frac{S_{ik}(T_{ij})}{S_{i}(T_{ij})} \right] - \sum_{g=1}^{G} \sum_{h=1}^{n_i} \delta_{gh} Y_{ij}(T_{gh}) \exp(b'Z_{ij}) \left[ Z_{ijk} - \frac{S_{ik}(T_{gh})}{S_{i}(T_{gh})} \right] \tag{4.8}
\]

Let the \( c_{bk} \)th element of the \( p \times p \) matrix \( C \) be explained by

\[
c_{bk} = \sum_{i=1}^{G} \sum_{j=1}^{n_i} \sum_{l=1}^{n_i} W_{jib} W_{lkb} \tag{4.9}
\]

4.6 Robust (Empirical) Estimation

The Robust estimator (Lin and Wei, 1989) for the usual single-event proportional hazards model is

\[
\hat{R}(\hat{\beta}) = \hat{\text{Var}}(\hat{\beta})[\hat{R}_s \hat{R}_s] \hat{\text{Var}}(\hat{\beta}) \tag{4.10}
\]

where \( \hat{\text{Var}}(\hat{\beta}) \) is the information matrix

\( R_s \) is matrix of score residuals obtained from MLE and

\( \hat{\beta} \) is an estimated regression coefficient

This estimation is used for adjusting the correlation among outcomes on the same subject. It also adjusts the estimated variances of regression coefficients obtained for fitted model to account for misspecification of a correlation structure assumed.

4.7 Parametric Frailty Model for Recurrent Event Data

Let the total number of patients be \( N \). A particular patient has different periods at risk during the total observation time, that are separated from each other by either an event (sputum culture positive to negative) that lasts one or more months, or by a period during which the patient was not under observation. If there are \( n_i \) at risk
periods for patient $i$, then the complete information about patient $i$ can be presented by $n_i$ triplets

$$(t_{i11}, t_{i21}, \delta_{i1}), \ldots, (t_{im_i1}, t_{im_i2}, \delta_{im_i})$$

where for the $j^{th}$ triplet, $t_{ij1}$ is the start of the $j^{th}$ at risk period, $t_{ij2}$ is the end of the $j^{th}$ period and $\delta_{ij}$ is the censoring indicator and $t_{i11}=0$.

### 4.7.1 Counting Process Time Frailty Model

The hazard function for the frailty model for the recurrent event (Duchateau et al., 2003) with counting process time is given by

$$h_i(t) = \begin{cases} h_0(t)Z_i \exp(\beta x_i) & \text{for } t_{ij1} \leq t \leq t_{ij2}, \ j = 1, 2, \ldots, n_i \\ 0 & \text{otherwise} \end{cases}$$

(4.11)

where $\lambda_0(t)$ is the baseline hazard which is assumed to be independent of both the event and the covariates considered for the patient and $Z_i$ is the frailty of the $i^{th}$ patient.

The frailties $Z_1, \ldots, Z_N$ are assumed to be independent with the common frailty density. As frailty density, it considers the one parameter gamma density with mean 1 and variance $\theta$.

$$f_{Z_i}(z) = \frac{1}{\theta^\theta \Gamma(\frac{1}{\theta})} z^{\theta-1} \exp(-\frac{z}{\theta})$$

(4.12)

The likelihood function to this hazard is given as

$$\prod_{i=1}^{N} \prod_{j=1}^{n_i} \left( h_i(t_{ij2}) \right)^{\delta_{ij}} \exp(-\Lambda_i(t_{ij1}, t_{ij2}))$$

(4.13)

The cumulative hazard is
\[ F_i(t_{ij1}, t_{ij2}) = \int_{t_{ij1}}^{t_{ij2}} h_i(t) dt \]  

(4.14)

The Weibull baseline hazard may be taken for the parametric frailty model for this data. The hazard function for the Weibull is given by (counting process time).

\[ h_i(t) = \begin{cases} 
    h_c \gamma t^{\gamma-1} Z_i \exp(\beta x_i) & \text{for } t_{ij1} \leq t \leq t_{ij2}, \ j = 1, 2, \ldots, n_i \\
    0 & \text{otherwise} 
\end{cases} \]  

(4.15)

and the likelihood function is

\[ \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left[ h_c \gamma t_{ij2}^{\gamma-1} Z_i \exp(\beta x_i) \delta_{ij} \exp(-h_c x(t_{ij2} - t_{ij1}^{\gamma}) Z_i \exp(\beta x_i)) \right] \]  

(4.16)

### 4.7.2 Gap Time Frailty Model

For the gap time case, part of the information in the triplets is redundant, and summarize the information in the triplets alternatively as

\[ ((t_{ij2} - t_{ij1}, \delta_{ij}), (t_{m_{i,1}, m_{i,1}, m_{i,j}})) \]

This can be represented for the length of the time at risk is needed, and not the particular time (relevant to study entry time) when the patient is at risk.

The hazard function for the frailty model with gap time is given by

\[ h_i(t) = \begin{cases} 
    h_o (t - t_{ij1}) Z_i \exp(-\beta x_i) & \text{for } t_{ij1} \leq t \leq t_{ij2}, \ j = 1, 2, \ldots, n_i \\
    0 & \text{otherwise} 
\end{cases} \]  

(4.17)

and the corresponding likelihood function is, as before,

\[ LL = \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left( h_i(t_{ij2}) \delta_{ij} \exp(-F_i(t_{ij1}, t_{ij2})) \right) \]  

(4.18)

For the different meaning for the hazard \( \lambda_i(\cdot) \) and cumulative hazard \( F_i(\cdot, \cdot) \).

The hazard function for gap time (Weibull-gap time) is
\[ h_i(t) = \begin{cases} h_{g \gamma} (t - t_{ij})^{\gamma - 1} Z_i \exp(\beta x_i) & \text{for } t_{ij} \leq t \leq t_{ij2}, \ j = 1,2,\ldots, n_i \\ 0 & \text{otherwise} \end{cases} \]  
(4.19)

and the resulting likelihood function

\[
LL = \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left[ \left( h_{g \gamma} (t_{ij2} - t_{ij})^{\gamma - 1} Z_i \exp(\beta x_i) \right)^{\delta_i} \exp \left( -h_{g \gamma} x (t_{ij2} - t_{ij})^{\gamma} Z_i \exp(\beta x_i) \right) \right] 
\]

(4.20)

For the exponential distribution, the constant baseline hazard ratio \( \lambda_0(t) = \lambda \), the likelihood functions for the gap time and counting process time are equal.

The previous models contain the frailty terms \( Z_i \), but not the frailty \( \theta \). In the parametric models the unobserved frailty terms is to integrate out the frailty density, to obtain the unobserved likelihood (Klein, 1992; Duchateau et al., 2002). A closed form for the observable likelihood for the one parameter gamma density with frailty terms \( Z_i \) and the frailty parameter \( \theta \) is

\[
LL = \prod_{i=1}^{N} \frac{\Gamma \left( \frac{1}{\theta} + d_i \right)}{(\theta^\theta \Gamma \left( \frac{1}{\theta} \right) \left( \frac{1}{\theta} + \sum_{j=1}^{n_i} \lambda_c x (t_{ij2} - t_{ij})^{\gamma} \exp(\beta x_i) \right)^{d_i \theta}} \left( h_{g \gamma} (t_{ij2} - t_{ij})^{\gamma - 1} \exp(\beta x_i) \right)^{\delta_i} 
\]

(4.21)

For the counting process time with \( d_i = \sum_{j=1}^{n_i} \delta_{ij} \), the number of recurrent events for patient \( i \).

The likelihood for the observable gap time model is given by

\[
LL = \prod_{i=1}^{N} \frac{\Gamma \left( \frac{1}{\theta} + d_i \right)}{(\theta^\theta \Gamma \left( \frac{1}{\theta} \right) \left( \frac{1}{\theta} + \sum_{j=1}^{n_i} \lambda_c x (t_{ij2} - t_{ij})^{\gamma} \exp(\beta x_i) \right)^{d_i \theta}} \left( h_{g \gamma} (t_{ij2} - t_{ij})^{\gamma - 1} \exp(\beta x_i) \right)^{\delta_i} 
\]

(4.22)
4.8 Semi-Parametric Frailty Models for Recurrent Event Data

In semi-parametric hazard model, the baseline hazard function is left unspecified. The difference between the counting process time and the gap time is in terms of risk sets.

The partial likelihood function for Cox PH for the counting process time is

\[
\prod_{i=1}^{N} \prod_{j=1}^{n_i} \left[ \frac{Z_i \exp(\beta x_i)}{\sum_{k=1}^{N} Y_k(t_{ij2}) Z_k \exp(\beta x_k)} \right]^{\delta_{ij}}
\]

(4.23)

with

\[
Y_k(t_{ij2}) = \begin{cases} 
1 & \text{if patient } k \text{ at risk at time } t_{ij2} \\
0 & \text{otherwise}
\end{cases}
\]

and for the gap time is

\[
\prod_{i=1}^{N} \prod_{j=1}^{n_i} \left[ \frac{Z_i \exp(\beta x_i)}{\sum_{k=1}^{N} \sum_{l=0}^{n_k} Y_{kl}(t_{ij2}) Z_k \exp(\beta x_k)} \right]^{\delta_{ij}}
\]

(4.24)

with

\[
Y_k(t_{ij2}) = \begin{cases} 
1 & \text{if } (t_{kl2} - t_{kl1}) \geq (t_{ij2} - t_{ij1}) \\
0 & \text{otherwise}
\end{cases}
\]

4.9 Application to the Clinical Trial Data

The database consists of 155 multi drug resistant (Resistant to the drugs Isoniazid and Rifampicin) tuberculosis patients admitted in a clinical trial at National Institute for Research in Tuberculosis, Chennai during the period 2002-2007. The total number of records created. The number of records have been created depends upon the number of recurring events met by each patient. The event of interest is sputum culture
conversion from positive to negative in every occasion. The five covariates considered for the models are age (in yrs), sex (Male = 1, Female = 0), Treat (0, 1), wt at baseline (kgs) and sensitivity. The age is classified into three groups viz. <30, 31-40 and ≥ 40. The treatments given to the patients are broadly classified into two categories based on Injection Kanamycin given daily and intermittent (thrice weekly) other five drugs common in both treatment. The other details can be found in the data layout is restructured as per the requirement of the models (Counting process (calendar time), Conditional A (Gap time), Conditional B (total time) and Marginal) for analysis purpose and shown in figure 4.1 and Table 4.1. The results are presented separately for the three age groups.
Table 4.1 Data layout for recurrent events (as given in Fig. 4.1)

<table>
<thead>
<tr>
<th>Model</th>
<th>Subject X</th>
<th></th>
<th>Subject Y</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time Interval</td>
<td>Event</td>
<td>Stratum (Interval #)</td>
<td>Time Interval</td>
</tr>
<tr>
<td>Counting Process</td>
<td>0,3</td>
<td>1</td>
<td>1</td>
<td>0,2</td>
</tr>
<tr>
<td></td>
<td>3,17</td>
<td>1</td>
<td>1</td>
<td>2,11</td>
</tr>
<tr>
<td></td>
<td>17,23</td>
<td>1</td>
<td>1</td>
<td>11,23</td>
</tr>
<tr>
<td></td>
<td>23,28</td>
<td>1</td>
<td>1</td>
<td>23,24</td>
</tr>
<tr>
<td></td>
<td>28,36</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gap time or Conditional A</td>
<td>0,3</td>
<td>1</td>
<td>1</td>
<td>0,2</td>
</tr>
<tr>
<td></td>
<td>3,17</td>
<td>1</td>
<td>2</td>
<td>2,11</td>
</tr>
<tr>
<td></td>
<td>17,23</td>
<td>1</td>
<td>3</td>
<td>11,23</td>
</tr>
<tr>
<td></td>
<td>23,28</td>
<td>1</td>
<td>4</td>
<td>23,24</td>
</tr>
<tr>
<td></td>
<td>28,36</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Conditional B</td>
<td>0,3</td>
<td>1</td>
<td>1</td>
<td>0,2</td>
</tr>
<tr>
<td></td>
<td>0,14</td>
<td>1</td>
<td>2</td>
<td>0,9</td>
</tr>
<tr>
<td></td>
<td>0,6</td>
<td>1</td>
<td>3</td>
<td>0,12</td>
</tr>
<tr>
<td></td>
<td>0,5</td>
<td>1</td>
<td>4</td>
<td>0,1</td>
</tr>
<tr>
<td></td>
<td>0,8</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>0,3</td>
<td>1</td>
<td>1</td>
<td>0,2</td>
</tr>
<tr>
<td></td>
<td>0,17</td>
<td>1</td>
<td>2</td>
<td>0,11</td>
</tr>
<tr>
<td></td>
<td>0,23</td>
<td>1</td>
<td>3</td>
<td>0,23</td>
</tr>
<tr>
<td></td>
<td>0,28</td>
<td>1</td>
<td>4</td>
<td>0,24</td>
</tr>
</tbody>
</table>
Modeling of this type of recurrent events data can be carried out by a Cox PH model, with two or more intervals for a same subject, the different lines of data contributed by the same subject are treated in the analysis as if they were independent contributions from different subjects, even though there are several outcomes on the same subject. In contrast, for the standard Cox PH model approach for nonrecurring survival data, different lines of data are treated as independent because they come from different subjects.

**Figure-4.1 Standard data layout for different models**

<table>
<thead>
<tr>
<th>Calendar Time or Counting Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap Time or Conditional A</td>
</tr>
<tr>
<td>Conditional B</td>
</tr>
<tr>
<td>Marginal</td>
</tr>
</tbody>
</table>

Time • Event ○ censored
Table 4.2 Mean and variance Estimates under different models

<table>
<thead>
<tr>
<th>Age in (yrs)</th>
<th>Counting Process</th>
<th>Gap time Model</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cova.</td>
<td>β Estim.</td>
<td>SE</td>
</tr>
<tr>
<td>&lt;30</td>
<td>treat</td>
<td>-0.20</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>sex</td>
<td>-0.24</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>wt</td>
<td>-0.009</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>dst</td>
<td>0.04</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL =491.80</td>
<td>AIC =499.80</td>
</tr>
<tr>
<td>31-40</td>
<td>treat</td>
<td>-0.03</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>sex</td>
<td>-0.35</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>wt</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>dst</td>
<td>0.14</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL =479.56</td>
<td>AIC =487.56</td>
</tr>
<tr>
<td>&gt;40</td>
<td>treat</td>
<td>0.23</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>sex</td>
<td>0.32</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>wt</td>
<td>-0.01*</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>dst</td>
<td>-0.03</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL =688.681</td>
<td>AIC =696.681</td>
</tr>
</tbody>
</table>

* p<0.05

among three models, the gap time model had lowest deviance in the age groups 30-40 and >40 and has deviance same as marginal model. Only weight had significant effect of the recurrent events.
Table 4.4 Exponential model with Gamma shared Frailty

<table>
<thead>
<tr>
<th>Age in (yrs)</th>
<th>Cov</th>
<th>Haz. Ratio</th>
<th>Std. Err</th>
<th>(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;30</td>
<td>Treat</td>
<td>2.08</td>
<td>0.69</td>
<td>1.089</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.94</td>
<td>0.32</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Wt</td>
<td>1.00</td>
<td>0.01</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Dst</td>
<td>1.04</td>
<td>0.32</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL=128.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31-40</td>
<td>Treat</td>
<td>1.65</td>
<td>0.66</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.83</td>
<td>0.31</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Wt</td>
<td>1.08*</td>
<td>0.02</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>Dst</td>
<td>0.66</td>
<td>0.28</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL=104.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;40</td>
<td>Treat</td>
<td>1.70</td>
<td>0.57</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.83</td>
<td>0.38</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>Wt</td>
<td>1.02</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Dst</td>
<td>0.62</td>
<td>0.20</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2LL=191.16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p<0.05

Except the exponential with gamma frailty model, for the other distributions weibull, lognormal and log-logistics the results were not derived due to the likelihood values not converged for my data. The exponential distribution with shared gamma model shows the significance for the covariate weight for the age group 31-40. The deviance -2LL for the age groups 31-40 is 104.91 and the deviance value is also less compared to the other ages. It indicates that this exponential model with gamma frailty for this age group yields better results compared to the age groups.
4.10 Summary

The purpose of this chapter is to give an overview of easily applicable statistical techniques that are available to analyze recurrent event data, to compare the results of each model with other models, and to give some recommendations on how to analyze recurrent event data. In general, standard hazard regression methods cannot be applied because of correlations between multivariate failures or recurrent event times within a subject. Adjustment is necessary for existing correlations, and more sophisticated analytic approaches are needed to obtain accurate estimates and efficient inferences. In the presence of the dependence between recurrent events times within a subject and subject-specific susceptibility across subjects, a variety of statistical methods have been proposed for the estimation of the covariate effect. The goal of robust estimation for the counting process approach is to obtain variance estimators that adjust for correlation within subjects when previously no such correlation was assumed. The robust estimator of the variance of an estimated regression coefficient allows tests of hypotheses and confidence interval estimation about model parameters to account for correlation within subjects. In survival analysis, multiplicative and additive hazards models provide the two principal frameworks for studying the association between risk factors and recurrent event durations for the analysis of multivariate failure time data. The majority of existing regression methods for analyzing multivariate failure or recurrent event time data assumes multiplicative covariate effects. Various authors have considered multivariate failure time models to be extensions of the Cox proportional hazards model. The multivariate model with a Markov assumption, the conditional approach, the marginal approach, and the random effects approach are among them. A widely used technique for adjusting for the correlation among outcomes on the same subject is called robust estimation.
The recurrent events may not be independent always for a same subject and the order of the events are important. Here the strata are the time interval numbers. The gap time or conditional A approach uses the exact same (start, stop) data layout format used for the counting process approach, except that for conditional A, an Stratified Cox (SC) model is used rather than a standard (uncertified) PH model. In this gap time model, the strata variable here is interval in this listing. The counting process approach to analysis typically used when recurrent events are treated as identical. The data taken from multi drug resistant TB patients trial was analyzed for two approaches are counting process and gap time modeling.

In overall comparison of the counting process, gap time and marginal model, the gap time model had lowest deviance in the age groups 30-40 and >40 and the next lowest deviance is from marginal model for the age group 31-40. Only weight had significant effect of the recurrent events. Hence for the MDR-TB recurrent data, the gap time model gives better results compared with the other models.

The parametric with shared gamma frailty models Weibull, Log Normal and Log Logistics likelihood values were not converged for my data. The exponential distributions with shared gamma model likelihood values converged and the results of this model shows the significance covariate weight for age group 31-40 and the smallest deviance -2LL value 104.90. It also shows the smallest variance \( \theta \) in the age group 31-40. It indicates that this exponential model with gamma frailty for this age group fits better compared to the age groups.