NTNU

Norwegian University of
Science and Technology

Modelling of Semi-Competing Risks Using the Illness-Death Model with Shared Frailty

Maria Lie Selle

Master of Science in Physics and Mathematics
Submission date: June 2016
Supervisor: Bo Henry Lindqvist, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem description

- Give an introduction to modelling and inference for semi-competing risks
- Study the illness-death model with shared frailty for dependent semi-competing risks
- Apply the model to both simulated and real data

Assignment given: January 15, 2016
Supervisor: Bo H. Lindqvist

Abstract

Semi-competing risks are a variation of competing risks where a terminal event censors a non-terminal event, but not vice versa. This thesis describes and studies modelling of semi-competing risks using the illness-death model with shared frailty suggested by Xu et al. (2010)[Biometrics, 66(3):716–725]. In their model the dependency between the terminal and non-terminal failure time is incorporated through the use of a shared frailty, which gives a model with conditional transition rates possessing the Markov property. We introduce the use of parametric models for the conditional transition rates and an expansion of the model where an additional terminal event is included.

Maximum likelihood estimation is performed to fit the model to data sets. First, a simulation study is carried out. Then the model is applied to two real data sets. The first data set contains observations of leukaemia patients after bone marrow transplantation, where the non-terminal event is relapse of the disease and the terminal event is death. For this data set we compare the use of a power law function and a log-linear law function as model for the conditional transition rates and find that relapse after bone marrow transplant for leukaemia patients is associated with increased probability of death. The second data set contains observations of patients admitted to hospital intensive care unit, where the non-terminal event is hospital-acquired pneumonia and there are two terminal events, alive discharge and death on the unit. For this data set we include an additional terminal state and covariates in the model, and we find that hospital-acquired pneumonia is associated with decreased rate of discharge from intensive care unit stay, while the probability of death increases as a consequence of prolonged stay. The method makes good estimates for the model parameters and by incorporating frailties, it is simple to construct a likelihood function, and to expand the model by adding more states. The interpretation of the marginal and conditional transition rates is different, which must be taken into account when interpreting the results. The frailty of each subject is usually not accessible. Nevertheless, both the marginal and conditional transition rates are of value and both should be considered in modelling of semi-competing risk.

Sammendrag

Semi-konkurrerende risikoer er en variant av konkurrerende risikoer der en terminerende hendelse kan sensurere en ikke-terminerende hendelse, men ikke omvendt. Denne avhandlingen beskriver og studerer modellering av semi-konkurrerende risikoer ved bruk av en "illness-death" modell med såkalt "shared frailty" foreslått av Xu et al. (2010)[Biometrics, 66 (3): 716-725]. I denne modellen blir avhengigheten mellom den terminerende hendelsen og den ikke-terminerende hendelsen innlemmet ved å bruke en "shared frailty". Dette gir en modell med betingede hasardrater med Markov-egenskap. Vi introduserer bruk av parametriske modeller for de betingede hasardratene og en utvidelse av modellen der en ekstra terminerende hendelse er inkludert.

Maksimal sannsynlighetsestimering blir brukt for å tilpasse modellen til ulike datasett. Først blir et simuleringsstudie gjennomført. Deretter blir modellen anvendt på to ekte datasett. Det første datasettet inneholder observasjoner fra leukemipasienter etter beinmargstransplantasjon, der den ikke-terminerende hendelsen er tilbakefall av sykdommen og den terminerende hendelsen er død. For dette datasettet sammenligner vi bruk av "power law" og "log-linear law" som modell for de betingede hasardratene. Vi finner at tilbakefall etter beinmargstransplantasjon for leukemipasienter er assosiert med økt sannsynlighet for død. Det andre datasettet inneholder observasjoner for pasienter innlagt på en intensivavdeling, hvor den ikke-terminerende hendelsen er lungebetennelse og det er to terminerende hendelser, utskriving fra intensivavdelingen og død. For dette datasettet inkluderer vi en ekstra terminerende hendelse og kovariater i modellen, og vi finner at lungebetennelse er assosiert med redusert rate for utskriving fra intensivavdelingen, mens sannsynligheten for død øker som følge av forlenget sykehusopphold. Metoden gir gode estimater for modellparametrene og ved å inkludere en "frailty" er det enkelt å konstruere sannsynlighetsmaksimeringsfunksjonen og å utvide modellen. Tolkningen av de marginale og betingede ratene er forskjellig og dette må tas i betraktning når man skal tolke resultater. Hvert enkelt subjekts "frailty" er vanligvis ikke er tilgjengelig. Likevel, er både de marginale og betingede hasardratene av verdi og begge bør vurderes i modellering av semi-konkurrerende risikoer

Preface

This Master's thesis concludes my Master of Science in Applied Physics and Mathematics with specialisation in Industrial mathematics at the Norwegian University of Science and Technology (NTNU). The work on this thesis has been carried out at the Department of Mathematical Sciences during my tenth and final semester as a student at NTNU.

During the past year I have received outstanding supervision from my supervisor professor Bo H. Lindqvist at the Department of Mathematical Sciences. I would like to thank him for all his inputs, ideas and helpful comments.

Maria Lie Selle,
Trondheim, June 2016

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In survival analysis the goal is usually to model the time until occurrence of an event of interest, often called a failure. In some situations there are more than one failure that can occur such that the event of interest is prevented from occurring. This leads to a form of missing data problem which makes modelling more complicated. The situation is called a competing risks situation, and today the theory of competing risks has applications in many fields including reliability and maintenance studies, medical research, demography, actuarial science, and in econometrics. An example of competing risks is in cancer where death due to cancer may be the event of interest, and death due to other causes such as surgical mortality or old age are competing risks.

In competing risks each observation consists of time to failure and the cause of the failure. The semi-competing risks situation is a generalisation of competing risks and usually one considers only two events, one terminal and one non-terminal. In this situation the terminal event censors the non-terminal event, while the occurrence of the non-terminal event does not prevent the terminal event from occurring. Because of this, each observation consists of either one or two failure times and which events that occurred. For example, in medical research the non-terminal event may be relapse of a disease and the terminal event may be death. Some of the essential questions in semi-competing risks are how is the effect of the non-terminal event on the terminal event, how to predict time to failure and how to estimate failure rates for specific causes.

Semi-competing risks have not yet become as common in literature as competing risks. It was first introduced by Fine et al. (2001) and was modelled by a copula model where it is assumed that the joint distribution of the terminal failure time and the non-terminal failure time is given by a copula, for example the gamma frailty copula. This model assumes latent failure times, also known as analysis by net quantities (Jiang et al., 2005), a representation in which only the failure time of the occurring event is identified, and the other potential failure times are not. Xu et al. (2010) argue that models involving latent failure times should be avoided. This is because they make untestable assumptions, the interpretation of the marginal distribution of the non-terminal event is hypothetical, and covariance analysis is complicated.

Xu et al. (2010) therefore suggest a model for semi-competing risks which avoids the use of latent failure times, and uses observable quantities instead, also known as analysis by crude quantities (Jiang et al., 2005). The model they suggest is an illness-death model with shared frailty. The illness-death model is a multi-state model, a model that allows subjects to move among a number of states over time. The functions of interest in a multi-state model are the transition rates which provide the instantaneous probabilities of transition from one state to another (Touraine et al., 2013). The illness-death model allows subjects to move among the so-called ‘health’, ‘illness’ and ‘death’ states, or state 0, 1 and 2 which we will use as state names. Subjects are initially healthy and then may become diseased and die, or die disease-free, which is the situation we want to model in semi-competing risks.

Furthermore, as the name of the model suggests, frailties are included. The frailty is an unobservable multiplicative effect on a hazard function. This means that subjects with a frailty higher than one will have a greater hazard of failure, whereas subjects with frailty lower than one will have a decreased hazard of failure (Gutierrez et al., 2002). In the shared-frailty illness-death model the frailty of each subject is shared between the subject’s transition rates creating a conditional model for the transition rates.

In this thesis we will investigate the illness-death model with shared frailty, by modelling semi-competing risks data. While Xu et al. (2010) use non-parametric models for the conditional transition rates, we introduce parametric models for the conditional transition rates and we also introduce an expansion of the model which allows for two competing terminal events. The model is tested on simulated data sets and two real data sets by using maximum likelihood estimation in R (R Development Core Team, 2008). The first data set contains observations of patients after a bone marrow transplantation as treatment for acute leukaemia where relapse of the cancer is the non-terminal event and death is the terminal event. The second data set contains information about patients admitted to an intensive care unit in hospitals where the non-terminal event is hospital-acquired pneumonia and there are two terminal events, death and discharge. With the last data set we have also included covariates in the models which is described by Xu et al. (2010).

The remainder of the thesis is organised as follows. In Section 2 we describe theory of competing risks and semi-competing risks. The illness-death model with shared frailty of Xu et al. (2010) is presented in Section 3, along with the parametric models and our expansion of the model. In Section 4 we present a simulation study with the illness-death model with shared frailty and the expanded model. The data analysis is presented in Section 5, and in Section 6 we make some concluding remarks about the model and some recommendations for further work. In Appendix A some more theory is described, including some basic theory of frailty models. Some derivations are included in Appendix B and in Appendix C the R-code for simulating semi-competing risks data and the log likelihood functions for most of the models are given.

Chapter 2

Theory

This chapter contains theory on standard survival analysis, competing risks and semi-competing risks. Section 2.1 and 2.2 are edited versions from Selle (2015) and contains theory from Lindqvist (2006) and Putter et al. (2007). In Appendix A we present an introduction to the use of frailties should it not be known to the reader, and we also present the likelihood ratio test as this is a test that will be used frequently throughout the thesis to compare model fits.

2.1 Survival analysis

In survival analysis one is interested in modelling the time until a specific event. This time is usually called the survival time, failure time or lifetime and is a random variable represented by T. The failure time will here be thought of as continuous with probability density function $f(t)$ and cumulative distribution function $F(t) = P(T \leq t)$, where $f(t) = \frac{d}{dt}F(t)$. The corresponding survival function is defined as $S(t) = P(T > t)$. The hazard function at time t, defined as

$$z(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t \,|\, T > t)}{\Delta t} = \frac{f(t)}{S(t)},$$

is the rate of failure at time t given that the failure time is larger than t. The cumulative hazard rate becomes $Z(t) = \int_0^t z(u)\, \mathrm{d}u$.

In survival analysis, some failure times can be censored, meaning that some of the failure times are only known to have occurred within a certain time interval. The failure times can be right-, left- or interval-censored. In administrative censoring, subjects who have not experienced failure beyond the closing of a study will be censored. This date is fixed such that the censoring time will be independent from the failure time.

A non-parametric estimator for $S(t)$ is the Kaplan-Meier estimator. Let the ith individual have potential failure time T_i and potential censoring time C_i for $i = 1, ..., k$. If $T_1, ..., T_k$ are assumed to be independent and identically distributed with survival function $S(t)$ and the censoring distribution is independent from the failure times, then the Kaplan-Meier estimator is the following

$$\hat{S}(t) = \prod_{i:T_{(i)} \leq t} \frac{n_i - d_i}{n_i}$$

where $T_{(1)} < T_{(2)} < \cdots$ are the times with at least one failure, n_i is the number at risk at $T_{(i)}$, and d_i is the number failing at $T_{(i)}$.

In some cases, there exists information from covariates which may help explain the failure times or predict why some units fail quickly and some units survive a long time. The failure time distribution can be related to covariates by regression models. A well known model proposed by Cox is $z(t; \mathbf{x}) = z_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x})$. This is known as the proportional hazards model, or Cox model, and $z_0(t)$ is called the baseline hazard function. This can be any positive function of t. The regression parameter $\boldsymbol{\beta}$ can be estimated by maximizing the log partial likelihood of $z(t; \mathbf{x})$ which gives maximum partial likelihood estimators. The cumulative baseline hazard function can be estimated by the Breslow estimator

$$\hat{Z}_0(t) = \sum_{T_{(j)} \leq t} \frac{1}{\sum_{j \in R_j} \exp(\hat{\boldsymbol{\beta}}^T \mathbf{x}_i)},$$

where R_j is the risk set at time $T_{(j)}$, and $\hat{\boldsymbol{\beta}}$ is the Cox partial likelihood estimator of $\boldsymbol{\beta}$.

2.2 Competing risks

In competing risks one observes a pair (T, ϵ), where T is the failure time and $\epsilon \in \{1, 2, ..., N\}$ is the cause of failure. When studying time to failure from a specific cause, failures from other causes are competing events. Since there may be many causes of failure, one can consider the observed T as the smallest of several latent failure times $\{T_1, T_2, ..., T_k\}$, where $T = \min_j T_j$ and $\epsilon = e$ if $T = T_e$. This is called the latent failure time representation.

2.2.1 The cumulative incidence function

The joint distribution of (T, ϵ) is specified by sub-distribution functions, or cumulative incidence functions as they are also called,

$$F_j(t) = P(T \leq t, \epsilon = j), \tag{2.1}$$

which are defined for $t > 0, j \in \{1, 2, ..., N\}$. By differentiation the sub-density functions become

$$f_j(t) = F_j'(t).$$

The marginal distribution of T is given by

$$F(t) = P(T \leq t) = \sum_{j=1}^{k} F_j(t).$$

This can be expressed by the survival function, $S(t) = 1 - F(t)$, which can be interpreted as the probability of not having failed from any cause at time t.

2.2.2 The cause-specific hazard function

The distribution of (T, ϵ) can also be specified by sub-hazard functions, also known as cause-specific hazard functions. The interpretation of the cause-specific hazard function is that it is the failure rate of cause j at time t in the presence of the other failure causes, given that the lifetime T is greater than t. The cause-specific hazard function for cause j is defined as

$$\lambda_j(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t, \epsilon = j \mid T > t)}{\Delta t} = \frac{f_j(t)}{S(t)}, \tag{2.2}$$

where $S(t)$ is the survival function of T. The overall hazard function of T then becomes

$$\lambda(t) = \sum_{j=1}^{N} \lambda_j(t).$$

A useful connection is

$$F_j(t) = \int_0^t \lambda_j(u) S(u) \, \mathrm{d}u, \tag{2.3}$$

which is the cumulative incidence function in terms of the cause-specific hazard. The cumulative cause-specific hazard functions are defined as

$$\Lambda_j(t) = \int_0^t \lambda_j(u) \, \mathrm{d}u.$$

Using this, the cumulative hazard function of T is $\Lambda(t) = \sum_{j=1}^{N} \Lambda_j(t)$. We now have the following relationship

$$S(t) = e^{-\Lambda(t)} = e^{-\sum_{j=1}^{N} \Lambda_j(t)}. \tag{2.4}$$

2.2.3 The identifiability problem

In a competing risks analysis, one is often interested in the joint and marginal distributions of the latent failure times $T_1, T_2, ..., T_k$. But in general, the distributions of the latent failure times are not identifiable from the distribution of the observable pair (T, ϵ). This is because there are many different joint distributions of $T_1, T_2, ..., T_k$, which can result in the same distribution of (T, ϵ).

It has been found (Tsiatis, 1975) that if the set of cumulative incidence functions $F_j(t)$ is given for some model with dependent risks, there exists a unique model, a so-called independent-risks proxy model, with independent risks which gives rise to the same $F_j(t)$. Thus, one cannot know which of the two models is correct from only the observations of (T, ϵ).

2.3 Semi-competing risks

The semi-competing risks problem was first introduced by Fine et al. (2001) and is a variation of competing risks. The problem refers to a situation where a subject can experience two types of failures, where one of the failure times censors the other but not vice versa. The censoring failure is referred to as a terminal event or terminal failure, while the other is referred to as the non-terminal event or non-terminal failure. A non-terminal event can for instance be relapse of some disease, while a terminal event can be death. A patient that has a relapse after treatment can die, while a patient dying cannot have a relapse after death. An overall independent censoring of all failure times is usually included in the semi-competing risks problem. A simple semi-competing risks situation is presented in Figure 2.1. State 0 is the initial state, state 1 is a transient state where the non-terminal event has occurred and state 2 is the state where the terminal-event has occurred.

Figure 2.2, adapted from Jiang et al. (2005), illustrates how semi-competing risks data compares to bivariate right-censored data and right-censored competing risks data. For the middle graph T_1 is the time until the non-terminal event and T_2 is the time until the terminal event. For the right-censored bivariate data in the

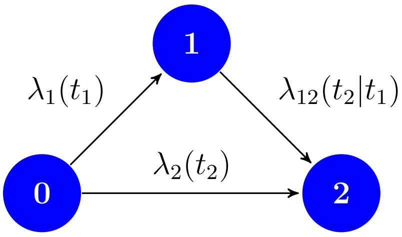


Figure 2.1: A semi-competing risks situation. The hazard rate, also known as transition rate, for each transition is included.

left figure both failure times can be observed, while for the ordinary competing risks data in the right figure, T_1 and T_2 cannot be observed together and the failure times can only be observed along the diagonal line. We note that for semi-competing risks data both failure times can be observed as long as T_1 is observed before T_2 such that the observations are restricted to $t_2 \geq t_1$, which is called the upper wedge (Xu et al., 2010; Jiang et al., 2005).

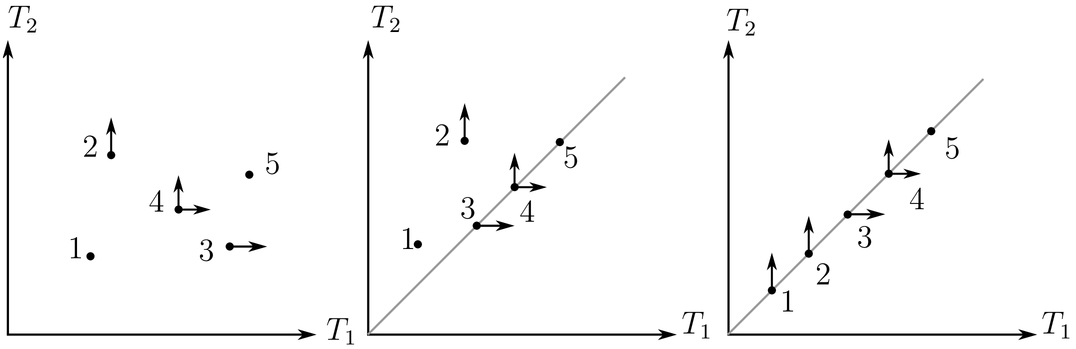


Figure 2.2: This figure illustrates possible observations in bivariate data (left), semi-competing risks data (middle) and competing risks data (right). A dot indicates that both T_1 and T_2 has been observed, and an arrow indicates censoring of the failure time in the pointing direction.

Table 2.1: Possible orderings of the events and the four observable cases they lead to.

Order	Y_1	Y_2	δ_1	δ_2	Case
T_1, T_2, C	T_1	T_2	1	1	1
T_2, T_1, C	T_2	T_2	0	1	2
T_2, C, T_1	T_2	T_2	0	1	2
T_1, C, T_2	T_1	C	1	0	3
C, T_1, T_2	C	C	0	0	4
C, T_2, T_1	C	C	0	0	4

2.3.1 Notation

In this thesis we will mostly consider semi-competing risks situations with two events, one terminal and one non-terminal. The time until the non-terminal event and the time until the terminal event are T_1 and T_2 respectively, and the censoring time is represented by C. The observation from each subject will be $(Y_1, Y_2, \delta_1, \delta_2)$, where $Y_1 = \min(T_1, Y_2), Y_2 = \min(T_2, C), \delta_1 = I(T_1 \leq Y_2)$ and $\delta_2 = I(T_2 \leq C)$. Here, I is the indicator function.

To better understand this notation we have set up the different orderings of the events in Table 2.1 and which observable cases they lead to. We see that Y_1 refers to the first observed event time and δ_1 indicates if this was T_1 or not. Y_2 refers to what was observed after Y_1 or the same as Y_1 and δ_2 indicates whether T_2 was censored or not. In the derivation of the likelihood function in Appendix B.2 the cases are explained in greater detail. The real data sets studied later in this thesis are transformed into the format described here.

Chapter 3

The illness-death model with shared frailty

An illness-death model is a multi-state model that is much used in the medical literature to describe disease progression (Meira-Machado et al., 2008), and it was first described by Fix and Neyman (1951). It can model a situation where there is one terminal event and one non-terminal event, as the situation in Figure 2.1. This means that each subject will have three possible transition rates. It can either have a transition directly to the terminal event or it can have a transition to the non-terminal event first and then to the terminal event.

We shall consider an illness-death model where the hazard rates, or transition rates as they will be referred to, of each subject have a shared frailty. This model is called the illness-death model with shared frailty. We will, for simplicity, from now on also refer to this model as the illness-death model although we mean the illness-death model with shared frailty. The model has been described and studied by Xu et al. (2010), and in this section the model will be described. In addition, parametric models for the transition rates and an expansion of the model will be introduced.

3.1 Transition rates

Recall the notation from Section 2.3.1, where T_1 is the non-terminal event time, T_2 is the terminal event time and C is an independent censoring time. From Figure 2.2 we have already seen that in semi-competing risks the observation of the non-terminal event is only available if $t_2 \geq t_1$, in other words if it occurs before the terminal event. If a subject experiences a terminal failure before the non-terminal failure has occurred, we define $T_1 = \infty$. This means that T_1 is always defined even though the terminal event occurs before the non-terminal event.

Furthermore, recall the semi-competing risks situation and the states presented in Figure 2.1. The illness-death model is completely defined by the transition rates, and these are

$$\begin{aligned}\lambda_1(t_1) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_1 \le T_1 < t_1 + \Delta t | T_1 \ge t_1, T_2 \ge t_1),\\ \lambda_2(t_2) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \le T_2 < t_2 + \Delta t | T_1 \ge t_2, T_2 \ge t_2),\\ \lambda_{12}(t_2|t_1) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \le T_2 < t_2 + \Delta t | T_1 = t_1, T_2 \ge t_2),\end{aligned} \tag{3.1}$$

where $0 < t_1 < t_2$. Here $\lambda_1(t_1)$ is the transition rate from state 0 to state 1 and $\lambda_2(t_2)$ is the transition rate from state 0 to state 2. In fact $\lambda_1(t_1)$ and $\lambda_2(t_2)$ are the same as the cause-specific hazards in competing risks. The transition rate from state 1 to state 2, $\lambda_{12}(t_2|t_1)$, can in general depend on both t_1 and t_2. If $\lambda_{12}(t_2|t_1)$ depends only on t_2 it is called a Markov model since the future and past are independent given the present. That is, the transition rate from state 1 to state 2 is independent of the time state 1 was reached, which makes $\lambda_{12}(t_2|t_1) = \lambda_{12}(t_2)$. The Markov model is most frequently used because of its simplicity (Meira-Machado et al., 2008). If $\lambda_{12}(t_2|t_1)$ depends on the time since state 1 was reached, $t_2 - t_1$, it is a semi-Markov model.

A dependent structure between T_1 and T_2 can be incorporated by using a shared frailty, denoted by γ. For some more background theory about shared frailty models the reader is referred to Appendix A.1 and Gutierrez et al. (2002). In this model it is not the subjects that share frailty, but the conditional transition rates for each subject have the same frailty. The conditional transition rates corresponding to (3.1) are defined as

$$\begin{aligned}\lambda_1(t_1|\gamma) &= \gamma\lambda_{01}(t_1),\ t_1 > 0\\ \lambda_2(t_2|\gamma) &= \gamma\lambda_{02}(t_2),\ t_2 > 0\\ \lambda_{12}(t_2|t_1, \gamma) &= \gamma\lambda_{03}(t_2),\ 0 < t_1 < t_2\end{aligned} \tag{3.2}$$

When $\lambda_{02}(t_2)$ and $\lambda_{03}(t_2)$ are arbitrary functions, this frailty model is described as the ‘general model’, and if $\lambda_{02}(t_2) = \lambda_{03}(t_2)$ it is referred to as the ‘restricted model’. Since $\lambda_{12}(t_2|t_1, \gamma)$ in (3.2) is independent of the time state 1 was reached, the conditional transition rates are Markovian. But as we will see, the corresponding marginal transition rates are not Markovian because of the dependent structure between T_1 and T_2 incorporated by γ.

The frailty is assumed to follow a Gamma distribution with expectation 1 and variance θ. This gives the following distribution function for γ

$$g(\gamma; 1/\theta, \theta) = \frac{1}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})}\gamma^{\frac{1}{\theta}-1}e^{-\frac{\gamma}{\theta}}, \ \gamma \geq 0 \tag{3.3}$$

We will refer to the distribution in (3.3) as $g(\gamma)$. Since a distribution family for γ has been assumed, models for the marginal transition rates can be derived. From Xu et al. (2010) these are

$$\lambda_1(t_1) = (1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_1)])^{-1}\lambda_{01}(t_1), \ t_1 > 0 \tag{3.4}$$

$$\lambda_2(t_2) = (1 + \theta[\Lambda_{01}(t_2) + \Lambda_{02}(t_2)])^{-1}\lambda_{02}(t_2), \ t_2 > 0 \tag{3.5}$$

$$\lambda_{12}(t_2|t_1) = (1 + \theta)(1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_2) + \Lambda_{03}(t_1, t_2)])^{-1}\lambda_{03}(t_2), \tag{3.6}$$

$$0 < t_1 < t_2,$$

where $\Lambda_{0i}(t)$ for $i = 1, 2, 3$, are the cumulative conditional transition rates, and $\Lambda_{03}(s, t) = \Lambda_{03}(t) - \Lambda_{03}(s)$. This definition will be used throughout the thesis. Our derivation of the marginal transition rates from the conditional transition rates can be found in Appendix B.1. The marginal transition rate from state 1 to state 2, $\lambda_{12}(t_2|t_1)$, in (3.6) depends on both t_1 and t_2 and is therefore not Markovian as opposed to its corresponding conditional transition rate, unless γ is constant which makes $\theta = 0$.

The dependence of T_2 on T_1 can be described by two measures. Either by the common frailty γ, or by the conditional explanatory hazard ratio which is $\lambda_{03}(t_2)/\lambda_{02}(t_2)$ when we have a Markov model for the conditional transition rates. The explanatory hazard ratio describes how the risk of event 2 changes over time given that event 1 has occurred (Lee et al., 2015a). With a semi-Markov it is not obvious how to interpret the conditional explanatory hazard ratio since $\lambda_{02}(t_2)$ and $\lambda_{03}(t_2 - t_1)$ are of different time scales. For the restricted model, the conditional explanatory hazard ratio is equal to 1 since the conditional transition rates to state 2 are the same. Therefore the dependence of T_2 on T_1 is fully captured by γ.

The main difference between marginal and conditional models is that marginal models are population-average models, while conditional models are subject-specific (Lee et al., 2004), or frailty-specific in our case. Since not all subjects have the same frailty, the conditional transition rates are only comparable within subjects sharing frailty. This causes the interpretation of the marginal and conditional transition rates to be different. We will sometimes, for simplicity refer to the marginal transition rates as the transition rates, while the conditional transition rates will always be referred to as the conditional transition rates.

3.1.1 Parametric models for the transition rates

Now, we introduce two parametric models for the conditional transition rates. The first parametric model is the power law and the conditional transition rates have the following parametric model

$$\lambda_{0i}(t) = \alpha_i \beta_i t^{\beta_i - 1}, i = 1, 2, 3, \tag{3.7}$$

where $\alpha_i, \beta_i > 0$, and the cumulative conditional transition rates will be on the form

$$\Lambda_{0i}(t) = \alpha_i t^{\beta_i}, i = 1, 2, 3.$$

Another choice for parametric model is a log-linear law. In this case the conditional transition rates will have parametric model

$$\lambda_{0i}(t) = e^{a_i + b_i t}, i = 1, 2, 3, \tag{3.8}$$

where $-\infty < a_i, b_i < \infty$, and the cumulative conditional transition rates will be on the form

$$\Lambda_{0i}(t) = \frac{e^{a_i}}{b_i}(e^{b_i} - 1), i = 1, 2, 3.$$

3.2 The likelihood functions

Recall the notation for the observed data $(Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2})$, $i = 1, ..., n$ from Section 2.3.1, and let n be the number of subjects. The likelihood function for the general illness-death model is based on the conditional transition rates and is as follows

$$\begin{aligned} L_g = \prod_{i=1}^{n} & \lambda_{01}(Y_{i1})^{\delta_{i1}} \lambda_{02}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})} \lambda_{03}(Y_{i2})^{\delta_{i1}\delta_{i2}} (1+\theta)^{\delta_{i1}\delta_{i2}} \\ & \Big(1 + \theta[\Lambda_{01}(Y_{i1}) + \Lambda_{02}(Y_{i1}) + \Lambda_{03}(Y_{i1}, Y_{i2})]\Big)^{-1/\theta - \delta_{i1} - \delta_{i2}}. \end{aligned} \tag{3.9}$$

The likelihood function is found by first finding the conditional contributions from each case and then integrating out γ using the distribution function $g(\gamma)$. A simple derivation can be found in Appendix B.2.

The likelihood function for the restricted model based on the observed data $(Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2})$ is found by letting $\lambda_{02} = \lambda_{03}$, and is

$$L_r = \prod_{i=1}^{n} \lambda_{01}(Y_{i1})^{\delta_{i1}} \lambda_{02}(Y_{i2})^{\delta_{i2}} (1+\theta)^{\delta_{i1}\delta_{i2}} \Big(1 + \theta[\Lambda_{01}(Y_{i1}) + \Lambda_{02}(Y_{i2})]\Big)^{-1/\theta - \delta_{i1} - \delta_{i2}}. \tag{3.10}$$

3.3 Probability functions in the illness-death model

In this section we present some of the probability functions used later in the thesis. The conditional survival functions are used for simulating semi-competing risks data which is presented in Section 4.1. The corresponding marginal survival functions are used in Section 5.1 to visualise parts of the results. Lastly, the marginal transition probabilities are used in Section 5.2, also to visualise results from the data analysis.

3.3.1 Conditional survival functions

In the next chapter we will simulate from the illness-death model. To do this we need to know the probability of staying in a specific state until time t. Recall the relation $S(t) = e^{-\sum_{j=1}^{N} \Lambda_j(t)}$ from equation (2.4) in the section about competing risks. The same relation can be found for semi-competing risks in the illness-death model using the conditional transition rates defined in (3.2). Then the conditional probability of staying in state 0 until time t is

$$S(t|\gamma) = P(\text{Still in state 0 at } t|\gamma) = e^{-\gamma(\Lambda_{01}(t) + \Lambda_{02}(t))}, \tag{3.11}$$

where $\Lambda_{01}(t)$ and $\Lambda_{02}(t)$ are the cumulative transition rates out of state 0. If there is a transition to state 1 at t_1, the conditional probability of staying in state 1 until time t_2 is

$$\begin{aligned} S_{12}(t_2|t_1, \gamma) &= P(\text{Still in state 1 at } t_2|\text{Transition to state 1 at time } t_1) \\ &= \exp\left(-\gamma \int_{t_1}^{t_2} \lambda_{03}(u)\,\mathrm{d}u\right) = e^{-\gamma \Lambda_{03}(t_1, t_2)} \end{aligned} \tag{3.12}$$

3.3.2 Marginal survival functions

Using the known distribution family of γ, which is $g(\gamma)$ from (3.3), the marginal probability function for staying in state 0 is found by averaging $S(t|\gamma)$ over γ and using the relation in equation (B.1) to solve the integral. This gives

$$S(t) = P(\text{Still in state 0 at } t) = (1 + \theta[\Lambda_{01}(t) + \Lambda_{02}(t)])^{-1/\theta} \tag{3.13}$$

A function that will be used in Section 5.1 is the marginal survival function for the time to the non-terminal event, $S_1(t)$. According to Xu et al. (2010) this is given by

$$S_1(t) = (1 + \theta\Lambda_{01}(t))^{-\frac{1}{\theta}}, \tag{3.14}$$

which is what we get if we let $\Lambda_{02}(t)$ in (3.13) be zero.

Furthermore, we compute the marginal survival function for time to the terminal event given the non-terminal event has occurred, $S_{12}(t_2|t_1)$. This is done by averaging $S_{12}(t_2|t_1, \gamma)$ in equation (3.12) over all values of γ

$$S_{12}(t_2|t_1) = \int_0^\infty \exp\left(-\gamma\Lambda_{03}(t_1, t_2)\right) g(\gamma)\, \mathrm{d}\gamma,$$

and by using the relation in equation (B.1), the survival function for time until the terminal event given transition to the non-terminal event at t_1 becomes

$$S_{12}(t_2|t_1) = (1 + \theta[\Lambda_{03}(t_1, t_2)])^{-1/\theta} \tag{3.15}$$

3.3.3 Marginal transition probabilities

In this section we describe the marginal transition probabilities in the illness-death model, $P_{01}(t)$, $P_{02}(t)$ and $P_{12}(t_2|t_1)$. The marginal transition probability $P_{01}(t)$ is the probability of being in state 1 at time t given that the previous state 0 was entered at time 0. Moreover, the marginal transition probability $P_{02}(t)$ is the probability of being in state 2 at time t given that the previous state was state 0 and that this was entered at time 0. Note that $P_{02}(t)$ is usually defined such that the process can have had a transition to state 1 in between state 0 and state 2. Finally, $P_{12}(t_2|t_1)$ is the probability of being in state 2 at time t_2 given that the previous state was 1 and this was entered at time t_1. Note that in this context, t_2 is not the time for the occurrence of the terminal failure, but is instead a time for which we are interested in the probability of being in state 2.

First, the derivation of $P_{01}(t)$ is presented. The probability of being in state 1 at time t is given by the probability of having a transition to state 1 before t

and staying in 1 until time t. One must also multiply with the probability of not having had a transition out of state 0 before the transition of interest. We use the conditional rates and must therefore average over γ. Since the time of transition to state 1 is unknown we must also integrate over all t. This gives

$$P_{01}(t) = \int_0^\infty \int_0^t e^{-\gamma(\Lambda_{01}(s)+\Lambda_{02}(s))}\gamma\lambda_{01}(s)e^{-\gamma\Lambda_{03}(s,t)}g(\gamma)\,\mathrm{d}s\mathrm{d}\gamma,$$

that is, the probability of staying in state 0 until time s, having a transition at time s to state 1, and staying in state 1 until time t. By solving the outer integral the transition probability becomes

$$P_{01}(t) = \int_0^t \lambda_{01}(s)(1+\theta[\Lambda_{01}(s)+\Lambda_{02}(s)+\Lambda_{03}(s,t)])^{-1/\theta-1}\,\mathrm{d}s \tag{3.16}$$

The derivation of $P_{02}(t)$ is done in the same manner and gives

$$P_{02}(t) = \int_0^t \lambda_{02}(s)(1+\theta[\Lambda_{01}(s)+\Lambda_{02}(s)])^{-1/\theta-1}\,\mathrm{d}s \tag{3.17}$$

In this function we do not have a contribution for staying in state 2 until time t, since state 2 is absorbing.

We move over to the transition probability $P_{12}(t_2|t_1)$, which is conditional on that state 1 was reached at time t_1. Then the probability of being in state 2 at time t_2 is given by the probability of staying in state 1 until some time s and having a transition to state 2 at time $s \leq t_2$. Again, since the time of transition from state 1 to state 2 is unknown we integrate over possible transition times. Using conditional transition rates and averaging over γ, this gives

$$P_{12}(t_2|t_1) = \int_{t_1}^{t_2}\int_0^\infty \gamma\lambda_{03}(s)e^{-\gamma\Lambda_{03}(t_1,s)}g(\gamma)\,\mathrm{d}\gamma\mathrm{d}s$$

After integrating out γ one gets

$$P_{12}(t_2|t_1) = \int_{t_1}^{t_2} \lambda_{03}(s)(1+\theta\Lambda_{03}(t_1,s))^{-1-1/\theta}\,\mathrm{d}s \tag{3.18}$$

The integrals in equations (3.16), (3.17) and (3.18) will be solved numerically.

3.4 Covariates in the illness-death model

Covariates are explanatory variables that may affect the survival times, and in this model we assume that the effect of the covariates is the same at all times. It is simple to include covariates in the shared frailty illness-death model by incorporating them in the conditional transition rates. Let $\boldsymbol{x} = (x_1, x_2, ..., x_p)^T$ be a vector of p covariates, and let $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \boldsymbol{\varphi}_3$ be three vectors of coefficients with length p. Then the conditional transition rates with covariates become

$$\begin{aligned}
\lambda_1(t_1|\gamma, \boldsymbol{x}) &= \gamma\lambda_{01}(t_1)\exp(\boldsymbol{\varphi}_1^T\boldsymbol{x}),\ t_1 > 0\\
\lambda_2(t_2|\gamma, \boldsymbol{x}) &= \gamma\lambda_{02}(t_2)\exp(\boldsymbol{\varphi}_2^T\boldsymbol{x}),\ t_2 > 0\\
\lambda_{12}(t_2|t_1, \gamma, \boldsymbol{x}) &= \gamma\lambda_{03}(t_2)\exp(\boldsymbol{\varphi}_3^T\boldsymbol{x}),\ 0 < t_1 < t_2,
\end{aligned}$$

which are the conditional transition rates in equation (3.2) where each rate has been multiplied by a transition-specific covariate term. The corresponding marginal transition rates with covariates are the same as in equations (3.4) - (3.6) where λ_{01} and Λ_{01} are multiplied with $\exp(\boldsymbol{\varphi}_1^T\boldsymbol{x})$, λ_{02} and Λ_{02} are multiplied with $\exp(\boldsymbol{\varphi}_2^T\boldsymbol{x})$, and λ_{03} and Λ_{03} are multiplied with $\exp(\boldsymbol{\varphi}_3^T\boldsymbol{x})$.

Another modelling strategy would be to include the covariates in the marginal transition rates by multiplying the transition rate in (3.4) with $\exp(\boldsymbol{\varphi}_1^T\boldsymbol{x})$, the transition rate in (3.5) with $\exp(\boldsymbol{\varphi}_2^T\boldsymbol{x})$ and the transition rate in (3.6) with $\exp(\boldsymbol{\varphi}_3^T\boldsymbol{x})$. This would give covariates with a different interpretation since the regression effect on these marginal transition rates would not be the same as the conditional approach described first. However, with the conditional approach it is much simpler to construct the likelihood function.

The likelihood function in the general model with covariates is presented in equation (3.19) below. If we in addition to observing $(Y_{i1}, Y_{i2}, \delta_{i1}, \delta_{i2})$ for each subject, also observe one vector of covariates $\boldsymbol{x}_i$ for each subject, the likelihood function becomes the following

$$\begin{aligned}
L_{cov} = \prod_{i=1}^{n} &\lambda_{01}(Y_{i1})^{\delta_{i1}}\lambda_{02}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})}\lambda_{03}(Y_{i2})^{\delta_{i1}\delta_{i2}}\\
&\cdot \exp[\delta_{i1}\boldsymbol{\varphi}_1^T\boldsymbol{x}_i + \delta_{i2}(1-\delta_{i1})\boldsymbol{\varphi}_2^T\boldsymbol{x}_i + \delta_{i1}\delta_{i2}\boldsymbol{\varphi}_3^T\boldsymbol{x}_i](1+\theta)^{\delta_{i1}\delta_{i2}} \qquad (3.19)\\
&\cdot \Big(1 + \theta[\Lambda_{01}(Y_{i1})e^{\boldsymbol{\varphi}_1^T\boldsymbol{x}_i} + \Lambda_{02}(Y_{i1})e^{\boldsymbol{\varphi}_2^T\boldsymbol{x}_i} + \Lambda_{03}(Y_{i1}, Y_{i2})e^{\boldsymbol{\varphi}_3^T\boldsymbol{x}_i}]\Big)^{-1/\theta-\delta_{i1}-\delta_{i2}}
\end{aligned}$$

Note that if we let the coefficients of the covariates be 0, we get the likelihood function for the general model presented in equation (3.9).

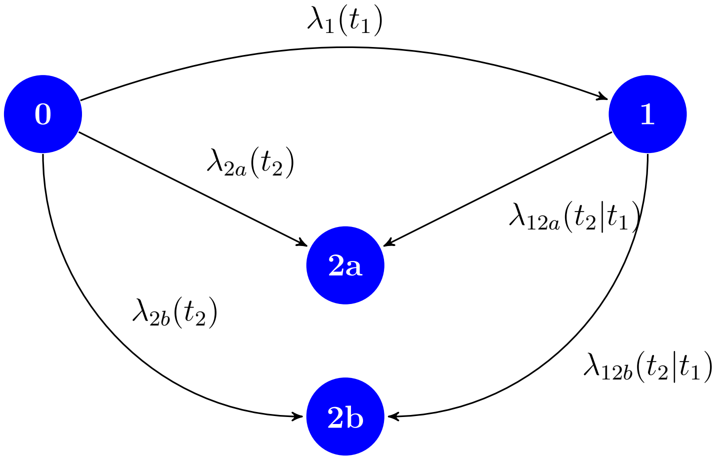


Figure 3.1: A semi-competing risks situation where the terminal failure consists of two separate events, $2a$ and $2b$. The transition rates for each transition are included.

3.5 Expansion of the illness-death model

In this section we introduce an expansion of the illness-death model with shared frailty, by letting the terminal event consist of two competing events. We still let T_1 be the time to the non-terminal event and T_2 be the time to the terminal event. In Figure 3.1 the expanded semi-competing risks situation is presented. The terminal event now consists of two events, event $2a$ and event $2b$, and the model now consists of five transition rates rather than three. The marginal transition rates are defined as

$$\begin{aligned}
\lambda_1(t_1) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_1 \leq T_1 < t_1 + \Delta t | T_1 \geq t_1, T_2 \geq t_1), \\
\lambda_{2a}(t_2) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \leq T_2 < t_2 + \Delta t, \mathcal{J} = a | T_1 \geq t_2, T_2 \geq t_2), \\
\lambda_{2b}(t_2) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \leq T_2 < t_2 + \Delta t, \mathcal{J} = b | T_1 \geq t_2, T_2 \geq t_2), \\
\lambda_{12a}(t_2|t_1) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \leq T_2 < t_2 + \Delta t, \mathcal{J} = a | T_1 = t_1, T_2 \geq t_2), \\
\lambda_{12b}(t_2|t_1) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t_2 \leq T_2 < t_2 + \Delta t, \mathcal{J} = b | T_1 = t_1, T_2 \geq t_2),
\end{aligned}$$

where $0 < t_1 < t_2$, and $\mathcal{J} \in \{a, b\}$ indicates which event was the terminal event. Furthermore, the corresponding conditional transition rates become

$$\begin{aligned}
&\lambda_1(t_1|\gamma) = \gamma\lambda_{01}(t_1),\ t_1 > 0\\
&\lambda_{2a}(t_2|\gamma) = \gamma\lambda_{02a}(t_2),\ t_2 > 0\\
&\lambda_{2b}(t_2|\gamma) = \gamma\lambda_{02b}(t_2),\ t_2 > 0\\
&\lambda_{12a}(t_2|t_1, \gamma) = \gamma\lambda_{03a}(t_2),\ 0 < t_1 < t_2\\
&\lambda_{12b}(t_2|t_1, \gamma) = \gamma\lambda_{03b}(t_2),\ 0 < t_1 < t_2
\end{aligned}$$

The likelihood function for the expanded model is given in equation (3.20) below. We have defined two indicator functions δ_{ia} and δ_{ib}, one for each terminal event. These indicate which event was the terminal event by having the value 1 if the event occurred and 0 if the event did not occur. If there was censoring such that neither of the terminal events occurred, the values of these indicators are of no interest and can be set to be 0 since they will not contribute to the likelihood function. Using the conditional transition rates, the likelihood for the expanded model is

$$\begin{aligned}
L_{exp} = \prod_{i=1}^{n} &\lambda_{01}(Y_{i1})^{\delta_{i1}}\lambda_{02a}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})\delta_{ia}}\lambda_{02b}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})\delta_{ib}}\lambda_{03a}(Y_{i2})^{\delta_{i1}\delta_{i2}\delta_{ia}}\\
&\cdot\lambda_{03b}(Y_{i2})^{\delta_{i1}\delta_{i2}\delta_{ib}}(1+\theta)^{\delta_{i1}\delta_{i2}}\Big(1+\theta[\Lambda_{01}(Y_{i1})+\Lambda_{02a}(Y_{i1})+\Lambda_{02b}(Y_{i1}) \qquad (3.20)\\
&+\Lambda_{03a}(Y_{i2})-\Lambda_{03a}(Y_{i1})+\Lambda_{03b}(Y_{i2})-\Lambda_{03b}(Y_{i1})]\Big)^{-1/\theta-\delta_{i1}-\delta_{i2}}
\end{aligned}$$

It is possible to include covariates in the expanded illness-death model in the same way as presented in Section 3.4. The conditional transition rates in the expanded likelihood function are then multiplied with covariate specific terms $\exp(\boldsymbol{\varphi}_i^T\boldsymbol{x})$.

The marginal transition probabilities in the model with only one terminal event, $P_{01}(t)$, $P_{02}(t)$, and $P_{12}(t_2|t_1)$, have already been presented in Section 3.3.3. The same approach is used when we find the transition probabilities in the expanded model which are $P_{01}(t), P_{02a}(t), P_{02b}(t), P_{12a}(t_2|t_1)$ and $P_{12b}(t_2|t_1)$.

For the probability of being in state 1 at time t we have

$$P_{01}(t) = \int_0^t \lambda_{01}(s)(1+\theta[\Lambda_{01}(s)+\Lambda_{02a}(s)+\Lambda_{02b}(s)+\Lambda_{03a}(s,t)+\Lambda_{03b}(s,t)])^{-1/\theta-1}\,\mathrm{d}s,$$

where $\Lambda_{02}(s)$ in equation (3.11) has been replaced by $\Lambda_{02a}(s)+\Lambda_{02b}(s)$ and $\Lambda_{03}(s,t)$ has been replaced with $\Lambda_{03a}(s,t) + \Lambda_{03b}(s,t)$. Similarly the probabilities of being in state $2a$ and state $2b$ at time t without having entered state 1 are

$$P_{02a}(t) = \int_0^t \lambda_{02a}(s)(1+\theta[\Lambda_{01}(s)+\Lambda_{02a}(s)+\Lambda_{02b}(s)])^{-1/\theta-1}\,\mathrm{d}s,$$

and

$$P_{02b}(t) = \int_0^t \lambda_{02b}(s)(1+\theta[\Lambda_{01}(s)+\Lambda_{02a}(s)+\Lambda_{02b}(s)])^{-1/\theta-1}\,\mathrm{d}s.$$

For the probability to be in state $2a$ and $2b$ at time t_2 given that state 1 was reached at time t_1 we have

$$P_{12a}(t_2|t_1) = \int_{t_1}^{t_2} \lambda_{03a}(s)(1+\theta[\Lambda_{03a}(s,t_2)+\Lambda_{03b}(s,t_2)])^{-1/\theta-1}\,\mathrm{d}s,$$

and

$$P_{12b}(t_2|t_1) = \int_{t_1}^{t_2} \lambda_{03b}(s)(1+\theta[\Lambda_{03a}(s,t_2)+\Lambda_{03b}(s,t_2)])^{-1/\theta-1}\,\mathrm{d}s.$$

Chapter 4

Simulation study

A simulation study with the illness-death model with shared frailty has been carried out and will be presented in this chapter. By simulating data from the model and then estimating the parameters, one can test the method. This will ensure that the method is correct and give an indication of how good the estimates are when the model is applied to real data sets.

The first section contains the algorithm used to simulate the data. The parameters used to generate data are estimated using maximum likelihood estimation and the quality of the estimates are evaluated. This is presented in the second section. In the third section we present the simulation algorithm and some simulation results in the expanded illness-death model with shared frailty.

4.1 Simulation algorithm

A procedure for simulating data from a semi-competing risks situation as the one presented in Figure 2.1 has been carried out. In this model the transition rates are on the parametric form as presented in Section 3.1.1, which means we have simulated data for both the power law model $\lambda_{0i}(t) = \alpha_i \beta_i t^{\beta_i - 1}$, and the log-linear law $\lambda_{0i}(t) = \exp(a_i + b_i t)$. The algorithm for simulating data from the illness-death model with shared frailty is presented in Algorithm 1. The algorithm starts with generating the frailty parameter γ. Next, on line 3 the first event time is simulated. This is done by drawing from the probability of not having any transitions before time t which is $\exp(-\gamma[\Lambda_{01}(t) + \Lambda_{02}(t)])$ from equation (3.11). This is done by equating the expression for the probability to a number uniformly distributed between 0 and 1, and solving for t. Then, on line 4, the probability that the transition at time t is to state 1 is calculated from

$$p = P(\text{Go to state 1}| \text{ Transition from state 0 in } [t, t + \Delta t)) =$$

$$\frac{P(\text{Go to state 1 in } [t, t + \Delta t))}{P(\text{Transition from state 0 in } [t, t + \Delta t))} = \frac{\lambda_{01}(t)}{\lambda_{01}(t) + \lambda_{02}(t)}.$$

If there is a transition to state 1, the second event time t_2 is simulated on line 9. This is calculated from $\exp(-\gamma\Lambda_{03}(t_1, t_2))$ given in (3.12), by equating it to a number uniformly distributed between 0 and 1, and solving for t_2. Lastly, an independent censoring time is simulated. The censoring distribution has been chosen to be a mixture distribution as in Xu et al. (2010). This distribution has equal weights on a uniform distribution between v and w, and a point mass at w. The values of v and w can be chosen to get the preferred amount of censoring.

The algorithm has been implemented as the functions `SimData.power()` and `SimData.loglinear()` which can be found in Appendix C.1.1 and C.1.2, respectively. These functions call on other functions which are given below in the same appendix.

4.2 Simulation results

Semi-competing risks data is generated using Algorithm 1 and maximum likelihood estimation is used to estimate the parameters. This is done using the built-in function `optim()` in R on the log likelihood functions for the models which are presented in Appendix C.2. The likelihood function in the general and restricted model are presented in equation (3.9) and (3.10) respectively. Further, the censoring mechanism has $v = 5$ and $w = 10$.

In this section the results from estimating the model parameters from simulated data sets are presented. The standard deviations are found from taking the square root of the inverse of the Hessian matrix of the log likelihood functions. The confidence intervals are estimated using the fact that maximum likelihood estimators are asymptotically normal (Casella and Berger, 2002, p. 472).

Algorithm 1 Simulate data from the illness-death model with shared frailty.

```
for each subject do
    $\gamma \sim Gamma(\frac{1}{\theta}, \theta)$
    $t_1 \sim$ first event time
    $p =$ probability of going to state 1
    $u \sim uniform(0, 1)$
    if $u \leq p$ then
        $Y_1 = t_1$
        $\delta_1 = 1$
        $t_2 \sim$ second event time
        $Y_2 = t_2$
    else
        $Y_1 = Y_2 = t_1$
        $\delta_1 = 0$
    end if
    $C \sim$ censoring distribution
    if $C > Y_2$ then
        $\delta_2 = 1$
    else if $C \geq Y_1$ and $C \leq Y_2$ then
        $Y_2 = C$
        $\delta_2 = 0$
    else
        $Y_1 = Y_2 = C$
        $\delta_1 = \delta_2 = 0$
    end if
end for
```

4.2.1 The general model with power law

We will first simulate data and fit them to the general model with power law for the conditional transition rates. The log likelihood function for this model is given in C.2.1.

Constant transition rates

We first simulate a data set where all parameter values are 1. This means that the three conditional transition rates are all constant and equal. The data set consists of 1000 observations and there is about 10% censoring. The initial values in `optim()` are set to be 1 for all the parameters. The model parameters and the results of estimating all seven parameters are presented in Table 4.1.

The maximum likelihood estimates from this model are close to the true parameter values, and all 95% confidence intervals cover the true values. The standard deviations of the $\hat{\alpha}$'s and $\hat{\theta}$ are higher than for the $\hat{\beta}$'s which indicates that these parameters are harder to estimate.

Table 4.1: Maximum likelihood estimates of the parameters in the general model with constant conditional transition rates described in Section 4.2.1. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	1	0.9778	0.1009	0.7799	1.1756
α_2	1	1.0243	0.1071	0.8143	1.2343
α_3	1	0.9244	0.0851	0.7576	1.0912
β_1	1	1.0085	0.0499	0.9108	1.1063
β_2	1	1.0751	0.0521	0.9729	1.1773
β_3	1	0.9833	0.0708	0.8445	1.1219
θ	1	1.0629	0.1208	0.8262	1.2996

Time-varying transition rates

In this model the parameters have been chosen such that the conditional transition rates from state 0 to state 1 and from state 0 to state 2 decrease with time, while the conditional transition rate from state 1 to state 2 increases with time. The simulated data set consists of 1000 observations and there is about 16% censoring. Again, the initial values in `optim()` are set to be 1 for all the parameters. The model parameters and the results of estimating all seven parameters are presented in Table 4.2.

We find that all the parameters are estimated quite correctly despite the initial values were not set to be the correct parameter values. We note that the standard deviations are of the same magnitude as in Table 4.1 and that all confidence intervals cover the true parameter value.

Table 4.2: Maximum likelihood estimates of parameters in the general model with time-varying conditional transition rates described in Section 4.2.1. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	1	1.1203	0.1143	0.8962	1.3444
α_2	1	0.9956	0.1055	0.7888	1.2023
α_3	1	1.0767	0.0775	0.9248	1.2287
β_1	0.5	0.5244	0.0262	0.4729	0.5758
β_2	0.5	0.5446	0.0280	0.4896	0.5995
β_3	2	1.9531	0.0856	1.7854	2.1209
θ	1	0.9713	0.1210	0.7341	1.2085

4.2.2 The restricted model with power law

In this section we will simulate data and fit them to the restricted model, where $\lambda_{02} = \lambda_{03}$ is assumed. This means that $\alpha_2 = \alpha_3$ and $\beta_2 = \beta_3$ such that only five parameters will be estimated. The log likelihood function for the model that is maximized is given in C.2.2.

Constant transition rates

For comparison with the general model we first use the same data set as in Section 4.2.1 where the parameters are all set to be 1, which means that the conditional transition rates are all constant and equal. The initial values in `optim()` are set to be 1 for all the parameters. The model parameters and the results of estimating all five parameters are presented in Table 4.3.

Again, all the estimates are good and the confidence intervals cover the true parameter values. When comparing with the results in the general model in Table 4.1 the parameters estimates in the restricted model are not outstandingly better. For some parameters the estimates in the general model are closer to the true parameter and for some parameters the estimates in the restricted model are closer to the true parameter. However, the standard deviations in the restricted model are lower than in the general model. This is as expected since there are fewer parameters to estimate in the restricted model.

Table 4.3: Maximum likelihood estimates of the parameters in the restricted model with constant conditional transition rates described in Section 4.2.2. The standard deviation (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	1	0.8973	0.0714	0.7573	1.0373
α_2	1	0.9182	0.0585	0.8035	1.0329
β_1	1	0.9659	0.0396	0.8884	1.0435
β_2	1	0.9996	0.0337	0.9334	1.0657
θ	1	0.9593	0.0881	0.7867	1.1319

Constant and time-varying transition rates

In the last model with power law the conditional transition rates to state 2 have been set constant, and the conditional transition rate from state 0 to state 1 is decreasing with time. In the simulated data set there are 1000 observations and about 20% censoring. The initial values in `optim()` are set to be 1 for all the parameters. The model parameters and the results of estimating all five parameters are presented in Table 4.4. As seen earlier, the difference between the true parameter values and the estimates are relatively small, and the standard deviations are good, especially considering the initial values were not the same as the true values.

Table 4.4: Maximum likelihood estimates of the parameters in the restricted model with constant and time-varying conditional transition rates described in Section 4.2.2. The standard deviation (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	2	2.0477	0.1946	1.6660	2.4291
α_2	1	1.0158	0.0780	0.8628	1.1686
β_1	0.8	0.8073	0.0294	0.7497	0.8649
β_2	1	1.0067	0.0334	0.9412	1.0723
θ	2	1.9319	0.1335	1.6703	2.1935

4.2.3 The general model with log-linear law

The log likelihood function for the general model with log-linear law is given in C.2.3. Using the log-linear law for the conditional transition rates we have simulated 1000 observations, where the conditional transition rates are equal for all transitions. The censoring is about 35% which is high. Here, all the initial values in `optim()` have been set to be the true parameter value. The model parameters and the results of estimating all seven parameters are presented in Table 4.5.

We find that all the parameters are estimated quite correctly except for $\hat{b}_1$ and $\hat{b}_2$. They have been estimated to be negative, while they are actually positive. This would mean that the estimated conditional transition rate would be decreasing with time, rather than increase. But considering the high proportion of censoring, the estimates and the confidence intervals are good.

Table 4.5: Maximum likelihood estimates of the parameters in the general model with conditional transition rates following the log-linear law. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
a_1	-1	-1.1041	0.0959	-1.2921	-0.9161
a_2	-1	-1.1478	0.0971	-1.3383	-0.9574
a_3	-1	-1.1808	0.1284	-1.4326	-0.9291
b_1	0.01	-0.0413	0.0485	-0.1365	0.0538
b_2	0.01	-0.0766	0.0509	-0.1764	0.0232
b_3	0.01	0.0043	0.0357	-0.0657	0.0744
θ	2	1.7196	0.2740	1.1823	2.2568

4.2.4 The restricted model with log-linear law

For the same data set as in Section 4.2.3 above we fit the restricted model with log-linear law for the transition rates. The log likelihood function for this model is given in C.2.4. Again, the true values are used as initial values. The results are presented in Table 4.6. We see an improvement in the estimates, especially for $\hat{\theta}$, and a small decrease in standard deviation compared to the general model in Table 4.5. This is as expected since there are fewer parameters to estimate in this model.

Table 4.6: Maximum likelihood estimates of the parameters in the restricted model with conditional transition rates following the log-linear law. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
a_1	-1	-1.0594	0.0903	-1.2366	-0.8823
a_2	-1	-1.1398	0.0806	-1.2979	-0.9818
b_1	0.01	0.0098	0.0347	-0.0582	0.0778
b_2	0.01	0.0006	0.0248	-0.0482	0.0493
θ	2	2.0253	0.1861	1.6607	2.3900

4.3 Simulating from the expanded model

In this section we present the simulation algorithm for semi-competing risks data in the expanded illness-death model with shared frailty. We also use simulated data to test the log likelihood function for the expanded illness-death model which is given in Appendix C.2.6.

4.3.1 Simulation algorithm in the expanded model

A procedure for simulating data from the expanded illness-death model as the one presented in Figure 3.1 has been carried out. The simulation algorithm is presented in Algorithm 2. We have only simulated data using the power law model, since the log-linear law is not used in the expanded model later. As in Algorithm 1, we start by generating the frailty parameter γ. On line 3 the first event time is simulated. This is done by drawing from the probability of not having any transitions before time t which is $\exp(-\gamma[\Lambda_{01}(t) + \Lambda_{02a}(t) + \Lambda_{02b}(t)])$ in the expanded model. The expression is set equal to a uniformly distributed number between 0 and 1, and solved for t. The probability of the first event being event 1 given a transition at time t is computed on line 4. This is calculated from

$$p_1 = \frac{\lambda_{01}(t)}{\lambda_{01}(t) + \lambda_{02a}(t) + \lambda_{02b}(t)}.$$

If there is a transition to state 1 at t, the second event time t_2 is simulated on line 9. This is calculated from $\exp(-\gamma[\Lambda_{03a}(t, t_2) + (\Lambda_{03b}(t, t_2))])$, by equating it to a uniformly distributed number and solving for t_2. Whether the second transition goes to state $2a$ or $2b$ is computed on lines 11-17 with p_3 being the probability of going to state $2a$ from state 1 at time t_2

$$p_3 = \frac{\lambda_{03a}(t_2)}{\lambda_{03a}(t_2) + \lambda_{03b}(t_2)}.$$

If the first transition was not to state 1, the probability of going to state $2a$ from state 0 at time t is computed. This is done on lines 21-26, where

$$p_2 = \frac{\lambda_{02a}(t)}{\lambda_{02a}(t) + \lambda_{02b}(t)},$$

is the probability of going to state $2a$ at time t when we know that state 1 is not possible for this transition.

Lastly, an independent censoring time is simulated. The censoring distribution has been chosen to be the same as the one described in Section 4.1, a mixture distribution with equal weights on a uniform distribution between v and w, and a point mass at w. The function for simulating the semi-competing risks data using this algorithm has been implemented in R as `SimData.expanded()`. This function is presented in Appendix C.1.3.

4.3.2 Simulation results in the expanded model

Semi-competing risks data from the expanded illness-death model with shared frailty is generated using Algorithm 2 and maximum likelihood estimation is used to estimate the parameters. Again, this is done using the built-in function `optim()` in R, now on the log of the likelihood function from equation (3.20). The log likelihood function has been implemented in R and the function is given in Appendix C.2.6. The standard deviations and condfidence intervals are computed using the Hessian matrix as in Section 4.2

We begin by simulating a data set of 1000 observations where all the parameters are set to have the value 1. The censoring mechanism is still $v = 5$ and $w = 10$ which gives about 6% censoring. The initial values in `optim()` are set to be 1 for all the parameters. The estimated parameters are given in Table 4.7, and we see that the estimates are close to the true values. The standard deviations are also good.

Table 4.7: Maximum likelihood estimates of the parameters in the expanded model with constant conditional transition rates. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	1	0.9771	0.1210	0.7400	1.2143
α_{2a}	1	1.0074	0.1248	0.7627	1.2520
α_{2b}	1	1.0456	0.1262	0.7981	1.2931
α_{3a}	1	0.9709	0.1144	0.7467	1.1951
α_{3b}	1	0.9513	0.1114	0.7329	1.1697
β_1	1	0.9817	0.0529	0.8780	1.0854
β_{2a}	1	1.0083	0.0536	0.9033	1.1133
β_{2b}	1	0.9540	0.0507	0.8547	1.0534
β_{3a}	1	1.0240	0.0939	0.8400	1.2081
β_{3b}	1	0.9675	0.0934	0.7845	1.1506
θ	1	1.0098	0.1168	0.7808	1.2388

Next, we simulate a data set of 1000 observations where $\lambda_{01}(t) = 0.2t$, $\lambda_{02a}(t) = \lambda_{02b}(t) = t$, $\lambda_{03a}(t) = \lambda_{03b}(t) = 0.5$, and $\theta = 2$. With the same values in the censoring mechanism, $v = 5$ and $w = 10$, the data set has about 10% censoring, and the initial values are set to be the true values. The maximum likelihood estimates of the parameters are given in Table 4.8. These estimates are not as good as the ones in Table 4.7, probably due to that there is a little more censoring and that there is more variation in this data set compared to the first data set that was simulated. However, all the confidence intervals cover the true value.

4.4 Summary of the simulation study

From the simulation study we find that maximum likelihood estimation works well to estimate the parameters in the illness-death model, both for the general and restricted models, and the expanded model. The confidence intervals of the estimated parameters all cover the true value. As we have fewer parameters in the model, the estimates are closer to the true values and the standard deviations also become lower, which is as expected.

When we in the next section use maximum likelihood estimation with `optim()` in R to find the model parameters, we can trust that the results are good and that the estimation procedure works correctly. This of course assumes that the data sets are large enough and without a high proportion of censoring.

Table 4.8: Maximum likelihood estimates of the parameters in the expanded model with constant and non-constant conditional transition rates. The standard deviations (SD) and bounds for the 95% confidence intervals are included.

Par.	True value	Est.	SD	Lower bound	Upper bound
α_1	0.1	0.1082	0.0178	0.0734	0.1431
α_{2a}	0.5	0.5087	0.0591	0.3929	0.6245
α_{2b}	0.5	0.4437	0.0532	0.3394	0.5479
α_{3a}	0.5	0.3569	0.1431	0.0764	0.6374
α_{3b}	0.5	0.5280	0.1510	0.2319	0.8242
β_1	2	1.9785	0.1513	1.6819	2.2751
β_{2a}	2	1.9694	0.1168	1.7405	2.1983
β_{2b}	2	2.0817	0.1217	1.8432	2.3202
β_{3a}	1	1.1753	0.2544	0.6766	1.6739
β_{3b}	1	1.0914	0.1936	0.7119	1.4708
θ	2	1.8643	0.1978	1.4765	2.2520

Algorithm 2 Simulate data from the expanded illness-death model with shared frailty.

```
for each subject do
    $\gamma \sim Gamma(\frac{1}{\theta}, \theta)$
    $t \sim$ first event time
    $p_1 =$ probability of going to state 1
    $u \sim uniform(0,1)$
    if $u \leq p_1$ then
        $Y_1 = t$
        $\delta_1 = 1$
        $t_2 \sim$ second event time
        $Y_2 = t_2$
        $p_3 =$ probability of going to state $2a$ from state 1
        $u \sim uniform(0,1)$
        if $u \leq p_3$ then
            $\delta_a = 1$, $\delta_b = 0$
        else
            $\delta_a = 0$, $\delta_b = 1$
        end if
    else
        $Y_1 = Y_2 = t$
        $\delta_1 = 0$
        $p_2 =$ probability of going to state $2a$ from state 0 given that can not go to state 1
        if $u \leq p_2$ then
            $\delta_a = 1$, $\delta_b = 0$
        else
            $\delta_a = 0$, $\delta_b = 1$
        end if
    end if
    $C \sim$ censoring distribution
    if $C > Y_2$ then
        $\delta_2 = 1$
    else if $C \geq Y_1$ and $C \leq Y_2$ then
        $Y_2 = C$
        $\delta_2 = 0$
    else
        $Y_1 = Y_2 = C$
        $\delta_1 = \delta_2 = 0$
    end if
end for
```

Chapter 5

Data analysis

In this chapter we will apply the illness-death model to real data sets consisting of semi-competing risks data. We will fit the data to the likelihood functions of the parametric models presented in Chapter 3 and tested in Chapter 4, to study the model and compare the results with other research. The first data set contains data on 137 bone marrow transplant patients and can be found in Klein and Moeschberger (1997). This data set will be modelled using the general and restricted model with both power law and log-linear law, and also some sub-models. The second data set contains data from 1313 randomly chosen patients from a cohort study from 2008 called SIR 3 (Spread of nosocomial Infections and Resistant pathogens) and is aimed at analysing the effect of hospital-acquired infections on the length of intensive care unit stay (Wolkewitz et al., 2008). This data set will be modelled using the general model with power law and the expanded model with power law. We also include covariates, and for the expanded model we will test some sub-models.

To maximise the log likelihood functions we use the function `optim()` in R, as done in the simulation study. In addition to the log likelihood functions used in the simulation study we also use two log likelihood functions for models with covariates. These are given in Appendix C.2 with the other log likelihood functions. In some cases in the data analysis we test sub-models of the main models. The log likelihood functions for these models are not presented, but have been implemented in the same way as the other log likelihood functions with minor changes.

The standard deviations and confidence intervals for the parameter estimates are mostly computed using the Hessian. However, in Section 5.2.3 and Section 5.2.4 the Hessian matrices of the log likelihood functions are not obtainable. In these sections we have therefore used non-parametric bootstrapping to compute the standard deviations (Givens and Hoeting, 2012, p. 288-289) and the confidence intervals are computed using the percentile method. For more about this see (Givens and Hoeting, 2012, p. 292-294). In all the other sections, the standard deviations and confidence intervals have been computed using the Hessian matrix.

5.1 Bone marrow transplant data

The data set for this section contains 137 observations of patients after a bone marrow transplantation as treatment for acute leukaemia. It is included in the R-library `KMsurv` as `bmt`, and in the R-library `SemiCompRisks` as `BMT`. The terminal event in this data set is death and the non-terminal event is relapse of cancer. The failure times are measured in days from the transplantation and for each of the four possible cases described in Table 2.1 it has been observed

- 40 of case 1, death following a relapse
- 40 of case 2, only death
- 3 of case 3, censoring following a relapse
- 54 of case 4, censoring before any event

Recall that the conditional transition rate from state 0 to state 1, which is relapse in this case, is modelled by λ_{01}, the conditional transition rate from state 1 to state 2, which is death here, is λ_{03}, and the conditional transition rate from state 0 to state 2 is λ_{02}, where all are multiplied with γ. This data set is small compared to the ones used in the simulation study. Furthermore the proportion of censored observations is larger than the data sets used to test the method. Because of this we cannot expect the estimates to be as precise as in the simulation study.

Before we start with the data analysis we shortly present a study by Fine et al. (2001) in which they have applied a copula model on the bone marrow transplant data set. They have modelled the dependency between the two failure times by assuming that the joint distribution of T_1 and T_2 follows a gamma frailty copula with observations only available in the upper wedge. Using their method they have made two estimates of the frailty variance. The frailty variance in their model, θ_F, can be related to the frailty variance in the shared-frailty illness-death model, θ_I by the relation $\theta_I = \theta_F - 1$. Their estimates of the frailty variance are $\hat{\theta}_{u,F} = 8.79(2.15)$ and $\hat{\theta}_{w,F} = 8.61(2.15)$, where $\hat{\theta}_{u,F}$ is an unweighted estimate and

$\hat{\theta}_{w,F}$ is a weighted estimate. The standard deviations are given in the parentheses. These estimates correspond to frailty variances of 7.79 and 7.61 in our model. Furthermore, they have estimated the marginal survival function for time to relapse, $\hat{S}_1(t_1)$, the function presented in (3.14), which will be compared to our results.

In the following subsections we fit the bone marrow transplant data to the illness-death model with shared frailty using maximum likelihood estimation. We will test models with both the power law and the log-linear law, and have both general and restricted models. We will also test whether some transition rates can be modelled as constant to get a simpler model with fewer parameters. Finally, we sum up with a comparison of the model fits, compare with the results of Fine et al. (2001), and make a conclusion about which model gives the best fit.

5.1.1 The power law

The general model

In this section we assume the conditional transition rates can be modelled using the power law from equation (3.7), and maximise the likelihood function in the general model from (3.9). That means maximising the log likelihood function in C.2.1. The obtained parameter estimates are presented in Table 5.1, and the standard deviation and the 95% confidence intervals are included in the table.

The parameter values for the $\hat{\alpha}_i$s become small compared to the ones used in the simulation study. This is because in the bone marrow transplant data set the time until failure is much longer than the ones we simulated. We get negative lower bounds for the $\hat{\alpha}_i$s. This is unrealistic since the transition rates cannot be negative, so the lower bounds are set to be zero. We also note that the standard deviation for $\hat{\theta}$ is large.

The estimated marginal transition rates for the transitions out of state 0 are presented in Figure 5.1a. We note that the rate to death is higher than the rate to relapse, but that the rates themselves are low. The estimated marginal survival function for time to relapse, $\hat{S}_1(t_1)$, is given in Figure 5.2a. Here, the 95% confidence interval is included, along with 50 bootstrapped survival functions found by re-sampling from the data set and estimating the parameters. The confidence intervals are estimated by linearising the logarithm of the survival function and computing the variance. This is described in Appendix B.3 and the equation for the limits are given in equation (B.11). As expected, some of the survival functions found by bootstrapping are outside the 95% confidence interval.

Table 5.1: Maximum likelihood estimates of parameters in the general illness-death model with the power law for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
α_1	0.000200	0.000239	0	0.000668
α_2	0.000420	0.000466	0	0.001333
α_3	0.000517	0.001101	0	0.002675
β_1	1.426809	0.251797	0.933287	1.920331
β_2	1.287615	0.235588	0.825863	1.749368
β_3	1.488679	0.364924	0.773428	2.203932
θ	3.984286	1.118510	1.792006	6.1765666

The restricted model

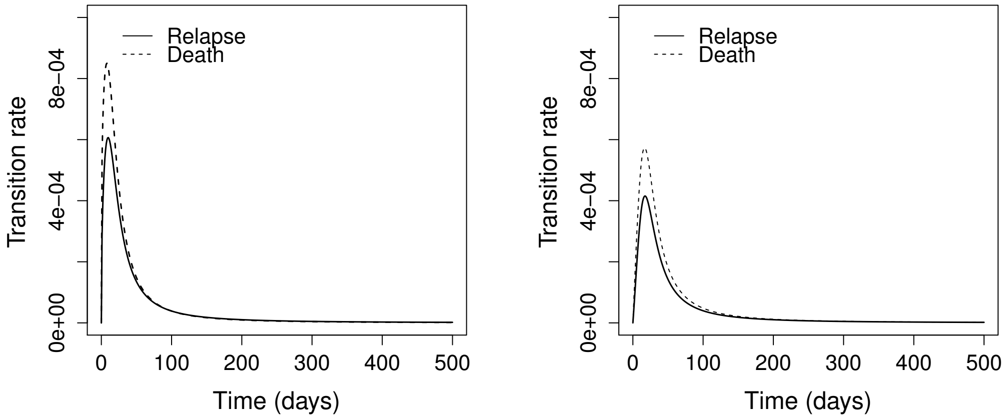
Now, we estimate the parameters in the restricted model by maximising the log likelihood function in C.2.2, which corresponds to maximising the restricted likelihood function given in equation (3.10). The obtained parameter estimates together with the standard deviations and the 95% confidence bounds are presented in Table 5.2. Again, the lower limits for the $\hat{\alpha}_i$s are negative so they are set to be zero. Furthermore we note that the $\hat{\alpha}_i$s in this model are ten times smaller than the ones in Table 5.1. This is however compensated by much higher $\hat{\beta}_i$s and a higher $\hat{\theta}$ in the restricted model.

The transition rates out of state 0 in the restricted model are presented in Figure 5.1b. As in the general model, the rate to death is higher than the rate to relapse. Compared to the general model, both rates are lower in the restricted model. The estimated survival function for time to relapse, $\hat{S}_1(t_1)$, in the restricted model is given in Figure 5.2b. Here, the 95% confidence interval is included, along with 50 bootstrapped survival functions. Again as expected, some of the survival functions found by bootstrapping are outside the 95% confidence interval. We see that the restricted and the general model give similar estimates for the marginal survival from relapse.

The two model fits can be compared using the maximum log likelihood values in each model as a measure. The maximum log likelihood values in the general model and restricted model with a power law are -942.61 and -950.24, respectively. We perform a likelihood ratio test with the null hypothesis being that the restricted model is the true model and the alternative hypothesis being that the general

Table 5.2: Maximum likelihood estimates of parameters in the restricted illness-death model with the power law for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
α_1	0.000021	0.000025	0	0.000070
α_2	0.000033	0.000036	0	0.000104
β_1	1.987193	0.252473	1.492346	2.482040
β_2	1.934969	0.230151	1.483872	2.386066
θ	6.515850	1.233012	4.099146	8.932554

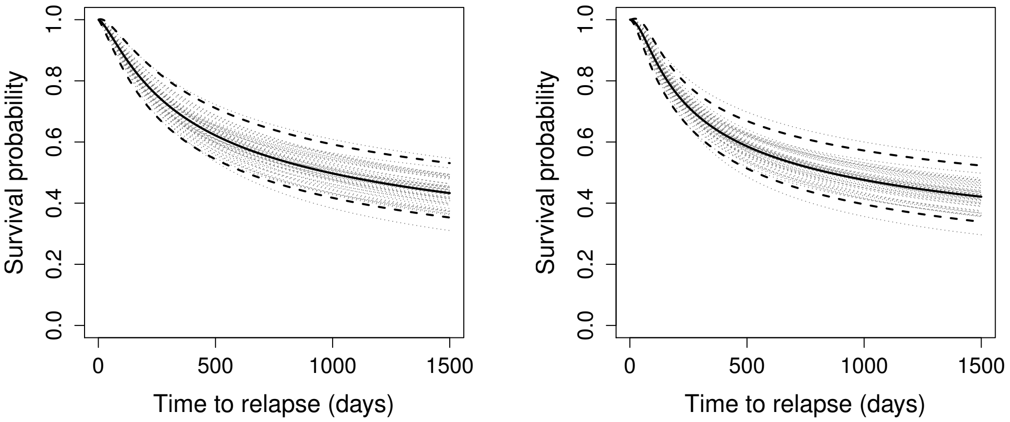


(a) The general model.

(b) The restricted model.

Figure 5.1: Estimated marginal transition rates for the transition to relapse (full line) and the transition to death (dashed line) not following relapse in the general and restricted model using the power law.

model is the true model. The likelihood ratio test and the formula of the test statistic are presented in Appendix A.2. The test statistic is 15.26, chi-squared distributed with 2 degrees of freedom. This means that we can reject the null hypothesis with a 5% significance level.



(a) The general model.

(b) The restricted model.

Figure 5.2: Estimated survival functions for time to relapse, $\hat{S}_1(t_1)$, for the bone marrow transplant data set using the power law. The solid lines are the estimated probabilities, the dotted lines are estimates obtained from non-parametric bootstrap samples and the dashed lines are the 95% confidence intervals.

Lastly, we compute the estimated survival function for time to death given that a patient has had a relapse at a specific time, $\hat{S}_{12}(t_2|t_1)$ from (3.15), in the general model. A plot of this survival function is presented in Figure 5.3 for different relapse times. From this figure it seems that the time of relapse has little impact on the chances of survival. It seems that with this model the chances of survival are a little lower for patients that experience a relapse after a long time compared to the ones who relapse early after transplantation, since the curves become steeper as the relapse time increases. However, the curves converge to a common probability such that the time of relapse only affects the survival probability the first year or so after relapse. We have not computed the estimated survival function for time to death given relapse in the restricted model since we believe the general model is a better fit due to the result of the likelihood ratio test.

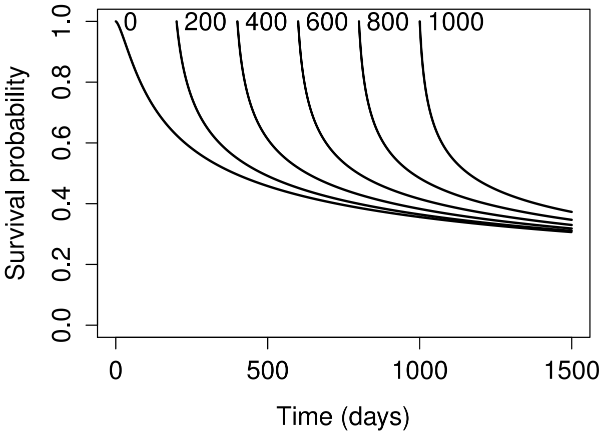


Figure 5.3: Estimated marginal survival functions for time to death given that a patient has had a relapse, $\hat{S}_{12}(t_2|t_1)$, for the power law. Each curve in this figure presents the probability for survival given that there was a relapse at the time given on top of the curve.

5.1.2 The log-linear law

The general model

In this section we assume that the conditional transition rates can be modelled using the log-linear law from (3.8). First, the likelihood function for the general model is maximised, which is done by maximising the log likelihood function in C.2.3. The obtained parameter estimates along with the standard deviations and the bounds for the 95% confidence intervals are presented in Table 5.3. We note that the estimated frailty variance $\hat{\theta}$ is much lower than in the other models considered so far.

The marginal transition rates out of state 0 in the general model with log-linear law are presented in Figure 5.4a. As opposed to the power law, the rate to relapse is now a little higher than the rate to death, and both rates are higher here than with the power law. The estimated marginal survival function for time to relapse, $\hat{S}_1(t_1)$, is presented in Figure 5.5a. Here, the 95% confidence interval is included, along with 50 bootstrap survival functions. The confidence interval is estimated in the same way as for the power law. This is described in Appendix B.3 and the equation for the limits is given in equation (B.12).

Table 5.3: Maximum likelihood estimates of parameters in the general illness-death model with the log-linear law for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
a_1	-6.541741	0.223584	-6.979967	-6.103515
a_2	-6.814524	0.236918	-7.278884	-6.350163
a_3	-4.410459	0.342400	-5.081563	-3.739355
b_1	-0.002298	0.000686	-0.003643	-0.000953
b_2	-0.001593	0.000578	-0.002728	-0.000459
b_3	-0.001189	0.000824	-0.002806	0.000427
θ	0.555935	0.455088	0	1.447908

The restricted model

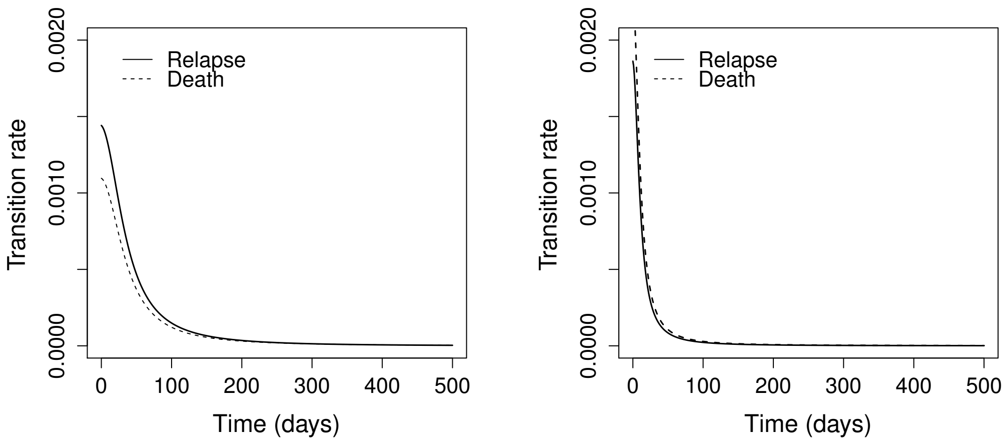
We also estimate the parameters in the restricted model with log-linear law by maximising the restricted likelihood function. The log likelihood function for this model is given in C.2.4, and maximising this results in the parameter estimates presented in Table 5.4, together with the standard deviations and the 95% confidence bounds.

Table 5.4: Maximum likelihood estimates of parameters in the restricted illness-death model with the log-linear law for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
α_1	-6.285297	0.271566	-6.817567	-5.753027
α_2	-6.074593	0.244646	-6.554099	-5.595087
β_1	0.000749	0.000836	-0.000888	0.002387
β_2	0.001139	0.000697	-0.000227	0.002506
θ	4.061232	1.021664	2.058769	6.063694

The transition rates in the restricted model with log-linear law are presented in Figure 5.4b. The two transition rates are almost equal in this model. The estimated marginal survival function for time to relapse, $\hat{S}_1(t_1)$, is given in Figure 5.5b. The 95% confidence interval is included, along with 50 bootstrapped survival functions. The curves are steeper than the ones for the general model in

Figure 5.5a.



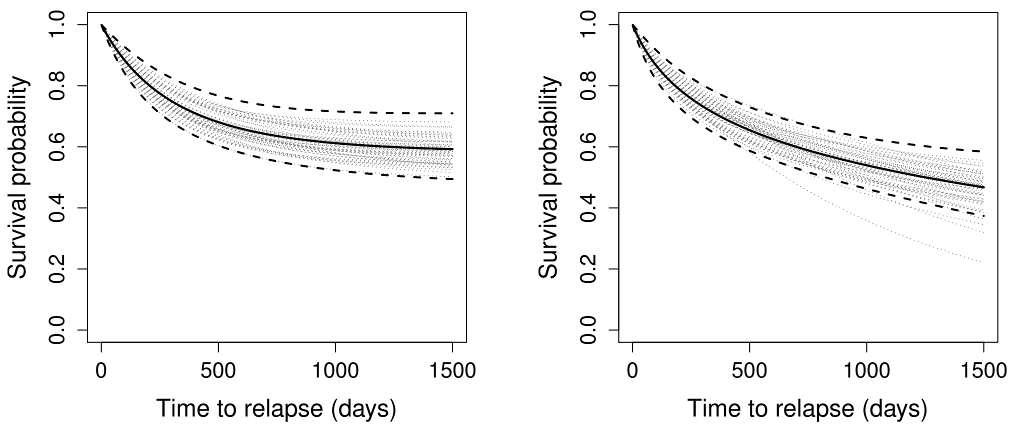
(a) The general model.

(b) The restricted model.

Figure 5.4: Estimated marginal transition rates for the transition to relapse (full line) and the transition to death (dashed line) not following relapse in the general and restricted model using the log-linear law.

The general and the restricted model with log-linear law can be compared using a likelihood ratio test, which is described in Appendix A.2. The maximum log likelihood values in the general model and restricted model with log-linear law are -939.66 and -964.80, respectively. A likelihood ratio test with the null hypothesis being that the restricted model is the true model and the alternative hypothesis being that the general model is the true model, gives a test statistic with value 50.28 being chi-squared distributed with 2 degrees of freedom. The null hypothesis is therefore rejected with a significance level of 5%, which means that the general model gives a statistically significant improvement on the maximum likelihood value compared to the restricted model.

Finally, we present the estimated survival function for time to death given that a patient has had a relapse, $\hat{S}_{12}(t_2|t_1)$. A plot of this in given in Figure 5.6. In this model, the curves have about the same shape for all relapse times which means that a patient with a late relapse has about the same survival chances as a patient with an early relapse. Compared to the curves from the power model in Figure 5.3 the estimated survival probability in the log-linear is much lower. Again, we have not computed the estimated survival function for time to death given relapse in the restricted model since we believe the general model is a much better fit.



(a) The general model.

(b) The restricted model.

Figure 5.5: Estimated survival functions for time to relapse, $\hat{S}_1(t_1)$, for the bone marrow transplant data set using the log-linear law. The solid lines are the estimated probabilities, the dotted lines are estimates obtained from 50 non-parametric bootstrap samples and the dashed lines are the 95% confidence intervals.

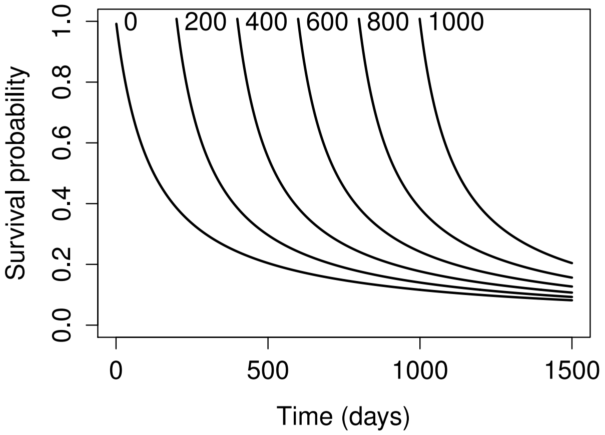


Figure 5.6: Estimated marginal survival functions for time to death given that a patient has had a relapse, $\hat{S}_{12}(t_2|t_1)$, for the log-linear law. Each curve in this figure presents the probability for survival given that there was a relapse at the time given on top of the curve.

5.1.3 Sub-models with constant conditional transition rates

Before we compare the model fits we want to investigate the possibilities of having some or all conditional transition rates constant. This corresponds to having $b_i = 0$ in the log-linear law or $\beta_i = 1$ in the power law. This would give constant conditional transition rates $\lambda_{0i}(t) = \exp(a_i)$, and cumulative conditional transition rates $\Lambda_{0i}(t) = t\exp(a_i), i = 1, 2, 3$, when we base on using the log-linear law. This assumption is reasonable since the estimated b_is in the log-linear law and the estimated β_is in the power law have confidence intervals indicating possibilities of constant conditional transition rates.

A log likelihood function with $b_i = 0, i = 1, 2, 3$ in the log-linear law in both the general and restricted model has been implemented in R. The result from maximising the log likelihood function in the general model with constant conditional transition rates for the bone marrow transplant data is presented in Table 5.5. The parameter estimates are close to the ones for the general log-linear model with $b_i \neq 0$ given in Table 5.3, except for $\hat{\theta}$ which has increased. The maximum log likelihood in this model is -944.48.

Table 5.5: Maximum likelihood estimates of parameters in the general illness-death model with the log-linear law where all $b_i = 0$ for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
a_1	-6.555849	0.240859	-7.027934	-6.083764
a_2	-6.626798	0.244386	-7.105795	-6.147801
a_3	-4.724115	0.260679	-5.235046	-4.213184
θ	2.457795	0.467264	1.541958	3.373631

We perform a likelihood ratio test with the null hypothesis being that all $b_i = 0$ and the alternative hypothesis being all $b_i \neq 0$. The log likelihood in the alternative hypothesis was -939.66, which is the general log-linear law model, and -944.48 in the null hypothesis. This gives a test statistic of 9.64 which is chi-squared distributed with 3 degrees of freedom, and leads to a rejection of the null hypothesis with 5% significance level.

The same is done for the restricted log-linear law with constant conditional transition rates, and the result from maximising the log likelihood is presented in Table 5.6. The parameter estimates are almost the same as in the restricted log-linear model with all $b_i \neq 0$ given in Table 5.4, except for $\hat{\theta}$ which has decreased. The maximum log likelihood in this model is -966.58.

Table 5.6: Maximum likelihood estimates of parameters in the restricted illness-death model with the log-linear law where all $b_i = 0$ for the bone marrow transplant data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
a_1	-6.363778	0.238929	-6.832079	-5.895477
a_2	-5.994624	0.212429	-6.410985	-5.578263
θ	2.960120	0.494791	1.990330	3.929911

Again, we perform a likelihood ratio test with the null hypothesis being that all $b_i = 0$ and the alternative hypothesis being that all $b_i \neq 0$. Since the restricted power law performed better than the restricted log-linear law, we use the restricted power law as alternative hypothesis. The maximum log likelihood value in the alternative hypothesis was -950.24, and -966.58 in the null hypothesis. This gives a test statistic of 32.68 which is chi-squared distributed with 2 degrees of freedom. We reject the null hypothesis with 5% significance level, which means that the improvement in the likelihood is statistically significant when going from a restricted model with all $b_i = 0$ to a restricted power law model with all $\beta_i \neq 1$.

Lastly, we test two general models where $\lambda_{03} = const$ and the other conditional transition rates follow the power law and the log-linear law. We test the general model because up until this point all general models have performed statistically significantly better than the restricted models. Furthermore, in both Table 5.1 and 5.3 the estimates have indicated that λ_{03} is not statistically significantly different from being constant.

In Table 5.7 the parameter estimates for a general power law model with $\beta_3 = 1$ are presented. The maximum log likelihood for this model is -943.55. We note that the maximum log likelihood for the general power law model from Section 5.1.1 was -942.61, a model with one parameter more than the one tested here. Again, we perform a likelihood ratio test. The null hypothesis is the general power law model with $\beta_3 = 1$ and the alternative hypothesis is the general power law model. This gives a test statistic of 1.88 chi-squared distributed with 1 degree of freedom. With 5% significance level the null hypothesis cannot be rejected, and one can conclude that the improvement in the maximum log likelihood value in the general model from Section 5.1.1 is not statistically significant compared to the general power law model with $\beta_3 = 1$.

Table 5.7: Maximum likelihood estimates of parameters in the general illness-death model with the power law for the bone marrow transplant data set with assumption $\beta_3 = 1$. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Estimate	SD	Lower bound	Upper bound
α_1	0.000419	0.000414	0	0.001231
α_2	0.000817	0.000747	0	0.002281
α_3	0.009503	0.002703	0.004204	0.014803
β_1	1.244138	0.191385	0.869023	1.619254
β_2	1.119950	0.177662	0.771730	1.468169
θ	3.122930	0.793938	1.566810	4.679050

In Table 5.8, the maximum log likelihood estimates of the parameters in the general log-linear law model with assumption $b_3 = 0$ are presented. The maximum log likelihood value in this model is -940.63. We note that this is close to the maximum log likelihood value in the general log-linear law model in Section 5.1.2, which was -939.66. A likelihood ratio test then has a test statistic of 1.94 chi-squared distributed with one degree of freedom. We therefore do not reject the null hypothesis with 5% level of significance. The conclusion from this is that going from the general log-linear law model with $b_3 = 0$ to the general model in Section 5.1.2 does not give a statistically significant improvement in the maximum log likelihood value.

Table 5.8: Maximum likelihood estimates of parameters in the general illness-death model with the log-linear law for the bone marrow transplant data set with assumption $b_3 = 0$. The standard deviations and limits for the 95% confidence interval are included.

Parameter	Estimate	SD	Lower bound	Upper bound
a_1	-6.481748	0.000414	-6.482560	-6.480937
a_2	-6.736707	0.000747	-6.738171	-6.735243
a_3	-4.798302	0.002703	-4.803601	-4.793002
b_1	-0.001864	0.191385	-0.376980	0.373251
b_2	-0.001224	0.177662	-0.349443	0.346995
θ	1.056057	0.793938	0	2.612176

5.1.4 Conclusion and comparing results

Eight variations of the illness-death model with shared frailty have now been fitted to the bone marrow transplant data set. In this subsection we summarise the results and findings from the model fitting. We start by giving an overview of the maximum log likelihood values for the models in Table 5.9.

As we have already seen, all the general models perform significantly better than the corresponding restricted models. Furthermore, from the table we can also note that a general model with all constant conditional transition rates gives a better fit than a restricted model with non-constant conditional transition rates, both with the power law and the log-linear law. This indicates that a relapse affects survival probability, and that the conditional transition rate to death should be different after a relapse. Moreover, by looking at the results in Table 5.1, Table 5.3, Table 5.5, Table 5.7 and Table 5.8 the conditional transition rate λ_{03} is higher than λ_{02} in all three models, which indicates that a relapse increases the probability of death. This result agrees with conclusions from other literature, where it has been found that the prognosis of relapsed disease is poor (Nguyen et al., 2008; Oriol et al., 2010).

In Figure 5.7, the estimated marginal survival probability from relapse, $\hat{S}_1(t_1)$, in the different models are displayed together. We see that the three power law models (blue, green, cyan) give similar results. The models with constant conditional transition rates also give similar curves (black, brown). The magenta and red curve are general log-linear law models where the magenta curve is from the model with λ_{03} being constant. The orange curve is the restricted log-linear law model. The magenta and red differ from the other curves in that the probability

Table 5.9: Maximum log likelihood values for the models in Section 5.1.

Model	Max log L
General power law	-942.61
Restricted power law	-950.24
General log-linear law	-939.66
Restricted log-linear law	-964.80
General constant rates	-944.48
Restricted constant rates	-966.68
General power law, $\lambda_{03} = const$	-943.53
General log-linear law, $\lambda_{03} = const$	-940.63

of not having a relapse in these models is higher.

There could be many reasons for why the models give different results when it comes to the estimated survival function for time to relapse as we see in Figure 5.7. First of all, the tables with parameter estimates show that there is uncertainty in all the parameters, and in some cases the 95% confidence intervals capture values which will have different impact on the functions we want to estimate. For example, in Table 5.1, the confidence intervals for all the $\hat{\beta}$s in the general power law model contain values both lower and higher than 1. Furthermore, most of the event times are between 0 and 500 days, which causes the estimates to be better at these time points and less good elsewhere. The data set also has many censored observations. In fact as much as about 40 % of the observations belong to case 4, where the observations are censored before any events occur.

In Section 5.1.3 we concluded that the general models with $\lambda_{03} = const$ gave model fits which were not statistically significantly improved on with the same models having λ_{03} non-constant. Because of this we would suggest that a general model with λ_{03} being constant gives the best fit for the bone marrow transplant data set. Since these models have fewer parameters we can expect that the parameter estimates are more precise, even though they are less flexible. In other words, we prefer models that sharply make correct estimates over models that accommodate a wide range of possible results.

We now compare the two best model fits from our data analysis to the result from Fine et al. (2001). This is done in Figure 5.8, and we see that our two estimated survival functions flatten out more than the curve estimated by Fine et al. (2001). Their non-parametric estimate is much steeper in the beginning

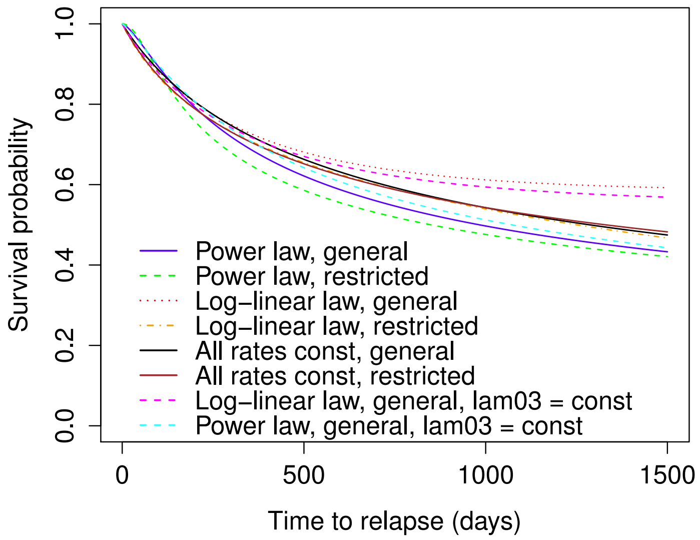


Figure 5.7: Estimated marginal survival function for time to relapse, $\hat{S}_1(t_1)$, for all eight models.

and seems to stabilise at a probability of about 0.4. Of the two models we have included in the figure, the general power law model with constant λ_{03} is closest to the non-parametric estimate. A cause for the difference can be that in our models we have used parametric models for the conditional transition rates which could be forcing the curves to be higher than the non-parametric estimate. Furthermore, it is not only the time to relapse we have modelled. We have also modelled the transitions to death from both the initial state and the relapsed state. This may cause a trade-off between getting the estimates as precise as possible in all three transition rates.

If we had included the restricted power law model in Figure 5.8, we would have seen that this resembles the curve estimated by Fine et al. (2001) most of all our models. In fact, according to Xu et al. (2010) the restricted illness-death model is essentially equal to the copula model used in Fine et al. (2001). However, in this section we have found that a general model is a better fit for the bone marrow transplant data than a restricted model. This strengthens our results in Figure 5.8 compared to the curve estimated by Fine et al. (2001), and is also another explanation for why the parametric marginal survival curves are not more similar to the result from Fine et al. (2001).

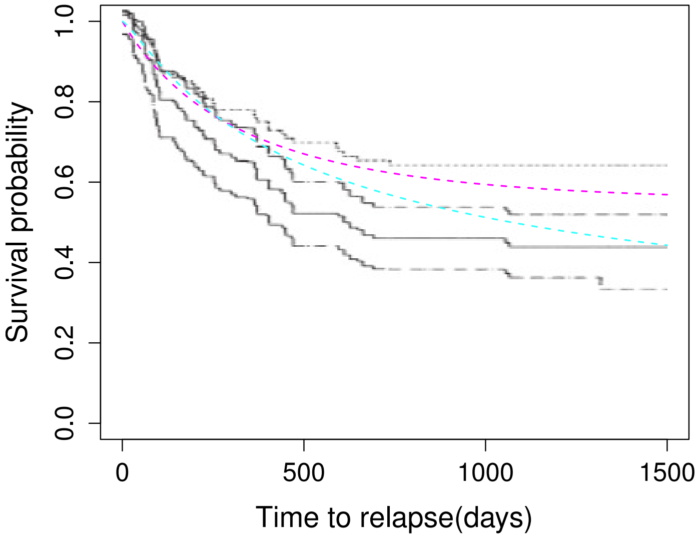


Figure 5.8: Estimated marginal survival function for time to relapse, $\hat{S}_1(t_1)$, in the general model with the assumption $\lambda_{03} = const$, for both the power law (cyan) and the log-linear law (magenta). Included is also the estimate from Fine et al. (2001) in solid lines (black), the dashed lines (black) are the 95% confidence limits, and the dotted, upper line (black) is the Kaplan-Meier estimate.

In all the models, the estimated frailty variance has been different and it has not been as high as in the model of Fine et al. (2001). Overall, $\hat{\theta}$ has been lower in the general models than the restricted models. It seems that the frailty therefore varies more when there are fewer parameters in the model, as if the frailty has more variation to describe when two transition rates are assumed to be equal. This may explain the $\hat{\theta}$ in the general log-linear law model in Table 5.3 which was only ~ 0.56. Since the model was a good fit, there was not much more variation in need of modelling and therefore the frailty variance was low.

From the results in this section it is clear that the interpretation of the conditional and marginal transition rates is not the same. In the power law model the conditional transition rates all increase with time, whereas the marginal transition rates given in Figure 5.1 have a peak at about 10-20 days and then decrease with time. For the log-linear law, some of the conditional transition rates decrease with time and some conditional transition rates increase with time, while the marginal transition rates in Figure 5.4 decrease with time. The reason for this difference between the marginal and the conditional transition rates is that the conditional transition rates are person-specific and the marginal transition rates are population-specific. A person-specific curve will usually increase with time because sooner or later a person must have a transition. The population-specific rates decrease with time because the most frail subjects have a transition first and the more robust are left behind and have a long time before having a failure. This will cause the the marginal transition rates to decrease even though the conditional transition rates are increasing.

Usually, the frailty of a specific person is not accessible, which means that we cannot say anything about the conditional transition rate of a specific patient. We can only use the marginal transition rates which tells us something about the whole population with the frailty averaged out. Therefore as long as we do not know the frailty of a person, the marginal rates may be useful for drawing conclusions.

In summary, we have found that relapse increases the probability of death due to the general models giving the best model fits, and the estimated conditional transition rates to death after relapse are higher than without relapse. This result corresponds with other studies which have found that prognosis of relapsed disease is poor (Nguyen et al., 2008; Oriol et al., 2010). The general model with λ_{03} constant is the best model fit for the bone marrow transplant data. The maximum log likelihood values for the models with power law and log-linear law are -943.53 and -940.63 respectively. Of these two models, the power law model resembles the result found by Fine et al. (2001) in survival from relapse more than the log-linear law.

5.2 Hospital-acquired pneumonia

In this section we will analyse the effect of hospital-acquired pneumonia on the length of intensive care unit stay and hospital mortality using the data set `icu.pneu` from the cohort study SIR 3. The data set can be found in the `kmi` library in R as `icu.pneu`. Hospital-acquired pneumonia is the most commonly reported infection in intensive care units (Wolkewitz et al., 2008), and it is of great interest to the health care to understand how infections are associated with increased length of hospital stay and the impact on the morbidity and mortality (Safdar et al., 2005). For more about the cohort study the reader is referred to Wolkewitz et al. (2008).

The non-terminal event in this data set is hospital-acquired pneumonia, while there are two terminal events, discharge from the hospital and death on the intensive care unit. We will consider both a model where these two are merged to one terminal event, and the expanded model where the two terminal events are considered separately, as presented in Section 3.5. We will refer to the illness-death model with combined terminal events as the ordinary model, and the illness-death model with two competing terminal events as the expanded model.

The failure times are measured in days from admittance on intensive care unit, and for each of the four possible cases for observations in competing risks described in Table 2.1 it has been observed

- 103 of case 1, out of these 82 were discharged alive and 21 died
- 1189 of case 2, out of these 1063 were discharged alive and 126 died
- 5 of case 3
- 16 of case 4

In this data set most of the failure times are between 1 and 20 days, and we therefore expect that the parametric models will give best results in this interval.

The data set has been analysed in several chapters of a book by Beyersmann et al. (2011). They have studied the impact of hospital-acquired pneumonia on length of intensive care unit stay and the impact on intensive care unit mortality, using a Markovian model. The main result from their study is that hospital-acquired pneumonia increases the hospital mortality via prolonged stay (Beyersmann et al., 2011, p. 182-192, p. 202-206, p. 216-217).

In Section 5.2.1 the data will be modelled using the ordinary model and in Section 5.2.3 the data will be modelled using the expanded model. The data set contains information about age and sex of the patients. This information will be included in the models as covariates in the ordinary and expanded model presented in Section 5.2.2 and Section 5.2.4 respectively. For the data analysis in this section we will only use the general model and one parametric model for the conditional transition rates since we are mostly interested in comparing the ordinary and expanded model, and this will reduce the number of analyses. The parametric model we will use is the power law.

5.2.1 The ordinary model

We first fit the ordinary illness-death model with power law to the pneumonia data by maximising the log likelihood function given in C.2.1. The parameter estimates are given in Table 5.10, and the standard deviations and limits for the 95% confidence intervals are included. From the parameter estimates it would seem that most patients have a transition to end of stay rather than pneumonia. Furthermore, we note that the conditional transition rate to end of stay with prior pneumonia infection will be small in the beginning, but grow much faster than the other rates due to the big $\hat{\beta}_3$.

Table 5.10: Maximum likelihood estimates of parameters in the ordinary illness-death model with the power law for the pneumonia data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Est.	SD	Lower bound	Upper bound
α_1	5.42957e-04	1.97523e-04	1.55813e-04	9.30102e-04
α_2	3.48583e-03	7.71232e-04	1.97422e-03	4.997445e-03
α_3	4.40665e-05	3.86585e-05	0	1.19837e-04
β_1	2.54557	0.17339	2.20573	2.88542
β_2	2.76053	0.13678	2.49244	3.02863
β_3	3.13487	0.26087	2.62356	3.64617
θ	1.74072	0.16168	1.42383	2.05762

To study the results visually, the marginal transition rates are computed. In Figure 5.9 the transition rates for pneumonia infection and end of stay without prior pneumonia infection are presented. The transition rate to hospital-acquired pneumonia is small, and since the transition rate for end of stay is much higher than the infection rate, most patients have an end of stay without pneumonia infection. The estimated transition rate for end of stay after pneumonia infection is given in Figure 5.10. This transition rate depends on the time of pneumonia infection t_1 in addition to the time since admission, and the curves are drawn for different infection times s. Included are also the estimated transition rates for end of stay without pneumonia infection for comparison. For all infection times s, an infection of pneumonia decreases the transition rate to end of stay compared to not being infected with pneumonia. Furthermore, as the time of pneumonia infection increases, the transition rate to end of stay with pneumonia infection becomes more similar to the transition rate to end of stay without pneumonia infection. This means that the effect of pneumonia is less prominent for patients being infected late in the admission. Since we in this model do not distinguish between the two terminal events, it is not possible to draw conclusions regarding the two terminal events. We will study the probabilities for the terminal events more closely in the Section 5.2.3 and Section 5.2.4 with the expanded model.

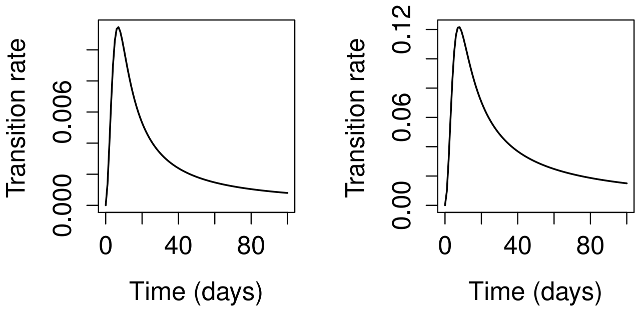


Figure 5.9: Estimated transition rate for pneumonia infection (left) and for end of stay prior to pneumonia infection (right) in the ordinary model.

The estimated probability of being in the infected state at time t, $\hat{P}_{01}(t_1)$ from equation (3.16), is presented in Figure 5.11, along with 50 bootstrap estimates. This probability has also been estimated by using the Aalen-Johanson estimator (Beyersmann et al., 2011, p. 185-186), and the estimate of Beyersmann et al. (2011) is included in Figure 5.11. As already seen in Figure 5.9, the rate to pneumonia infection is small and so is the probability. We see that the probability is highest at 10 to 15 days after admission, which coincides well with the transition rate in the left plot in Figure 5.9, where the transition rate for pneumonia infection is at its highest after 10 days.

The parametric estimate of the probability follows the non-parametric estimate closely and is inside the confidence intervals of Beyersmann et al. (2011) at all times. The bootstrap curves are also inside the confidence intervals most of the time. This indicates a good model fit for the infection of pneumonia.

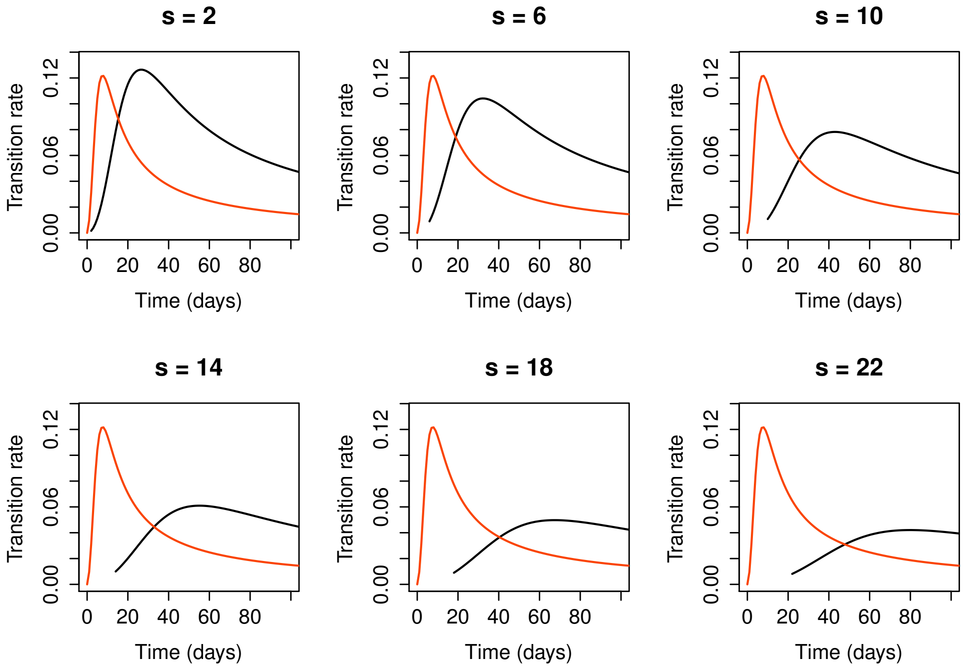


Figure 5.10: Estimated transition rates for end of stay with prior pneumonia infection at time s (black). Included is also the estimated transition rate for end of stay without prior pneumonia infection (red).

5.2.2 Including covariates in the ordinary model

In this section we include covariates in the ordinary model. As described in Section 3.4, to include covariates in the model, each conditional transition rate is multiplied with a transition-specific covariate term.

Since we have information about age and sex of the patients in the data set we include covariates for this. Each conditional transition rate then has two coefficients, $\varphi_{i,age}$ and $\varphi_{i,sexM}$, in addition to θ and the parameters in the power law. The likelihood function for this model is given in equation (3.19), and the log likelihood function for this model is presented in C.2.5. This function is maximised, and in Table 5.11 the parameter estimates for this model are presented along with the standard deviations and limits for the 95% confidence bounds.

The covariate coefficient for age, $\varphi_{1,age}$ indicates that increasing age increases the conditional transition rate for pneumonia infection. Age also slightly increases the conditional transition rate to end of stay without pneumonia, while it has a slight decreasing effect on the conditional transition rate to end of stay after

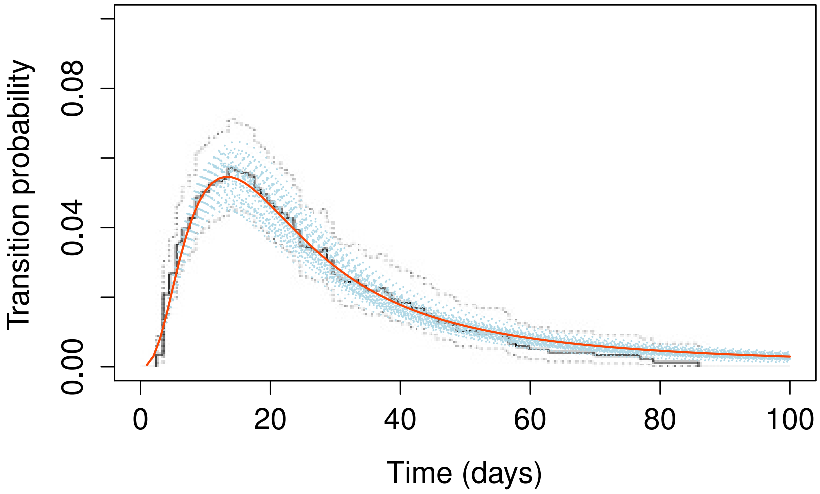


Figure 5.11: Estimated transition probability $\hat{P}_{01}(t_1)$ for pneumonia infection in the general model (red) and 50 bootstrap estimates (blue). Included is the Aalen-Johanson estimator (black) (Beyersmann et al., 2011, p. 186) with 95% confidence intervals (dotted) based on a complementary log-log transformation.

pneumonia infection.

The covariate coefficients for sex indicate that men have an increased conditional transition rate to pneumonia infection compared to women, while they have a decreased conditional transition rate to end of stay compared to women. For end of stay after pneumonia infection, men have an increased conditional transition rate. We will look closer into the effects of covariates on end of stay when using covariates in the expanded model in Section 5.2.4. The remaining parameter estimates in Table 5.11 and their standard deviations have not changed much from the model without covariates in Table 5.10.

We perform a hypothesis test on the effect of the covariates. The null hypothesis is that there is no effect of the covariates, $\varphi = 0$, and the alternative hypothesis is $\varphi \neq 0$. The test statistic, Z, for each covariate is the estimated covariate coefficient divided by the standard deviation of the estimate. The p-values are then found by computing the probability that a standard normal random variable is greater than Z or smaller than $-Z$. This is a two-sided test, and the Z-statistic and the p-values for the covariates are presented in Table 5.12. From the p-values in the

Table 5.11: Maximum likelihood estimates of parameters in the ordinary illness-death model with covariates for the pneumonia data set. The standard deviations (SD) and limits for the 95% confidence interval are included.

Par.	Est.	SD	Lower bound	Upper bound
α_1	2.32649e-04	1.52349e-04	0	5.31252e-04
α_2	3.47003e-03	1.23334e-03	1.05268e-03	5.88738e-04
α_3	5.89920e-05	6.48255e-05	0	1.86050e-04
β_1	2.54457	0.17414	2.20326	2.88588
β_2	2.76202	0.13718	2.49314	3.03089
β_3	3.17824	0.26624	2.65641	3.70008
θ	1.73758	0.16217	1.41973	2.05542
$\varphi_{1,age}$	1.01057e-02	6.21834e-03	-2.08226e-03	2.22936e-02
$\varphi_{2,age}$	4.10830e-03	3.27232e-03	-2.30544e-03	1.05220e-02
$\varphi_{3,age}$	-9.21382e-03	9.87593e-03	-2.85706e-02	1.01430e-02
$\varphi_{1,sexM}$	1.63175e-01	2.25482e-01	-2.78769e-01	6.05120e-01
$\varphi_{2,sexM}$	-1.49466e-01	1.20292e-01	-3.85238e-01	8.63063e-02
$\varphi_{3,sexM}$	6.79684e-02	3.27536e-01	-5.74001e-01	7.09938e-01

table we see that none of the covariates are statistically significant with 5% level of significance.

We use a likelihood ratio test to see if the model with covariates gives a statistically significant improvement on the maximum likelihood value compared to the ordinary model without covariates from Section 5.2.1. The maximum likelihood value for the model with covariates is -5189.87, which is the alternative hypothesis. The null hypothesis is the ordinary model with no covariates which has a maximum likelihood value of -5193.88. The theory for likelihood ratio tests and the test statistic are presented in Appendix A.2. The test statistic has a value of 8.02 and is chi-squared distributed with 6 degrees of freedom, since there are six more parameters in the model with covariates. This means that we do not reject the null hypothesis with 5% significance level. In other words adding covariates to the ordinary model does not improve the maximum likelihood value enough to be statistically significant.

Now, we take a closer look at the covariate for age for the conditional transition rate to pneumonia infection, $\varphi_{1,age}$. Several sources state that the elderly have a higher risk of being infected with pneumonia (Koivula et al., 1994; Paul et al.,

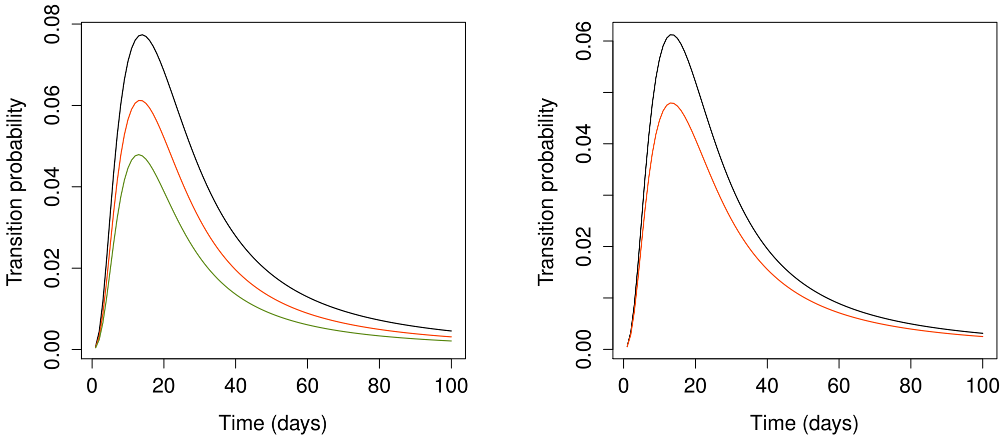
Table 5.12: The Z-statistics and p-values for the covariates in the ordinary illness-death model for the pneumonia data set.

Par.	Z	p-value
$\varphi_{1,age}$	1.6251	0.1041
$\varphi_{2,age}$	1.2555	0.2093
$\varphi_{3,age}$	-0.9329	0.3509
$\varphi_{1,sexM}$	0.7237	0.4692
$\varphi_{2,sexM}$	-1.2425	0.2141
$\varphi_{3,sexM}$	0.2075	0.8356

2015). With this information available, it will be natural to perform a one-sided hypothesis test with the null hypothesis being $\varphi_{1,age} = 0$, and the alternative hypothesis being $\varphi_{1,age} > 0$, which is that increasing age increases the conditional transition rate to pneumonia infection. The Z-statistic for $\varphi_{1,age}$ will still be the same as in Figure 5.12, but the p-value will be half of the p-value for $\varphi_{1,age}$ in Table 5.12. This is because the p-value is now found by computing the probability that a standard normal random variable is greater than Z. The p-value for $\varphi_{1,age}$ is then about 0.0521, which is not far from being statistically significant with a significance level of 5%.

Even though the covariates are not statistically significant with a significance level of 5%, it does not mean that there is no difference in the conditional transition rates for different ages or between male and female patients. It only means that we cannot with enough certainty reject the null hypothesis which says that there is no difference for different covariate values. There may still be effects of the covariates and it should be kept in mind that the high p-values do not imply that there are no differences between groups of patients (Rothman et al., 2008).

We have computed $\hat{P}_{01}(t_1)$ using the estimates in Table 5.11 for different values of the covariates even though they are not statistically significant. This is presented in Figure 5.12. In Figure 5.12a we have estimated the transition probability for pneumonia infection for three men aged 30, 60 and 90 years, and find that the probability of being infected with pneumonia is higher for older patients, which corresponds to the results in Table 5.11. Figure 5.12b compares the transition probability for a man and a woman both 60 years old, and it shows that men have a higher probability of being infected with pneumonia than women.



(a) Three men aged 90 (black), 60 (red) and 30 (green) years old.

(b) Man (black) and woman (red) both 60 years old.

Figure 5.12: Transition probabilities for being in the infected state at time t using the estimates in the ordinary model with covariates. Note that the covariate for sex is not statistically significant, while the covariate for age is close to statistically significant with a one-sided test.

5.2.3 The expanded model

In this section the pneumonia data set is modelled using the expanded model such that death and discharge can be studied separately. Recall the expanded model in Figure 3.1, and that in this model the conditional transition rate to pneumonia infection is $\gamma\lambda_{01}$. Moreover, $\gamma\lambda_{02a}$ and $\gamma\lambda_{02b}$ are the conditional transition rates to discharge and death respectively prior to pneumonia infection, and $\gamma\lambda_{03a}$ and $\gamma\lambda_{03b}$ are the conditional transition rates to discharge and death respectively after pneumonia infection.

We begin by fitting the expanded model to the data using the log likelihood function given in C.2.6. The parameter estimates for the model are presented in Table 5.13. The Hessian matrix in this model is non-invertible and therefore the standard deviations have been calculated using bootstrapping, and the confidence intervals have been estimated using the percentile method. These have been calculated using 200 bootstrap estimates, which is usually a sufficient amount (Efron and Tibshirani, 1994). The parameter estimates in the expanded model are similar to the estimates in the ordinary model in Table 5.10, in the sense that $\hat{\lambda}_{01}$, $\hat{\lambda}_{02}$, $\hat{\lambda}_{03}$

in the ordinary model are approximately equal to $\hat{\lambda}_{01}$, $\hat{\lambda}_{02a} + \hat{\lambda}_{02b}$, $\hat{\lambda}_{03a} + \hat{\lambda}_{03b}$ in the expanded model.

Table 5.13: Maximum likelihood estimates of parameters in the expanded illness-death model with power law for the pneumonia data set. The standard deviations from bootstrapping (SD_B) and limits for the 95% percentile confidence intervals are included.

Par.	Est.	SD_B	Lower bound$_B$	Upper bound$_B$
α_1	5.30027e-04	2.03684e-05	5.29565e-04	5.40682e-04
α_{2a}	3.38892e-03	1.89421e-05	3.38668e-03	3.42467e-03
α_{2b}	1.35486e-04	4.65512e-06	1.35300e-04	1.38724e-04
α_{3a}	4.40643e-05	5.29643e-05	1.21220e-05	1.79707e-04
α_{3b}	3.18628e-06	1.65178e-04	4.44203e-09	1.58845e-04
β_1	2.56097	0.05588	2.45723	2.67052
β_{2a}	2.73386	0.04468	2.64599	2.82567
β_{2b}	3.14435	0.05496	3.02925	3.24641
β_{3a}	3.07312	0.18628	2.67806	3.38627
β_{3b}	3.42693	0.70983	2.37545	5.14524
θ	1.75650	0.08734	1.57816	1.91565

Before we visualise the results, we test some sub-models and compare their maximum log likelihood values to the maximum log likelihood value in the expanded model which is -5641.35. The $\hat{\beta}_i$s in Table 5.13 all have confidence intervals that do not cover the value 1. Therefore, it is not likely that a model with constant conditional rates is suitable. The parameter estimates for λ_{01} and λ_{2b} are most similar with regards to $\hat{\alpha}_1$ and $\hat{\alpha}_{2b}$. We therefore fit a model with the restriction $\lambda_{01} = \lambda_{02b}$. The log likelihood function for this model is not given. We find a maximum log likelihood value of -5648.49. Using a likelihood ratio test with null hypothesis being the sub-model and the alternative hypothesis being the expanded model, the test statistic is 14.28 chi-squared distributed with 2 degrees of freedom. This leads to a rejection of the null hypothesis with 5% significance level.

The second sub-model we test has the restriction $\lambda_{02b} = \lambda_{03a}$, which is the null hypothesis. It gives a maximum log likelihood value -5655.31. A likelihood ratio test then gives a test statistic with value 27.92 which is chi-squared distributed with 2 degrees of freedom. The alternative hypothesis is still the expanded model. This test statistic is high so the null hypothesis is rejected with 5% level of significance.

Since we did not find any reasonable sub-models that were almost as good as the expanded model itself, we continue our analysis with the expanded model. In Figure 5.13 the estimated transition rates for pneumonia infection, and discharge and death prior to pneumonia infection are presented. From this, we find that most patients who has an end of stay without pneumonia are discharged. The transition rate to death without pneumonia is about the same as the transition rate to pneumonia infection, although a little higher.

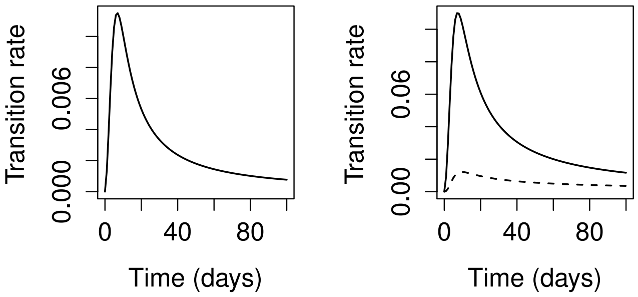


Figure 5.13: Estimated transition rates in the expanded model for pneumonia infection (left) and end of stay prior to pneumonia infection (right) caused by discharge (full line) and death (dashed line).

The transition rates to death and discharge after pneumonia infection at a specific time s are given in Figure 5.14. Included are also the transition rates for death and discharge without pneumonia infection. Recall that what is presented here are the instantaneous rates at which patients experience failure, given that they have survived up to time t, and possibly was infected at time s. This makes the curves a little hard to interpret and compare. However, we find that as s increases the transition rates to discharge and death after pneumonia infection decrease and become more similar to the rates to discharge and death without pneumonia. This means that the effect of pneumonia is less apparent for patients who are infected with pneumonia late in their hospital stay than patients being infected early, as we also saw in Section 5.2.1.

Furthermore, for all s the curves without infection are higher than the curves with pneumonia infection. This means that the patients who are infected with pneumonia have a lower transition rate to both discharge and death for the first 15 to 20 days after infection compared to patients who have not been infected.

After about 15 to 20 days after infection, the curves cross. However, as we will see in the next paragraph this does not mean that the probability of being discharged is bigger for patients who have experienced pneumonia infection.

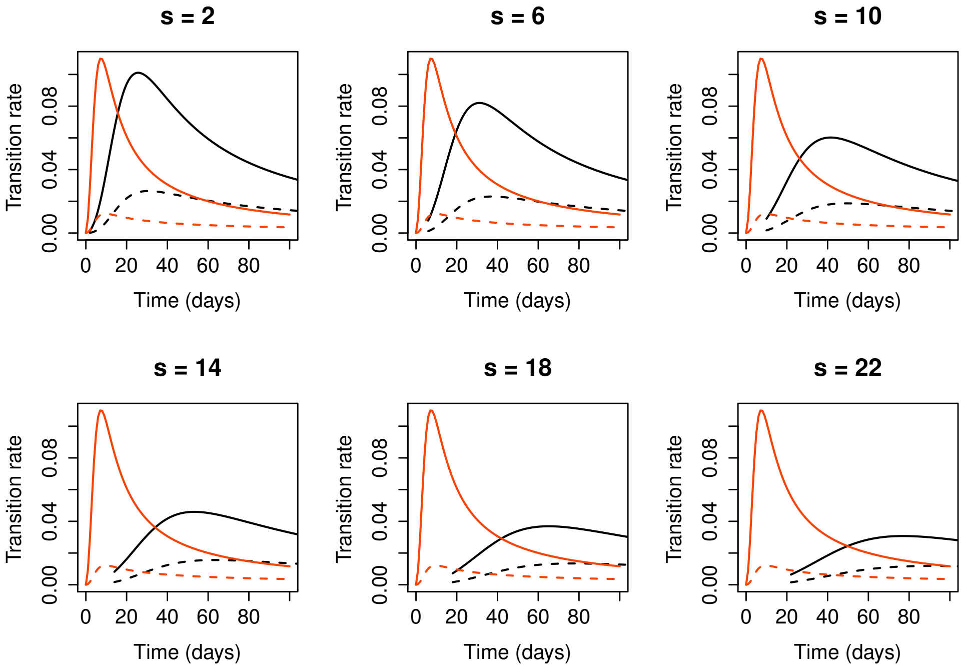


Figure 5.14: Estimated transition rates for discharge (full line) and death (dashed line) with prior pneumonia infection at time s in black. The red curves are the transition rates to discharge (full line) and death (dashed line) without prior pneumonia infection.

In Figure 5.15 we have plotted the estimated transition probabilities which were presented in Section 3.5. These are the probabilities of being discharged and dead at time t for patients who have been infected at time s, $\hat{P}_{12a}(t|s)$ and $\hat{P}_{12b}(t|s)$. In the same plots are also $\hat{P}_{02a}(t)$ and $\hat{P}_{02b}(t)$, the probabilities of being discharged and dead without pneumonia infection. From this figure we clearly see that pneumonia infection decreases the probability of being discharged, especially right after infection. Furthermore, pneumonia infection increases the probability of death for patients being hospitalised more than 20-30 days.

For death, the difference in the estimated transition probability between infected and non-infected patients is not as big as the difference between the discharge transition probabilities for infected and non-infected patients for the first 30 days. This indicates that the effect of pneumonia is bigger on discharge than

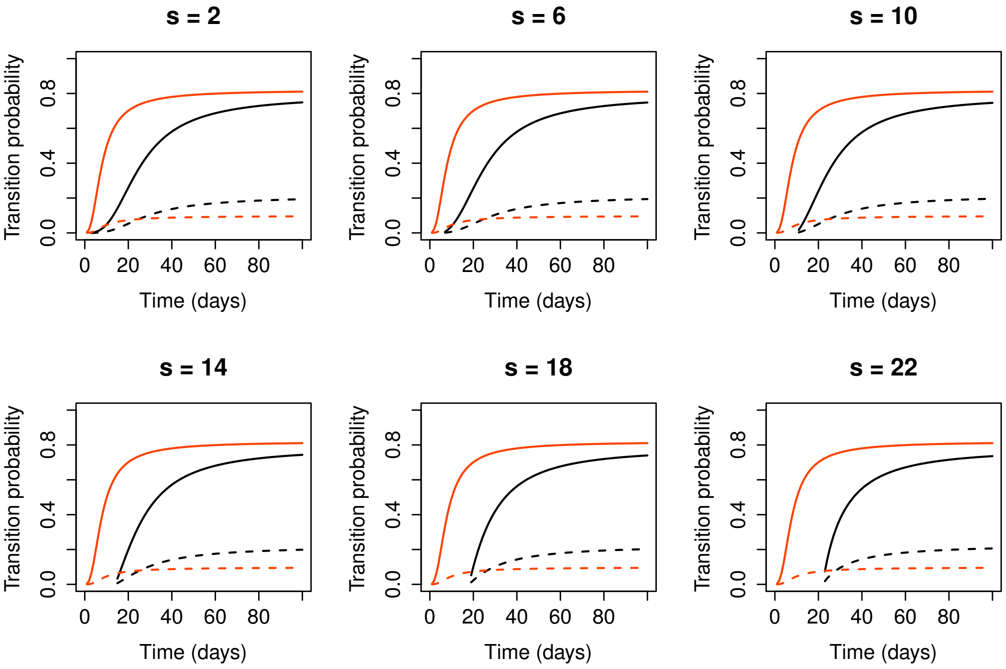


Figure 5.15: Estimated transition probabilities for death (black, dashed) and discharge (black, full line) with prior pneumonia infection at time s. Included are also the estimated transition probabilities for death (red, dashed) and discharge (red, full line) without prior pneumonia infection.

in death.

As s increases, the curves for infected patients become steeper. This means that the effect of pneumonia infection is smaller as the time of infection increases. This corresponds to what we saw from the transition rates in Figure 5.10 and Figure 5.14 where the transition rates with and without pneumonia become more similar as s increases.

The non-parametric transition probabilities have been estimated using the Aalen-Johansen estimators (Beyersmann et al., 2011, p. 187-188). These are the transition probabilities for end of stay in the model where the two terminal events are taken together, $P_{02}(t_2)$ and $P_{12}(t_2|t_1)$. Note that $P_{02}(t_2)$ in Beyersmann et al. (2011) is the transition probability of end of stay where end of stay after acquiring pneumonia is included, as opposed to our model where the pneumonia infected are not included in $P_{02}(t_2)$. However, since the probability of acquiring pneumonia is low, the difference will be small. When we add together the black, dashed curve with the black, full curve from Figure 5.15, we should get the same as the Aalen-Johansen estimators, and when we add the red, dashed curve with the red, full curve we should get almost the same results as the Aalen-Johansen estimators. We have not presented our curves together with the non-parametric estimates, but in fact the curves look similar. After about 30 days however, our parametric estimates underestimate the probabilities compared to the Aalen-Johansen estimators, which is not surprising since there is less data after 30 days than before 30 days.

5.2.4 Including covariates in the expanded model

In this section we include covariates in the expanded model by multiplying the conditional transition rates and the conditional cumulative transition rates in the likelihood function from (3.20) with transition-specific covariate terms as described in Section 3.4. Since there are five conditional transition rates, we get ten covariate coefficients in the model, two for each conditional transition rate describing the effect of age and sex.

The log likelihood function for this model, which is given in Appendix C.2.7, is maximised for the pneumonia data set, and the parameter estimates are presented in Table 5.14. Since the Hessian is non-invertible, the standard deviations and confidence intervals have been calculated using 200 bootstrap samples.

We start by looking at the coefficients for age. From the table, high age seems to have an increasing effect on the conditional transition rate to pneumonia infection. Moreover, the coefficient is close to the same coefficient in the ordinary model presented in Table 5.11, which is as expected. Furthermore, age also seems to increase the conditional transition rate to death without pneumonia infection and slightly increase the conditional transition rate to discharge without pneumonia. However, increasing age seems to have a decreasing effect on the discharge conditional transition rate and a slight decreasing effect on the conditional transition rate to death with pneumonia infection. The decreasing effect here is ten times bigger on the conditional transition rate to discharge than the conditional transition rate to death.

The table also shows the coefficients for sex. The conditional transition rate to pneumonia infection is higher for men than women. Men have a lower conditional transition rate to discharge than women, but a higher conditional transition rate to death. However, for the ones who are infected with pneumonia, the men have a higher conditional transition rate to discharge and a lower conditional transition

Table 5.14: Maximum likelihood estimates of parameters in the expanded illness-death model with covariates for the pneumonia data set. The bootstrap standard deviations (SD_B) and bounds for the 95% percentile confidence interval are included.

Par.	Est.	SD_B	Lower bound_B	Upper bound_B
α_1	2.16086e-04	1.55495e-04	5.87773e-05	6.26368e-04
α_{2a}	3.84399e-03	1.30667e-03	1.82886e-03	7.032831e-03
α_{2b}	2.30373e-05	2.03631e-05	4.69397e-06	7.58126e-05
α_{3a}	4.72314e-05	1.29020e-04	2.42308e-06	3.35065e-04
α_{3b}	7.84645e-06	5.76185e-04	1.98458e-08	9.35666e-04
β_1	2.56516	0.13816	2.33456	2.84209
β_{2a}	2.73627	0.13441	2.52208	3.04728
β_{2b}	3.17108	0.17253	2.87368	3.51829
β_{3a}	3.12801	0.30497	2.66782	3.79347
β_{3b}	3.46747	0.72497	2.21358	5.04740
θ	1.75684	0.14952	1.50786	2.07809
$\varphi_{1,age}$	1.06046e-02	6.37346e-03	-1.60974e-03	2.35827e-02
$\varphi_{2a,age}$	2.34243e-03	3.40434e-03	-4.59823e-03	9.06632e-03
$\varphi_{2b,age}$	2.62833e-02	6.79551e-03	1.33539e-02	3.98136e-02
$\varphi_{3a,age}$	-1.07193e-02	1.32614e-02	-3.58406e-02	1.47958e-02
$\varphi_{3b,age}$	-2.49579e-03	2.13658e-02	-4.50545e-02	4.33985e-02
$\varphi_{1,sexM}$	1.69696e-01	2.23558e-01	-2.88141e-01	5.74772e-01
$\varphi_{2a,sexM}$	-1.68873e-01	1.27475e-01	-4.05406e-01	7.68661e-02
$\varphi_{2b,sexM}$	8.60188e-02	2.36424e-01	-3.53856e-01	5.66228e-01
$\varphi_{3a,sexM}$	2.34488e-01	4.07662e-01	-5.09044e-01	9.92058e-01
$\varphi_{3b,sexM}$	-5.71287e-01	6.71115e-01	-2.02984	7.30264e-01

rate to death than women.

Of the remaining parameter estimates, $\hat{\alpha}_{2b}$ is the only estimate that has changed much from the estimates in the expanded model without covariates in Table 5.13. This is because the estimated covariate coefficient for age, $\hat{\varphi}_{2b,age}$, is quite large and the scaling of $\hat{\lambda}_{02b}$ must therefore be changed. The standard deviations in this model are bigger than in the expanded model without covariates in Table 5.13, because there are ten more parameters in this model.

We perform a hypothesis test where the null hypothesis is that there is no effect of the covariates, $\varphi = 0$, while the alternative hypothesis is $\varphi \neq 0$. The

Table 5.15: The Z-statistics and p-values for the covariates in the expanded illness-death model for the pneumonia data set.

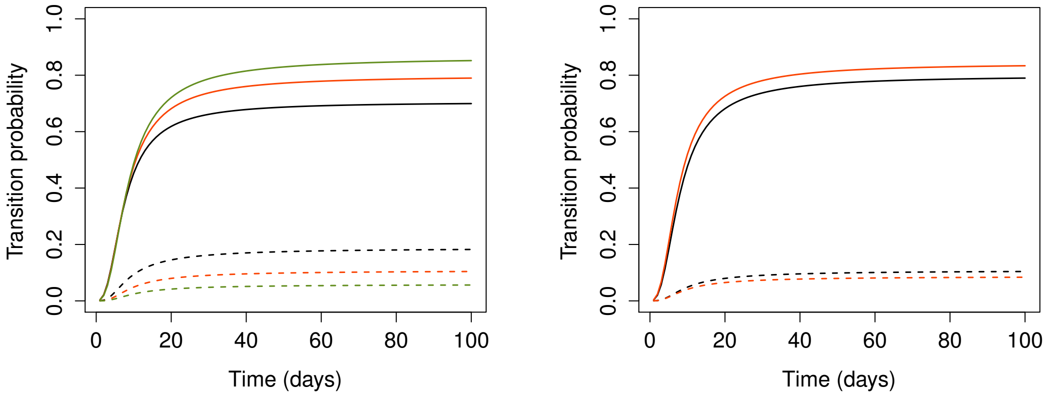
Par.	Z	p-value
$\varphi_{1,age}$	1.6639	0.0961
$\varphi_{2a,age}$	0.6881	0.4914
$\varphi_{2b,age}$	3.8677	0.0001
$\varphi_{3a,age}$	-0.8083	0.4189
$\varphi_{3b,age}$	-0.1168	0.9070
$\varphi_{1,sexM}$	0.7591	0.4478
$\varphi_{2a,sexM}$	-1.3248	0.1853
$\varphi_{2b,sexM}$	0.3638	0.7159
$\varphi_{3a,sexM}$	0.5752	0.5652
$\varphi_{3b,sexM}$	-0.8513	0.3946

Z-statistic and p-values for the covariates in the expanded model are computed as in the ordinary model, by dividing the estimated coefficients by their standard deviation. The Z-statistics and p-values are presented in Table 5.15. The p-values indicate that only the covariate for age on the conditional transition rate to death without pneumonia, $\hat{\varphi}_{2b,age}$, is statistically significant with 5% significance level.

If we use the same argumentation as in Section 5.2.2, which is that the elderly have an increased risk of pneumonia infection, a one-sided hypothesis test gives a p-value of about 0.0481 for $\hat{\varphi}_{1,age}$. This indicates statistical significance of the covariate for age for the conditional transition rate to death without pneumonia.

To compare the expanded model with covariates to the expanded model without covariates, we again use a likelihood ratio test where we let the former be the alternative hypothesis and the latter be the null hypothesis. The maximum likelihood value in the models are -5641.35 and -5626.71 in the null hypothesis and alternative hypothesis respectively. This gives a test statistic of 29.28 with 10 degrees of freedom, which suggest a rejection of the null hypothesis with a significance level of 5%. This means that the expanded model with covariates gives a significantly better fit than the expanded model without covariates.

In Figure 5.16 the effect of the covariates on $\hat{P}_{02a}(t)$ and $\hat{P}_{02b}(t)$ is presented. We start by exploring the statistically significant covariate $\varphi_{2b,age}$. Figure 5.16a shows $\hat{P}_{02b}(t)$ for three men aged 30, 60 and 90 years. We find that older patients have a higher probability of being dead than young patients, which corresponds to that age has an increasing effect on the conditional transition rate to death without pneumonia.



(a) Three men aged 90 (black), 60 (red) and 30 (green) years old.

(b) Man (black) and woman (red), both 60 years old.

Figure 5.16: Transition probability $\hat{P}_{02a}(t)$ (full line) and $\hat{P}_{02b}(t)$ (dashed line) using the estimates in the expanded model with covariates. Only the covariate for age on the transition rate to death is statistically significant here.

Furthermore, Figure 5.16a also contains $\hat{P}_{02a}(t)$ for male patients of different ages. We see that young male patients have a higher probability of being discharged than old patients after they have been admitted for more than about 10 days. This does not correspond with the result in Table 5.14, where we saw that higher age is slightly increasing the conditional transition rate to discharge. However, age increases the conditional transition rate to death more than it increases the conditional transition rate to discharge. This means that the conditional transition rate to death for young patients is much lower than for old patients, while the conditional transition rate to discharge for young patients is only slightly lower than for old patients. Since the conditional transition rate to discharge is higher than the conditional transition rate to death regardless of age, young patients will have a higher probability of discharge than old patients.

In the Figure 5.16b, $\hat{P}_{02a}(t)$ and $\hat{P}_{02b}(t)$ are presented for one man and one woman both 60 years old. The probability of being discharged is higher for women than men and the probability of being dead is higher for men than women, which corresponds to the covariate coefficients for the conditional transition rates. Nevertheless, the effect of sex is small, especially for death which is far from statistically significant, and is not showing before a few days after admission.

The effect of the covariates on $\hat{P}_{12a}(t|s)$ and $\hat{P}_{12b}(t|s)$ is also presented. In Figure 5.17, $\hat{P}_{12a}(t|s)$ and $\hat{P}_{12b}(t|s)$ are presented for different times s for men aged 30, 60 and 90 years old. For discharge we find the same as we did in Figure 5.16a, that older patients have a decreased probability of being discharged compared to young patients, although this effect is not statistically significant.

For death we also find the same as we did in 5.16a, which is that older patients have an increased probability of dying compared to young patients. However, the coefficients from Table 5.14 indicate that increasing age has a decreasing effect on the conditional transition rate to death. Here we get the same effect as we did for $\hat{P}_{02a}(t)$, which is that increasing age decreases the conditional transition rate to discharge more than it decreases the conditional transition rate to death. This means that the ratio between the conditional discharge and death rate for older patients is smaller than for young patients, and by this, older patients have a higher probability of dying after pneumonia than young patients.

In Figure 5.18, $\hat{P}_{12a}(t|s)$ and $\hat{P}_{12b}(t|s)$ are presented for different times s for a man and a woman both 60 years old. Here, the results are such that a man has a higher probability of being discharged after pneumonia infection, while a woman has a higher probability of dying after pneumonia infection. This corresponds to the fact that men has a lower conditional transition rate to death after pneumonia than women, and that men have a higher conditional transition rate to discharge than women. However, this effect is not statistically significant.

5.2.5 Conclusion and comparing results

In this section we compare the four model fits together and see if we can make a conclusion about the effect of hospital-acquired pneumonia on hospital mortality and morbidity.

With the expanded model, we find that hospital-acquired pneumonia reduces the transition rate to discharge and the probability of discharge, and also increases the probability of dying on the intensive care unit. From other literature (Beyersmann et al., 2011, p. 190) it has already been concluded that the increased probability of death after pneumonia infection is a cause of the prolonged stay in the hospital and that the death hazard remains more or less unchanged. We compute the conditional explanatory hazard ratio for the dependence of discharge and death on hospital-acquired pneumonia in Figure 5.19. The curves show how

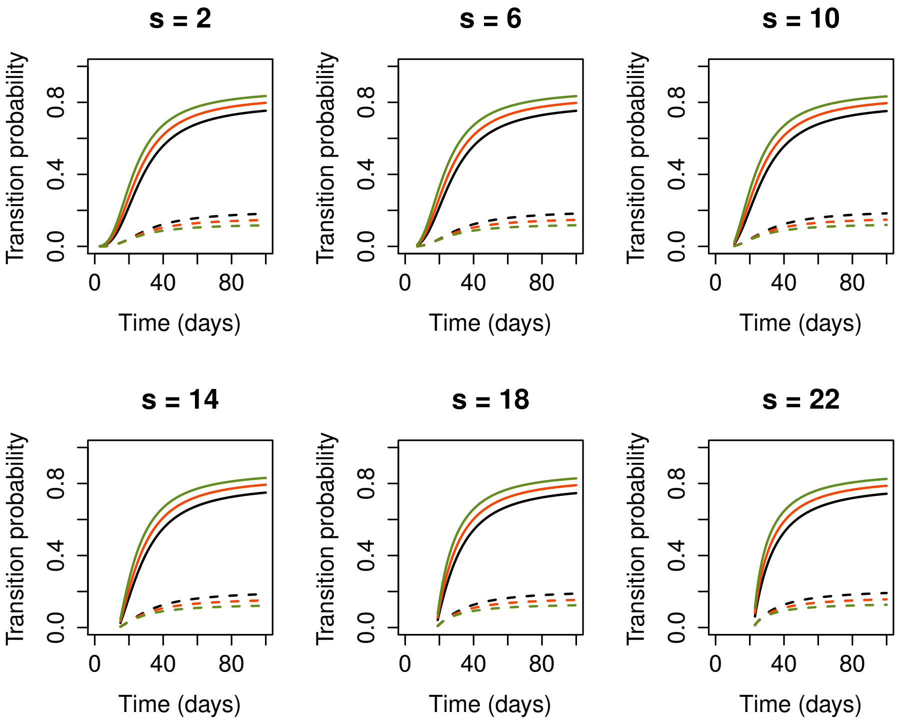


Figure 5.17: Estimated transition probabilities $\hat{P}_{12a}(t|s)$ (full line) and $\hat{P}_{12b}(t|s)$ (dashed line) using the estimates in the expanded model with covariates for three men aged 90 (black), 60 (red) and 30 (green) years old. None of the covariates here are statistically significant.

many times the risk of getting discharged or dying is increased over time by acquiring pneumonia. We find that the conditional risk of dying is actually decreased by acquiring pneumonia, however not as much as discharge. Since this measure is a conditional measure and also Markov, we can only interpret it on a person-specific level. For the population-specific interpretation we therefore also consider the marginal transition rates for discharge and death.

The marginal transition rates in Figure 5.14 show a similar result as the one found by Beyersmann et al. (2011). The difference between the two transition rates to death for patients with and without pneumonia infection are not equal, but they are similar when we compare with the difference between the transition rates to discharge with and without pneumonia. This supports the conclusion drawn by Beyersmann et al. (2011), that pneumonia does not have a direct effect on the hospital mortality, at least not a strong one. The increased probability of

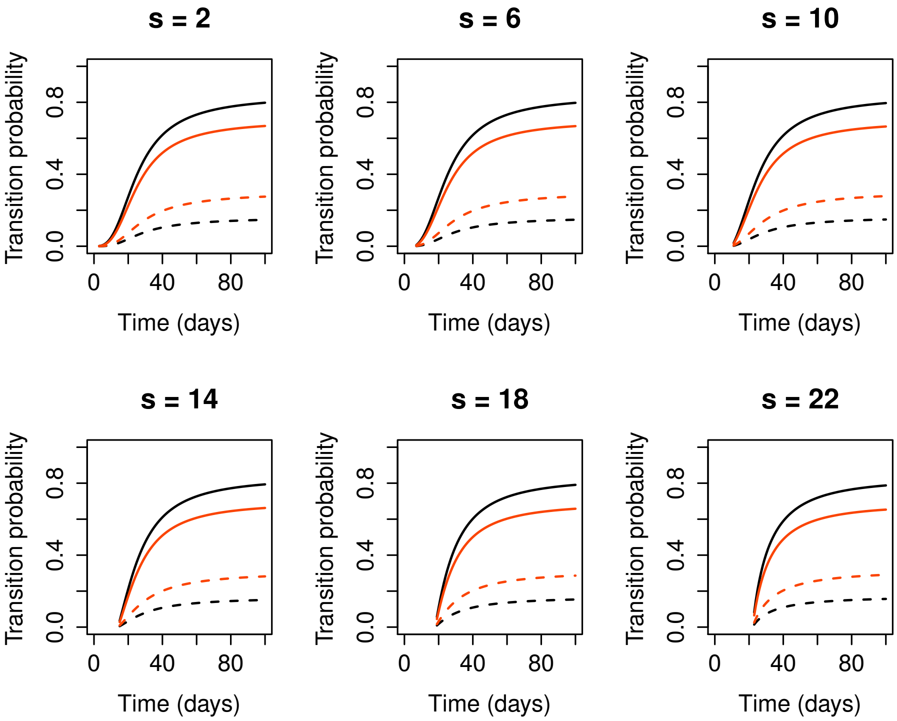


Figure 5.18: Estimated transition probabilities $\hat{P}_{12a}(t|s)$ (full line) and $\hat{P}_{12b}(t|s)$ (dashed line) using the estimates in the expanded model with covariates for a man (black) and a woman (red) both 60 years old. None of the covariates here are statistically significant.

death is mostly caused by a prolonged stay for the infected patients.

We have included covariates for age and sex in the models. Using a likelihood ratio test it was found that the ordinary model without covariates was a statistically significantly better fit than with covariates. However, in the expanded model the model with covariates was a better fit than the model without covariates. This is because it was only the age-covariate for death without prior pneumonia $\varphi_{2b,age}$ that was statistically significant, and age turned out to have larger effect on death without pneumonia than on discharge without pneumonia. In the ordinary model, the covariates for the combined endpoints had to explain both death and discharge, and since these were quite different, the covariate was not significant.

Furthermore, the coefficients for the covariates must be interpreted carefully as they only apply on a subject-specific level. In the analysis in Section 5.2.4 we saw that the sign of the covariate coefficient for the conditional transition rates are

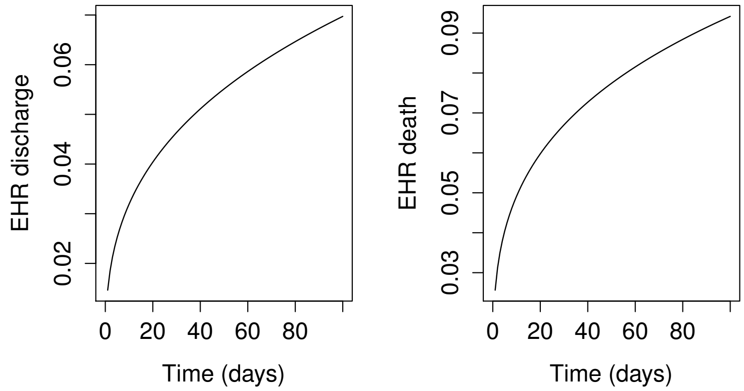


Figure 5.19: The conditional explanatory hazard ratio (EHR) for the dependence of discharge (left) and death (right) on pneumonia infection.

not directly associated with increased or decreased probability in the population. Even though increasing age had a slight positive effect on the conditional transition rate to discharge, that is we found $\hat{\varphi}_{2a,age} > 0$, young patients still had a higher probability of being discharged. This is because the probability of being discharged also depends on the conditional transition rate to death prior to pneumonia and the conditional transition rate to pneumonia infection.

The analysis of the pneumonia data set with the expanded model with covariates could be improved by trying to reduce the amount of covariates in the model. Some of the covariates in the model have high p-values, and a model without these can therefore be tested out. Reducing the parameters in the model will give parameter estimates that are more precise. A likelihood ratio test could then tell if the model with covariates for all transition rates is statistically significantly better than a model with covariates for only a few conditional transition rates.

Chapter 6

Concluding remarks

In this thesis we have explored the illness-death model with shared frailty introduced by Xu et al. (2010). Parametric functions for the conditional transition rates have been included, whereas Xu et al. (2010) use non-parametric conditional transition rates, and the model has been expanded by including an additional terminal state. We have also included covariates in the model using the conditional approach. In this chapter the main results from the thesis with regards to the model are discussed, and we present some suggestions for further work with the model.

6.1 Conclusion and main results

The illness-death model with shared frailty has a good way to handle the dependency between the non-terminal failure time and the terminal failure time by incorporating a frailty. With the conditional approach, the transition rates are Markov which makes it easy to construct a likelihood function. Since this approach does not go beyond the observable data, the problem with latent failure times is avoided.

Moreover, the simulation study in Chapter 4 shows that maximum likelihood estimation produces good estimates for the model parameters. The estimates for the real data sets also seem reasonable and produce sensible probability functions. Furthermore, it is simple to add more terminal states to the model and to include covariates. Because of this, the structure of the model can be simply customised to fit other situations than the ones studied in this thesis. For example a model with two non-terminal states.

It it important to note that the conditional transition rates and the marginal transition rates have different interpretations. In our opinion the conditional transition rates are not as useful as the marginal transition rates since γ is usually not known. Furthermore, the conditional transition rates only apply on a subject-specific level and not for the whole population of subjects. To make interpretation simpler the frailty can be integrated out of probability functions found using the conditional approach, which gives marginal measures. Marginal measures are often of interest to the health care, who wants to understand how non-terminal events like disease relapse or infections can affect other events for a population On the other hand, a patient may be more interested in the conditional measures as this has a person-specific interpretation. Since the interpretation of the marginal and conditional measures are different we recommend to use both side by side to obtain a full understanding of the semi-competing risks situation.

Finally, we have not performed a simulation study with the models where covariates have been included. This should be done so that we can evaluate the quality of the estimates and use the method with confidence, also when including covariates. However, since the estimates with covariates were close to the estimates without covariates, we can assume that the estimates are trustworthy.

6.2 Further work

A possible further expansion of the model could be to use another distribution class for the frailty. In frailty models, the standard assumption is to use a Gamma distribution for the frailty. The Gamma distribution is well known and has simple distributions, but there are no biological reasons for this choice (Hougaard, 1995). Other distribution classes that are used for the frailty, and could be possible to try are the lognormal distribution, the inverse Gaussian distribution and the positive stable distribution (Duchateau and Janssen, 2007).

A modelling framework for analysis of semi-competing risks based on the shared-frailty illness-death model by Xu et al. (2010) has recently been implemented in R, in the library `SemiCompRisks`. This modelling framework is proposed in Lee et al. (2015a) and Lee et al. (2015b), and allows for estimation and inference for regression parameters, the investigation of dependence structure, and prediction given covariates. The underlying model to characterise the dependency between T_1 and T_2 is the illness-death model with shared frailty, which means that they use the same definitions of the marginal and conditional transition rates as we have in this thesis. Furthermore, the novel modelling framework supports both frequentist and Bayesian analysis. Prior distributions must therefore be chosen for the frailty variance, the covariates and some other additional parameters of the model in the Bayesian analysis. The framework permits cluster-correlated data and both parametric and non-parametric specifications for a range of components which gives much flexibility in the model. The framework also allows for both the Markov model and the semi-Markov model for the conditional transition rates. Because this modelling framework essentially is the same as what we have implemented in this thesis, it would be interesting to study the framework more in detail. It would also be interesting to apply it to the bone marrow transplant data set analysed in Section 5.1 and the pneumonia data set analysed in Section 5.2 to see if we obtain the same results as in this thesis.

Bibliography

Beyersmann, J., Allignol, A., and Schumacher, M. (2011). *Competing risks and multistate models with R.* Springer Science & Business Media.

Casella, G. and Berger, R. L. (2002). *Statistical inference*, volume 2. Duxbury Pacific Grove, CA.

Duchateau, L. and Janssen, P. (2007). *The frailty model.* Springer Science & Business Media.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap.* CRC press.

Fine, J. P., Jiang, H., and Chappell, R. (2001). On semi-competing risks data. *Biometrika*, 88(4):907–919.

Fix, E. and Neyman, J. (1951). A simple stochastic model of recovery, relapse, death and loss of patients. *Human Biology*, 23(3):205–241.

Givens, G. H. and Hoeting, J. A. (2012). *Computational statistics*, volume 710. John Wiley & Sons.

Gutierrez, R. G. et al. (2002). Parametric frailty and shared frailty survival models. *Stata Journal*, 2(1):22–44.

Hougaard, P. (1995). Frailty models for survival data. *Lifetime data analysis*, 1(3):255–273.

Jiang, H., Fine, J. P., and Chappell, R. (2005). Semiparametric analysis of survival data with left truncation and dependent right censoring. *Biometrics*, 61(2):567–575.

Klein, J. P. and Moeschberger, M. L. (1997). *Survival analysis Techniques for Censored and truncated data*. Springer Science & Business Media.

Koivula, I., Sten, M., and Makela, P. H. (1994). Risk factors for pneumonia in the elderly. *The American journal of medicine*, 96(4):313–320.

Lee, K. H., Dominici, F., Schrag, D., and Haneuse, S. (2015a). Hierarchical models for semi-competing risks data with application to quality of end-of-life care for pancreatic cancer. *arXiv preprint arXiv:1502.00526*.

Lee, K. H., Haneuse, S., Schrag, D., and Dominici, F. (2015b). Bayesian semiparametric analysis of semicompeting risks data: investigating hospital readmission after a pancreatic cancer diagnosis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(2):253–273.

Lee, Y., Nelder, J. A., et al. (2004). Conditional and marginal models: another view. *Statistical Science*, 19(2):219–238.

Lindqvist, B. H. (2006). Competing risks. Department of Mathematical Sciences, Norwegian University of Science and Technology.

Meira-Machado, L. F., de Uña-Álvarez, J., Cadarso-Suárez, C., and Andersen, P. (2008). Multi-state models for the analysis of time-to-event data. *Statistical methods in medical research*.

Nguyen, K., Devidas, M., Cheng, S.-C., La, M., Raetz, E. A., Carroll, W. L., Winick, N. J., Hunger, S. P., Gaynon, P. S., and Loh, M. L. (2008). Factors influencing survival after relapse from acute lymphoblastic leukemia: a children's oncology group study. *Leukemia*, 22(12):2142–2150.

Nielsen, G. G., Gill, R. D., Andersen, P. K., and Sørensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian journal of Statistics*, pages 25–43.

Oriol, A., Vives, S., Hernández-Rivas, J.-M., Tormo, M., Heras, I., Rivas, C., Bethencourt, C., Moscardó, F., Bueno, J., Grande, C., et al. (2010). Outcome after relapse of acute lymphoblastic leukemia in adult patients included in four consecutive risk-adapted trials by the pethema study group. *Haematologica*, 95(4):589–596.

Paul, K. J., Walker, R. L., and Dublin, S. (2015). Anticholinergic medications and risk of community-acquired pneumonia in elderly adults: A population-based case–control study. *Journal of the American Geriatrics Society*, 63(3):476–485.

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

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rothman, K. J., Greenland, S., and Lash, T. L. (2008). *Modern epidemiology*. Lippincott Williams & Wilkins.

Safdar, N., Dezfulian, C., Collard, H. R., and Saint, S. (2005). Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. *Critical care medicine*, 33(10):2184–2193.

Selle, M. L. (2015). Competing risks regression in R. TMA4500 Industrial mathematics specialization project.

Touraine, C., Helmer, C., and Joly, P. (2013). Predictions in an illness-death model. *Statistical methods in medical research*, page 0962280213489234.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62.

Wolkewitz, M., Vonberg, R. P., Grundmann, H., Beyersmann, J., Gastmeier, P., Bärwolff, S., Geffers, C., Behnke, M., Rüden, H., and Schumacher, M. (2008). Risk factors for the development of nosocomial pneumonia and mortality on intensive care units: application of competing risks models. *Critical Care*, 12(2):R44.

Xu, J., Kalbfleisch, J. D., and Tai, B. (2010). Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics*, 66(3):716–725.

Appendix A

Additional theory

A.1 Shared frailty models

A frailty is an unobservable multiplicative effect on a hazard function which is assumed to be individual or group-specific (Gutierrez et al., 2002). It is a random variable with unit mean and a variance which can be estimated along with the other model parameters. The distribution of the frailty is often assumed to belong to a known distribution class, usually the Gamma distribution. Subjects with a frailty greater than one experience a greater hazard of failure than subjects with a frailty less than one.

There are two main types of frailty models. They are the *frailty model* and the *shared frailty model*. The frailty model is used with univariate data, while the shared frailty model is used with multivariate data and the frailty is shared among individuals such that it models intra-group correlation (Nielsen et al., 1992). Sharing a frailty generates dependency between the subjects. In the illness-death model with shared frailty however, the frailty is shared between the transition rates of each subject generating dependency between the transition rates.

Let the frailty be denoted by γ. Since it is a multiplicative effect, the hazard function conditional on the frailty is $\lambda(t|\gamma) = \gamma\lambda_0(t)$. $\lambda(t|\gamma)$ is then a conditional hazard function. The frailty γ can be integrated out by using the known distribution class of γ. Then the resulting hazard function, $\lambda(t)$, is called the marginal hazard function and is a function of the parameters in the frailty distribution and $\lambda_0(t)$.

A.2 Likelihood ratio test

In Chapter 5, the likelihood ratio test is used to compare the goodness of fit of two models, where one is a special case of the other. These two models are often referred to as the null hypothesis and the alternative hypothesis, where the null hypothesis is a special case of the alternative hypothesis.

The test is based on a test statistic, $\lambda(\boldsymbol{x})$ which is the likelihood ratio between the two models, and it expresses how many times more likely the data are under one model than the other,

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\Theta_0} L(\mu|\boldsymbol{x})}{\sup_{\Theta} L(\mu|\boldsymbol{x})} \tag{A.1}$$

where Θ_0 are the possible parameter values under the null hypothesis, and Θ are the possible parameter values under the alternative hypothesis (Casella and Berger, 2002).

The test statistic can be used to compute a p-value, or compared to a critical value to decide whether to reject the null hypothesis in favour of the alternative model. A result by Wilks (1938), says that as the sample size n approaches ∞, the test statistic $-2\log\lambda(\boldsymbol{x})$ will be asymptotically chi-squared distributed with degrees of freedom equal to the difference in parameters in the null hypothesis and the alternative hypothesis. In this thesis we use $-2\log\lambda(\boldsymbol{x})$ as test statistic in the likelihood ratio tests we perform.

Appendix B

Derivations

This appendix contains derivations and explanations for some of the equations in the thesis.

B.1 Deriving the marginal transition rates

In this section we will derive the formulas in equations (3.4)-(3.6). Since the derivation of $\lambda_1(t_1)$ and $\lambda_2(t_2)$ will be the same, we will not make the derivation for $\lambda_2(t_2)$.

All the integrals in this section have the same form

$$\int_0^\infty \gamma^r e^{\kappa\gamma} \frac{1}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \gamma^{\frac{1}{\theta}-1} e^{-\frac{\gamma}{\theta}} \,\mathrm{d}\gamma = \frac{\Gamma(\frac{1}{\theta}+r)}{\Gamma(\frac{1}{\theta})\theta^{\frac{1}{\theta}}(\kappa+\frac{1}{\theta})^{\frac{1}{\theta}+r}} \tag{B.1}$$

To develop the solution further we use the identity $\Gamma(t+1) = t\Gamma(t)$. To show how we arrive at the solution we solve the integral on the left side of (B.1). We start by doing a substitution and let $z = \gamma(\kappa + \frac{1}{\theta}), \mathrm{d}z/\mathrm{d}\gamma = \kappa + \frac{1}{\theta}$. This gives

$$\int_0^\infty \gamma^r e^{\kappa\gamma} \frac{1}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \gamma^{\frac{1}{\theta}-1} e^{-\frac{\gamma}{\theta}} \,\mathrm{d}\gamma = \frac{1}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \frac{1}{(\kappa+\frac{1}{\theta})^{\frac{1}{\theta}+r}} \int_0^\infty z^{\frac{1}{\theta}+r-1} e^{-z} \,\mathrm{d}z \tag{B.2}$$

The integral on the right side of (B.2) is the well-known definition of the Gamma function, $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\mathrm{d}x$, which makes

$$\frac{1}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \frac{1}{(\kappa+\frac{1}{\theta})^{\frac{1}{\theta}+r}} \int_0^\infty z^{\frac{1}{\theta}+r-1} e^{-z} \,\mathrm{d}z = \frac{\Gamma(\frac{1}{\theta}+r)}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})(\kappa+\frac{1}{\theta})^{\frac{1}{\theta}+r}}$$

which is what we wanted to show.

We start with the definition of $\lambda_1(t_1)$, which is

$$\lambda_1(t_1)dt_1 = P(t_1 \leq T_1 < t_1 + dt_1 | T_1 \geq t_1, T_2 \geq t_1).$$

Next, we use the law of total probability and integrate over γ, which gives

$$\lambda_1(t_1)dt_1 = \int_0^\infty P(t_1 \leq T_1 < t_1 + dt_1 | T_1 \geq t_1, T_2 \geq t_1, \gamma) P(\gamma | T_1 \geq t_1, T_2 \geq t_2) \, \mathrm{d}\gamma.$$

This is the same as integrating the conditional transition rate $\gamma\lambda_{01}(t_1)$ and the density of γ given T_1 and T_2, so we get

$$\lambda_1(t_1)dt_1 = \int_0^\infty \gamma\lambda_{01}(t_1)dt_1 \cdot f(\gamma | T_1 \geq t_1, T_2 \geq t_1) \, \mathrm{d}\gamma. \tag{B.3}$$

Now, we want to find an expression for the conditional density of γ. We use Bayes theorem which gives

$$f(\gamma | T_1 \geq t_1, T_2 \geq t_1) = \frac{g(\gamma) P(T_1 \geq t_1, T_2 \geq t_1 | \gamma)}{P(T_1 \geq t_1, T_2 \geq t_1)}. \tag{B.4}$$

Here, $g(\gamma)$ is the Gamma distribution given in (3.3),

$$P(T_1 \geq t_1, T_2 \geq t_1 | \gamma) = e^{-\gamma(\Lambda_{01}(t_1) + \Lambda_{02}(t_1))},$$

since the transition rate out of state 0 at time t_1 is $\gamma(\lambda_{01}(t_1) + \lambda_{02}(t_1))$, and by averaging $P(T_1 \geq t_1, T_2 \geq t_1 | \gamma)$ over γ using (B.1) with $r = 0$ we get

$$P(T_1 \geq t_1, T_2 \geq t_1) = [1 + \theta(\Lambda_{01}(t_1) + \Lambda_{02}(t_1)])^{-1/\theta}.$$

Inserting $f(\gamma | T_1 \geq t_1, T_2 \geq t_1)$ into (B.3) this yields

$$\lambda_1(t_1)dt_1 = \int_0^\infty \gamma\lambda_{01}(t_1)dt_1 g(\gamma) e^{-\gamma(\Lambda_{01}(t_1) + \Lambda_{02}(t_1))} (1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_1)])^{1/\theta} \mathrm{d}\gamma,$$

which is on the same form as (B.1) with $r = 1$. Finally, after solving the integral, the unconditional transition rate from state 0 to state 1 becomes

$$\lambda_1(t_1) = (1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_1)])^{-1} \lambda_{01}(t_1), \; t_1 > 0.$$

Next, we will do the derivation of $\lambda_{12}(t_2|t_1)$. Again, we start with the definition of the transition rate and integrate over γ

$$\lambda_{12}(t_2|t_1)dt_2 = \int_0^\infty P(t_2 \leq T_2 < t_2 + dt_2 | T_1 = t_1, T_2 \geq t_2, \gamma) P(\gamma | T_1 = t_1, T_2 \geq t_2) \, \mathrm{d}\gamma$$

where $t_1 < t_2$. This is the same as integrating the conditional transition rate $\gamma\lambda_{03}(t_2)$ from (3.2) and the density of γ given T_1 and T_2, so we get

$$\lambda_{12}(t_2|t_1)dt_2 = \int_0^\infty \gamma\lambda_{03}(t_2)dt_2 \cdot f(\gamma|T_1 = t_1, T_2 \geq t_2)\,\mathrm{d}\gamma. \tag{B.5}$$

Now, we want to find an expression for the conditional density of γ,

$$f(\gamma|T_1 = t_1, T_2 \geq t_2) = \frac{g(\gamma)P(t_1 \leq T_1 < t_1 + dt_1, T_2 \geq t_2|\gamma)}{P(t_1 \leq T_1 < t_1 + dt_1, T_2 \geq t_2)}. \tag{B.6}$$

Again, $g(\gamma)$ is the Gamma distribution $g(\gamma; 1/\theta, \theta)$,

$$P(t_1 < T_1 < t_1 + dt_1, T_2 \geq t_2|\gamma) = e^{-\gamma(\Lambda_{01}(t_1)+\Lambda_{02}(t_1)+\Lambda_{03}(t_1,t_2))}\gamma\lambda_{01}(t_1)dt_1,$$

and the denominator of (B.6) comes from integrating out γ from $P(t_1 < T_1 < t_1 + dt_1, T_2 \geq t_2|\gamma)$ using (B.1) with $r = 1$,

$$P(t_1 \leq T_1 < t_1 + dt_1, T_2 \geq t_2) = \frac{\lambda_{01}(t_1)}{(1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_1) + \Lambda_{03}(t_1, t_2)])^{1/\theta+1}}.$$

By inserting $f(\gamma|T_1 = t_1, T_2 \geq t_2)$ into (B.5) and solving the integral using the integral in equation (B.1) with $r = 2$, the unconditional transition rate from state 1 to state 2 becomes

$$\lambda_{12}(t_2|t_1) = (1 + \theta)(1 + \theta[\Lambda_{01}(t_1) + \Lambda_{02}(t_1) + \Lambda_{03}(t_1, t_2)])^{-1}\lambda_{03}(t_2),\ 0 < t_1 < t_2.$$

B.2 Deriving the likelihood function

In this section we show how the likelihood in the general model is constructed. Recall the cases in Table 2.1. The contribution to the likelihood function from each subject will be one of these cases. To get a better understanding of why the likelihood is as it is we will go through all four cases and show the likelihood contribution from each case given the frailty γ. We have also integrated out the frailty to see the marginal contribution.

Case 1

In this case a subject will have a transition to state 1 at Y_1 and then a transition to state 2 at Y_2. None of the event times are censored i.e. $\delta_{i1} = \delta_{i2} = 1$. One must therefore consider the probability of no events until Y_1, then a transition to state 1 at Y_1 and the probability of transition to state 2 at Y_2 after Y_1 which is

$$L_{i,1}|\gamma = e^{-\gamma(\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1}))}\gamma\lambda_{01}(Y_{i1})e^{-\gamma\Lambda_{03}(Y_{i1},Y_{i2})}\gamma\lambda_{03}(Y_{i2}).$$

When averaging out γ the likelihood contribution becomes

$$L_{i,1} = \lambda_{01}(Y_{i1})\lambda_{03}(Y_{i2})(1+\theta)(1+\theta[\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1})+\Lambda_{03}(Y_{i1},Y_{i2})])^{-1/\theta-2} \quad \text{(B.7)}$$

Case 2

In this case a subject will have a transition to state 2 at Y_1 and no events until this, which means that $Y_{i1} = Y_{i2}$, $\delta_{i1} = 0$ and $\delta_{i2} = 1$. One must therefore consider the probability of no events until Y_1, then a transition to state 2 at Y_1 which is

$$L_{i,2}|\gamma = e^{-\gamma(\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1}))}\gamma\lambda_{02}(Y_{i2}).$$

Note that we have used both Y_1 and Y_2 even though they are the same in this case. This is only for simplicity later. When averaging out γ the likelihood contribution becomes

$$L_{i,2} = \lambda_{02}(Y_{i2})(1+\theta[\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1})])^{-1/\theta-1} \quad \text{(B.8)}$$

Case 3

In this case a subject will have a transition to state 1 at Y_1 and then censoring at Y_2, which means that $\delta_{i1} = 1$ and $\delta_{i2} = 0$. One must therefore consider the probability of no events until Y_1, then a transition to state 1 followed by no events until Y_2 which is

$$L_{i,3}|\gamma = e^{-\gamma(\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1}))}\gamma\lambda_{01}(Y_{i1})e^{-\gamma\Lambda_{03}(Y_{i1},Y_{i2})}.$$

When averaging out γ the likelihood contribution becomes

$$L_{i,3} = \lambda_{01}(Y_{i1})(1+\theta[\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1})+\Lambda_{03}(Y_{i1},Y_{i2})])^{-1/\theta-1} \quad \text{(B.9)}$$

Case 4

In this case a subject will be censored before having any events. This means that $\delta_{i1} = \delta_{i2} = 0$. In this case as in case 2, $Y_1 = Y_2$. One must consider the probability of no events until Y_1 which is

$$L_{i,4}|\gamma = e^{-\gamma(\Lambda_{01}(Y_{i1})+\Lambda_{02}(Y_{i1}))}.$$

When averaging out γ the likelihood contribution becomes

$$L_{i,4} = (1 + \theta[\Lambda_{01}(Y_{i1}) + \Lambda_{02}(Y_{i1})])^{-1/\theta} \tag{B.10}$$

By multiplying together the results from equations (B.7)-(B.10), raising to appropriate powers of δ_{i1} and δ_{i2} and taking the product over all $i = 1, ..., n$, the likelihood in the general model becomes the one in equation (3.9) which is

$$\begin{aligned} L_g = \prod_{i=1}^{n} \lambda_{01}(Y_{i1})^{\delta_{i1}} \lambda_{02}(Y_{i2})^{\delta_{i2}(1-\delta_{i1})} \lambda_{03}(Y_{i2})^{\delta_{i1}\delta_{i2}} (1+\theta)^{\delta_{i1}\delta_{i2}} \\ \Big(1 + \theta[\Lambda_{01}(Y_{i1}) + \Lambda_{02}(Y_{i1}) + \Lambda_{03}(Y_{i1}, Y_{i2})]\Big)^{-1/\theta - \delta_{i1} - \delta_{i2}}. \end{aligned}$$

B.3 Confidence intervals for the survival function of state 1

In this section we describe how the confidence intervals of the marginal survival function for time to state 1 are calculated for the power law and the log-linear law. Recall the marginal survival function for time to relapse $S_1(t)$ from equation (3.14). The logarithm of this survival function is

$$\log S_1(t) = -\frac{1}{\theta} \log(1 + \theta \Lambda_{01}(t)),$$

and the estimate is $\widehat{\log S_1}(t) = -\frac{1}{\hat\theta} \log(1 + \hat\theta \hat\Lambda_{01}(t))$. To compute the variance of this, we linearise by expanding around the true parameters of Λ_{01} and θ. For the power law, the logarithm of the marginal survival function is

$$\log S_1^p(t) = -\frac{1}{\theta} \log(1 + \theta \alpha_1 t^{\beta_1})$$

Expanding around the true parameters θ, α_1 and β_1 gives

$$\begin{aligned} \widehat{\log S_1^p}(t) \simeq & -\frac{1}{\theta} \log(1 + \theta \alpha_1 t^{\beta_1}) - (\hat\alpha_1 - \alpha_1) \frac{d}{d\alpha_1} \frac{1}{\theta} \log(1 + \theta \alpha_1 t^{\beta_1}) \\ & - (\hat\beta_1 - \beta_1) \frac{d}{d\beta_1} \frac{1}{\theta} \log(1 + \theta \alpha_1 t^{\beta_1}) - (\hat\theta - \theta) \frac{d}{d\theta} \frac{1}{\theta} \log(1 + \theta \alpha_1 t^{\beta_1}) \end{aligned}$$

By computing the derivatives and taking the variance of $\widehat{\log S_1^p(t)}$ we find

$$\begin{aligned}\mathrm{Var}(\widehat{\log S_1^p(t)}) \simeq & \left(-\frac{t^{\beta_1}}{1+\theta\alpha_1 t^{\beta_1}}\right)^2 \mathrm{Var}(\hat{\alpha}_1) + \left(-\frac{\alpha_1 t^{\beta_1}\log t}{1+\theta\alpha_1 t^{\beta_1}}\right)^2 \mathrm{Var}(\hat{\beta}_1) \\ &+ \left(\frac{\log(1+\theta\alpha_1 t^{\beta_1})}{\theta^2} - \frac{\alpha_1 t^{\beta_1}}{\theta(1+\theta\alpha_1 t^{\beta_1})}\right)^2 \mathrm{Var}(\hat{\theta}) \\ &+ \left(-\frac{t^{\beta_1}}{1+\theta\alpha_1 t^{\beta_1}}\right)\left(-\frac{\alpha_1 t^{\beta_1}\log t}{1+\theta\alpha_1 t^{\beta_1}}\right) \mathrm{Cov}(\hat{\alpha}_1, \hat{\beta}_1) \\ &+ \left(-\frac{t^{\beta_1}}{1+\theta\alpha_1 t^{\beta_1}}\right)\left(\frac{\log(1+\theta\alpha_1 t^{\beta_1})}{\theta^2} - \frac{\alpha_1 t^{\beta_1}}{\theta(1+\theta\alpha_1 t^{\beta_1})}\right) \mathrm{Cov}(\hat{\alpha}_1, \hat{\theta}) \\ &+ \left(-\frac{\alpha_1 t^{\beta_1}\log t}{1+\theta\alpha_1 t^{\beta_1}}\right)\left(\frac{\log(1+\theta\alpha_1 t^{\beta_1})}{\theta^2} - \frac{\alpha_1 t^{\beta_1}}{\theta(1+\theta\alpha_1 t^{\beta_1})}\right) \mathrm{Cov}(\hat{\beta}_1, \hat{\theta})\end{aligned}$$

In practice the variance and covariance are found from the Hessian and the estimated parameters are used for the parameters.

Next, we look at the logarithm of the marginal survival function in the log-linear law, which is

$$\log S_1^l(t) = -\frac{1}{\theta}\log\Big(1+\frac{\theta}{b_1}(e^{a_1+b_1 t} - e^{a_1})\Big)$$

Again, an approximation is done by expanding around the estimates $\hat{\theta}, \hat{a}_1$ and $\hat{b}_1$. The linearisation of $\log S_1^l(t)$ is

$$\begin{aligned}\widehat{\log S_1^l(t)} \simeq & -\frac{1}{\theta}\log\Big(1+\frac{\theta}{b_1}(e^{a_1+b_1 t} - e^{a_1})\Big) - (\hat{a}_1 - a_1)\frac{d}{da_1}\frac{1}{\theta}\log\Big(1+\frac{\theta}{b_1}(e^{a_1+b_1 t} - e^{a_1})\Big) \\ & - (\hat{b}_1 - b_1)\frac{1}{\theta}\log\Big(1+\frac{\theta}{b_1}(e^{a_1+b_1 t} - e^{a_1})\Big) - (\hat{\theta} - \theta)\frac{1}{\theta}\log\Big(1+\frac{\theta}{b_1}(e^{a_1+b_1 t} - e^{a_1})\Big)\end{aligned}$$

By computing the derivatives and taking the variance of $\widehat{\log S_1^l(t)}$ we find

$$\begin{aligned}
\mathrm{Var}(\widehat{\log S_1^l(t)}) \simeq & \left(-\frac{e^{a_1}(e^{b_1 t}-1)}{b+\theta e^{a_1}(e^{b_1 t}-1)}\right)^2 \mathrm{Var}(\hat{a}_1) + \left(-\frac{b_1 t e^{a_1+b_1 t}-(e^{a_1+b_1 t}-e^{a_1})}{b_1(\theta(e^{a_1+b_1 t}-e^{a_1})+b_1)}\right)^2 \\
& \cdot \mathrm{Var}(\hat{b}_1) + \left(\frac{\log\frac{\theta(e^{a_1+b_1 t}-e^{a_1})+b_1}{b_1}}{\theta^2} - \frac{e^{a_1+b_1 t}-e^{a_1}}{b_1\theta\left(\frac{\theta(e^{a_1+b_1 t}-e^{a_1})}{b_1}+1\right)}\right)^2 \mathrm{Var}(\hat{\theta}) \\
& + \left(-\frac{e^{a_1}(e^{b_1 t}-1)}{b+\theta e^{a_1}(e^{b_1 t}-1)}\right)\left(-\frac{b_1 t e^{a_1+b_1 t}-(e^{a_1+b_1 t}-e^{a_1})}{b_1(\theta(e^{a_1+b_1 t}-e^{a_1})+b_1)}\right)\mathrm{Cov}(\hat{a}_1,\hat{b}_1) \\
& + \left(-\frac{e^{a_1}(e^{b_1 t}-1)}{b_1+\theta e^{a_1}(e^{b_1 t}-1)}\right)\left(\frac{\log\frac{\theta(e^{a_1+b_1 t}-e^{a_1})+b_1}{b_1}}{\theta^2} - \frac{e^{a_1+b_1 t}-e^{a_1}}{b_1\theta\left(\frac{\theta(e^{a_1+b_1 t}-e^{a_1})}{b_1}+1\right)}\right)\mathrm{Cov}(\hat{a}_1,\hat{\theta}) \\
& + \left(-\frac{b_1 t e^{a_1+b_1 t}-(e^{a_1+b_1 t}-e^{a_1})}{b_1(\theta(e^{a_1+b_1 t}-e^{a_1})+b_1)}\right)\left(\frac{\log\frac{\theta(e^{a_1+b_1 t}-e^{a_1})+b_1}{b_1}}{\theta^2} - \frac{e^{a_1+b_1 t}-e^{a_1}}{b_1\theta\left(\frac{\theta(e^{a_1+b_1 t}-e^{a_1})}{b_1}+1\right)}\right) \\
& \cdot \mathrm{Cov}(\hat{b}_1,\hat{\theta})
\end{aligned}$$

The limits for the confidence intervals for the marginal survival time to relapse are then

$$\exp\left[\log S_1(t) \pm 1.96\sqrt{\mathrm{Var}(\widehat{\log S_1^p(t)})}\right] \tag{B.11}$$

in the power law and

$$\exp\left[\log S_1(t) \pm 1.96\sqrt{\mathrm{Var}(\widehat{\log S_1^l(t)})}\right] \tag{B.12}$$

in the log-linear law.

Appendix C

R code

C.1 Simulation functions

C.1.1 The power law

The function for simulating data from the illness-death model with the power law as model for the conditional transition rates. The function calls on four other functions which are given below `SimData.power()`.

```
SimData.power=function(alpha1,alpha2,alpha3,beta1,beta2,beta3,theta,n){
  # Simulate n observations of semi-competing risks data with
  # censoring. The conditional transition rates have parametric formula
  # lambda(t) = alpha*beta*t^(beta-1)
  # The censoring distribution is a mixture distribution with equal
  # weights on a uniform distribution between 5 and 10, and a
  # point mass at 10
  # Input:
  # alpha1: the alpha parameter in lambda_01
  # alpha2: the alpha parameter in lambda_02
  # alpha3: the alpha parameter in lambda_03
  # beta1: the beta parameter in lambda_01
  # beta2: the beta parameter in lambda_02
  # beta3: the beta parameter in lambda_03
  # theta: the variance of the frailty parameter
  # n: the number of observations
  # Output:
```

```
# a data frame with the components Y1, Y2, delta1, delta2

data_sim = matrix(data=NA, nrow = n, ncol = 4)
for (i in 1:n){
  # Generate a frailty for each subject
  if (theta == 0){
    gamma = 1
  }
  else {
    gamma = rgamma(1, shape = 1/theta, scale = theta)
  }
  # Simulate the first event time
  t0 = 0
  while(t0 ==0){
    u0 = runif(1)
    f0 =function(t){Find_t0(t,alpha1,alpha2,beta1,beta2,gamma,u0)}
    t0 = uniroot(f0, c(0,1000000000000))$root
    }
  # Find the first transition
  prob = lambda(alpha1,beta1,t0) / (lambda(alpha1, beta1, t0) +
                                     lambda(alpha2, beta2, t0))
  uni = runif(1)
  if (uni <= prob){
    # Transition to state 1
    Y1 = t0
    delta1 = 1
    # Simulate the second event time
    u2 = runif(1)
    f2 = function(t2) {Find_t2(t2, t0, alpha3, beta3, gamma, u2)}
    t2 = uniroot(f2, c(t0, 10000000000))$root
    Y2 = t2
  }
  else{
    # Transition directly to state 2
    Y2 = t0
    Y1 = Y2
    delta1 = 0
    }
    # Include independent censoring
    cens_time = cens_time_sim()
    if(cens_time > Y2){
      delta2 = 1
    }else if(cens_time >= Y1  && cens_time<=Y2){
      delta2 = 0
      Y2 = cens_time
    }else{
      Y1 = cens_time
      Y2 = cens_time
      delta2 = 0
      delta1 = 0
    }
    data_sim[i,1] = Y1
    data_sim[i,2] = Y2
    data_sim[i,3] = delta1
    data_sim[i,4] = delta2
  }
  return(data.frame(Y1 = data_sim[,1], Y2 = data_sim[,2],
                    delta1 = data_sim[,3], delta2 = data_sim[,4]))
}

Find_t0 = function(t, alpha1, alpha2, beta1, beta2, gamma,u){
  # The root of this function is the first faliure time
  alpha1*t^beta1 + alpha2*t^beta2 + log(u)/gamma
}

Find_t2 = function(t2,t0, alpha3, beta3, gamma,u){
  # The root of this function is the second faliure time
  alpha3*t2^beta3 - alpha3*t0^beta3 + log(u)/gamma
}

lambda = function(alpha, beta, t){
  # The conditional transition rate
  lam = alpha*beta*t^(beta-1)
  return(lam)
}
```

```
cens_time_sim = function(){
  # A function that generates the censoring time
  uniform = runif(1)
  if(uniform < 0.5){
    cens = runif(1, min = 5, max =10)
    return(cens)
  }
  else{
    cens = 10
    return(cens)
  }
}
```

C.1.2 The log-linear law

The function for simulating data from the illness-death model with the log-linear law as model for the conditional transition rates. The function calls on four other functions which are given below `SimData.loglinear()`.

```
SimData.loglinear=function(alpha1,alpha2,alpha3,beta1,beta2,beta3,
theta,n){
  # Simulate n observations of semi-competing risks data with
  # censoring
  # The conditional transition rates have parametric formula
  # lambda(t) = exp(alpha+beta*t)
  # The censoring distribution is a mixture distribution with equal
  #  weights on a uniform distribution between 5 and 10, and a
  # point mass at 10
  # Input:
  # alpha1: the alpha parameter in lambda_01
  # alpha2: the alpha parameter in lambda_02
  # alpha3: the alpha parameter in lambda_03
  # beta1: the beta parameter in lambda_01
  # beta2: the beta parameter in lambda_02
  # beta3: the beta parameter in lambda_03
  # theta: the variance of the frailty parameter
  # n: the number of observations
# Output:
# a data frame with the components Y1, Y2, delta1, delta2
# NOTE: the function is not suited for using beta's = 0

data_sim = matrix(data=NA, nrow = n, ncol = 4)
for (i in 1:n){
  # Generate a frailty for each subject
  if (theta == 0){
    gamma = 1
  }
  else {
    gamma = rgamma(1, shape = 1/theta, scale = theta)
  }
  # Simulate first event time
  t0 = 0
  while(t0 ==0){
     u0 = runif(1)
     f0 =function(t){Find_t0_2(t,alpha1,alpha2,beta1,beta2,gamma,u0)}
     t0 = uniroot(f0, c(0,1000000000000))$root
  }
  # Find first transition
  prob = lambda2(alpha1,beta1,t0) / (lambda2(alpha1, beta1, t0) +
  lambda2(alpha2, beta2, t0))
  uni = runif(1)
  if (uni <= prob){
    # Transition to state 1
    Y1 = t0
    delta1 = 1
    # Simulate second event time
    u2 = runif(1)
    f2 = function(t2) {Find_t2_2(t2, t0, alpha3, beta3, gamma, u2)}
    t2 = uniroot(f2, c(t0, 10000000000))$root
    Y2 = t2
  }
  else{
    # Transition directly to state 2
    Y2 = t0
```

```
      Y1 = Y2
      delta1 = 0
    }
    # Include independent censoring
    cens_time = cens_time_sim()
    if(cens_time > Y2){
      delta2 = 1
    }else if(cens_time >= Y1  && cens_time<=Y2){
      delta2 = 0
      Y2 = cens_time
    }else{
      Y1 = cens_time
      Y2 = cens_time
      delta2 = 0
      delta1 = 0
    }
    data_sim[i,1] = Y1
    data_sim[i,2] = Y2
    data_sim[i,3] = delta1
    data_sim[i,4] = delta2
  }
  return(data.frame(Y1 = data_sim[,1], Y2 = data_sim[,2],
  delta1 = data_sim[,3], delta2 = data_sim[,4]))
}

Find_t0_2 = function(t, alpha1, alpha2, beta1, beta2, gamma,u){
  # The root of this function is the first faliure time
  (exp(alpha1+t*beta1)-exp(alpha1))/beta1 + (exp(alpha2+t*beta2) -
   exp(alpha2))/beta2 + log(u)/gamma
}

Find_t2_2 = function(t2,t0, alpha3, beta3, gamma,u){
  # The root of this function is the second faliure time
  (exp(alpha3 + t2*beta3)-exp(alpha3))/beta3 - (exp(alpha3+
  t0*beta3)- exp(alpha3))/beta3 + log(u)/gamma
}

lambda2 = function(alpha, beta, t){
  # The condtitional transition rate
  lam = exp(alpha + beta*t)
  return(lam)
}

cens_time_sim = function(){
  # A function that generates the censoring time
  uniform = runif(1)
  if(uniform < 0.5){
    cens = runif(1, min = 5, max =10)
    return(cens)
  }
  else{
    cens = 10
    return(cens)
  }
}
```

C.1.3 The expanded model with power law

The function for simulating data from the expanded illness-death model with the power law for the conditional transition rates. The function calls four other functions which are given below `SimData.expanded()`.

```
SimData.expanded = function(alpha1, alpha2, alpha3,alpha4,
alpha5, beta1, beta2, beta3,beta4, beta5, theta, n){
  # Simulate n observations of semi-competing risks data for the
  # expanded illness-death model with censoring
  # The conditional transition rates have parametric formula
  # lambda(t) = alpha*beta*t^(beta-1)
  # The censoring distribution is a mixture distribution with equal
  # weights on a uniform distribution between 5 and 10, and a
  # point mass at 10
  # Input:
  # alpha1: the alpha parameter in lambda01
  # alpha2: the alpha parameter in lambda02a
  # alpha3: the alpha parameter in lambda02b
  # alpha4: the alpha parameter in lambda03a
# alpha5: the alpha parameter in lambda03b
# beta1: the beta parameter in lambda01
# beta2: the beta parameter in lambda02a
# beta3: the beta parameter in lambda02b
# beta4: the beta parameter in lambda03a
# beta5: the beta parameter in lambda03b
# theta: the variance of the frailty
# parameter
# n: the number of observations
# Output:
# a data frame with the components Y1, Y2, delta1, delta2,
# delta_a, delta_b

data_sim = matrix(data=NA, nrow = n, ncol = 7)
for (i in 1:n){
  # Generate a frailty for each subject
  if (theta == 0){
    gamma = 1
  }
  else {
    gamma = rgamma(1, shape = 1/theta, scale = theta)
  }
  # Simulate the first event time
  t0 = 0
  while(t0 ==0){
    u0 = runif(1)
    f0 = function(t) {Find_t0_exp(t, alpha1, alpha2,alpha3 ,beta1,
     beta2, beta3,gamma,u0)}
    t0 = uniroot(f0, c(0,1000000000000))$root
  }
  # Find the first transition
  prob1 = lambda(alpha1,beta1,t0) / (lambda(alpha1, beta1, t0) +
      lambda(alpha2, beta2, t0)   +lambda(alpha3, beta3, t0))
  uni = runif(1)
  if (uni <= prob1){
    # Transition to state 1
    Y1 = t0
  delta1 = 1
  # Simulate the second event time
  u2 = runif(1)
  f2 = function(t2) {Find_t2_exp(t2, t0, alpha4,alpha5, beta4,
  beta5, gamma, u2)}
  t2 = uniroot(f2, c(t0, 10000000000))$root
  Y2 = t2
  # Find which was the terminating event - 2a or 2b
  prob3 = lambda(alpha4,beta4,t2) / (lambda(alpha4, beta4, t2) +
          lambda(alpha5, beta5, t2))
  uni = runif(1)
  if(prob3 <= uni ){ # 2a
  event1 =1
  event2 = 0
  }
  else{ # 2b
    event1 = 0
      event2 = 1
  }
}
else{ # If did not have a transition to state 1
  # Find which was the terminating event - 2a or 2b
  prob2 = lambda(alpha2,beta2,t0) / (lambda(alpha2, beta2, t0) +
   lambda(alpha3, beta3, t0))
  uni = runif(1)
  if(uni <= prob2){
  # Transition directly to state 2a
  Y2 = t0
  Y1 = Y2
  delta1 = 0
  event1 = 1
  event2 = 0
  }
  else{
  # Transition directly to state 2b
    Y2 = t0
    Y1 = Y2
    delta1 = 0
    event1 = 0
    event2 = 1
   }
  }
  # Include independent censoring
  cens_time = cens_time_sim()
  if(cens_time > Y2){
   delta2 = 1
  }else if(cens_time >= Y1  && cens_time<=Y2){
   delta2 = 0
   Y2 = cens_time
  }else{
   Y1 = cens_time
   Y2 = cens_time
   delta2 = 0
   delta1 = 0
  }
  data_sim[i,1] = Y1
  data_sim[i,2] = Y2
  data_sim[i,3] = delta1
  data_sim[i,4] = delta2
  data_sim[i,5] = event1
  data_sim[i,6] = event2
 }
 return(data.frame(Y1 = data_sim[,1], Y2 = data_sim[,2],
   delta1 = data_sim[,3], delta2 = data_sim[,4],  event1 =
   data_sim[,5], event2 = data_sim[,6] ) )
}


Find_t0_exp = function(t, alpha1, alpha2, alpha3, beta1,
beta2,beta3, gamma,u){
 # The root of this function is the first faliure time in the
 # expanded model
 alpha1*t^beta1 + alpha2*t^beta2+  alpha3*t^beta3 +
  log(u)/gamma
```

```
}

Find_t2_exp = function(t2,t0, alpha4,alpha5,beta4,
beta5, gamma,u){
  # The root of this function is the second faliure time in
  # the expanded model
  alpha4*t2^beta4 - alpha4*t0^beta4 +alpha5*t2^beta5 -
  alpha5*t0^beta5 + log(u)/gamma
}

lambda = function(alpha, beta, t){
  # The transition rate
  lam = alpha*beta*t^(beta-1)
  return(lam)
}

cens_time_sim = function(){
  # A function that generates the censoring time
  uniform = runif(1)
  if(uniform < 0.5){
    cens = runif(1, min = 5, max =10)
    return(cens)
  }
  else{
    cens = 10
    return(cens)
  }
}
```

C.2 The log likelihood functions

The functions for computing the log likelihood in the general model and the restricted model with both the power law and the log-linear law.

C.2.1 The general model with power law

```
LogLikelihood.power = function(par, datasett){
  # A function that computes the log likeihood in the general model
  # where the transition rates have parametric formula
  # lambda(t) = alpha*beta*t^(beta-1)
  # Input:
  # par: a vector containing the parameters of the conditional
  # transition rates in the following sequence alpha1, alpha2,
  # alpha3, beta1, beta2, beta3, theta
  # datasett: a data frame with columnwise components in the
  # following sequence Y1, Y2, delta1, delta2
  # Output:
  # the log likelihood

  alpha1 = par[1]
  alpha2 = par[2]
  alpha3 = par[3]
  beta1 = par[4]
  beta2 = par[5]
  beta3 = par[6]
  theta = par[7]

  Y1 = datasett[,1]
  Y2 = datasett[,2]
  d1 = datasett[,3]
  d2 = datasett[,4]

  temp = (alpha1*beta1*Y1^(beta1-1))^d1*(alpha2*beta2*Y2^
  (beta2-1))^(d2*(1-d1))*(alpha3*beta3*Y2^(beta3-1))^
  (d1*d2)*(theta + 1)^(d1*d2)*(1 + theta*(alpha1*Y1^(beta1) +
   alpha2*Y1^(beta2) + alpha3*Y2^(beta3) - alpha3*Y1^(beta3)))^
   (-d1-d2- 1/theta)

  lglik = sum(log(temp))
  return(lglik)
}
```

C.2.2 The restricted model with power law

```
LogLikelihood.power_res = function(par, datasett){
  # A function that computes the log likeihood in the restricted
  # model where the conditional transition rates have parametric
  # formula lambda(t) = alpha*beta*t^(beta-1)
  # Input:
  # par: a vector containing the parameters of the conditional
  # transition rates in the  following sequence alpha1, alpha2,
  # beta1, beta2, theta
  # datasett: a data frame with columnwise components in the
  # following sequence Y1, Y2, delta1, delta2
  # Output:
  # the log likelihood

  alpha1 = par[1]
  alpha2 = par[2]
  beta1 = par[3]
  beta2 = par[4]
  theta = par[5]

  Y1 = datasett[,1]
  Y2 = datasett[,2]
  d1 = datasett[,3]
  d2 = datasett[,4]

  temp = (alpha1*beta1*Y1^(beta1-1))^d1 *(alpha2*beta2*Y2^
  (beta2-1))^d2 * (theta + 1)^(d1*d2) * (1 + theta*(alpha1*Y1^
  (beta1) + alpha2*Y2^(beta2)))^(-d1-d2- 1/theta)

  lglik = sum(log(temp))
  return(lglik)
}
```

C.2.3 The general model with log-linear law

```
LogLikelihood.loglinear = function(par, datasett){
  # A function that computes the log likelihood in the general model
  # where the conditional transition rates have parametric formula
  # lambda(t) = exp(alpha + beta*t)
  # Input:
  # par: a vector containing the parameters of the conditional transition
  # rates in the following sequence alpha1, alpha2, alpha3, beta1, beta2,
  # beta3, theta
  # datasett: a data frame with columnwise components in the
  # following sequence Y1, Y2, delta1, delta2
  # Output:
  # the log likelihood

  alpha1 = par[1]
  alpha2 = par[2]
  alpha3 = par[3]
  beta1 = par[4]
  beta2 = par[5]
  beta3 = par[6]
  theta = par[7]

  Y1 = datasett[,1]
  Y2 = datasett[,2]
  d1 = datasett[,3]
  d2 = datasett[,4]

  temp = log( exp( d1*(alpha1 + beta1*Y1) )*exp( d2*(1-d1)*
  (alpha2 + beta2*Y2) )*exp( d1*d2*(alpha3 + beta3*Y2) ) *
  (1+theta)^(d1*d2) *(1+ theta*( (exp(alpha1 + beta1*Y1) -
  exp(alpha1))/beta1 + (exp(alpha2 + beta2*Y1) - exp(alpha2))/
  beta2 + (exp(alpha3 + beta3*Y2) -exp(alpha3))/beta3 -
  (exp(alpha3 + beta3*Y1)- exp(alpha3))/beta3))^(-d1-d2-1/theta))

  lglik = sum(temp)
  return(lglik)
}
```

C.2.4 The restricted model with log-linear law

```
LogLikelihood.loglinear_res = function(par, datasett){
  # A function that computes the log likelihood in the
  # restricted model  where the conditional transition rates
  # have parametric formula lambda(t) = exp(alpha + beta*t)
  # Input:
  # par: a vector containing the parameters of the conditional
  # transition rates in the following sequence alpha1, alpha2,
  # alpha3, beta1, beta2, beta3, theta
  # datasett: a data frame with columnwise components in
  # the following sequence Y1, Y2, delta1, delta2
  # Output:
  # the log likelihood
  alpha1 = par[1]
  alpha2 = par[2]
  beta1 = par[3]
  beta2 = par[4]
  theta = par[5]

  Y1 = datasett[,1]
  Y2 = datasett[,2]
  d1 = datasett[,3]
  d2 = datasett[,4]


  temp =log( exp( d1*(alpha1 + beta1*Y1) + d2*(alpha2 +
   beta2*Y2))*(1+theta)^(d1*d2) * (1+ theta*( (exp(alpha1+
   beta1*Y1)- exp(alpha1))/beta1+(exp(alpha2 + beta2*Y2)-
   exp(alpha2))/beta2 ))^(-d1-d2-1/theta))

  lglik = sum(temp)
  return(lglik)
}
```

C.2.5 The general model with power law and covariates

```
LogLikelihoodFun_cov = function(par, datasett){
  # A function that computes the log likeihood in the general
  # model with covariates where the conditional transition
  # rates have parametric formula
  # lambda(t) = alpha*beta*t^(beta-1)*exp(cov * coeff)
  # Input:
  # par: a vector containing the parameters of the conditional
  # transition rates in the following sequence alpha1, alpha2,
  # alpha3,beta1, beta2, beta3, theta, cov1, cov2, cov3, where
  # each cov is a vector with parameter for age and sex
  # datasett: a data frame with columnwise components in
  # the following sequence Y1, Y2, delta1, delta2, x, y,
  # where x and y are the age and sex
  # Output:
  # the log likelihood

  alpha1 = par[1]
  alpha2 = par[2]
  alpha3 = par[3]
  beta1 = par[4]
  beta2 = par[5]
  beta3 = par[6]
  theta = par[7]
  cov1 = c(par[8], par[11])
  cov2 = c(par[9], par[12])
  cov3 = c(par[10],par[13])

  Y1 = datasett[,1]
  Y2 = datasett[,2]
  d1 = datasett[,3]
  d2 = datasett[,4]
  x = datasett[,8]
  y = datasett[,9]
  if(alpha1 > 0 & alpha2 > 0 & alpha3 > 0){
      temp = (alpha1*beta1*Y1^(beta1-1))^d1*(alpha2*
      beta2*Y2^(beta2-1))^(d2*(1-d1))*(alpha3*beta3*
      Y2^(beta3-1))^(d1*d2)*exp(d1*t(cov1[1])*x +
      d1*cov1[2]*y  + d2*(1-d1)*cov2[1]*x +d2*(1-d1)*
      cov2[2]*y +d1*d2*cov3[1]*x  + d1*d2*cov3[2]*y)*
      (theta + 1)^(d1*d2) *(1 + theta*(alpha1*Y1^(beta1)*
      exp(cov1[1]*x + cov1[2]*y) + alpha2*Y1^(beta2)*
      exp(cov2[1]*x + cov2[2]*y) + alpha3*Y2^(beta3)*
      exp(cov3[1]*x +cov3[2]*y) -alpha3*Y1^(beta3)*
      exp(cov3[1]*x + cov3[2]*y)))^(-d1-d2- 1/theta)

     lglik = sum(log(temp))
     return(lglik)
  }
  else{
     return(-Inf)
  }
}
```

C.2.6 The expanded model with power law

```
LogLikelihoodFun.exp = function(par, datasett){
  # A function that computes the log likeihood in the expanded
  # model where the
  # transition rates have parametric formula: lambda(t) =
  # alpha*beta*t^(beta-1)
  # For the terminal event there are two events, 2a and 2b
  # Input:
  # par: a vector containing the parameters of the transition
  # rates in the following sequence
  #     alpha1: the alpha parameter in lambda_01
  #     alpha2: the alpha parameter in lambda_02a
  #     alpha3: the alpha parameter in lambda_02b
  #     alpha4: the alpha parameter in lambda_03a
  #     alpha5: the alpha parameter in lambda_03b
  #     beta1: the beta parameter in lambda_01
  #     beta2: the beta parameter in lambda_02a
  #     beta3: the beta parameter in lambda_02b
#     beta4: the beta parameter in lambda_03a
#     beta5: the beta parameter in lambda_03b
#     theta: the variance of the frailty parameter
# datasett: a dataframe with columnwise components in
# the following sequence
# Y1, Y2, delta1, delta2, di, de.
# di (discharge) and de (death) are 0 and 1 depending on
# which is the terminal event
# Output:
# the log likelihood

alpha1 = par[1]
alpha2 = par[2]
alpha3 = par[3]
alpha4 = par[4]
alpha5 = par[5]
beta1 = par[6]
beta2 = par[7]
beta3 = par[8]
beta4 = par[9]
beta5 = par[10]
theta = par[11]

Y1 = datasett[,1]
Y2 = datasett[,2]
d1 = datasett[,3]
d2 = datasett[,4]
di = datasett[,5]
de = datasett[,6]

if(alpha1 > 0 & alpha2 > 0 & alpha3 > 0 & alpha4 >0 &
alpha5 > 0){
 temp = log( (lam(alpha1, beta1, Y1))^d1 * (lam(alpha2,
 beta2, Y2))^(d2*(1-d1)*di) *(lam(alpha3, beta3, Y2))^
 (d2*(1-d1)*de)*(lam(alpha4, beta4,Y2))^(d1*d2*di) *
 (lam(alpha5, beta5,Y2))^(d1*d2*de)*(1 + theta)^(d1*d2)*
 (1 + theta*( Lam(alpha1,beta1, Y1) + Lam(alpha2, beta2,
   Y1) + Lam(alpha3, beta3, Y1) +Lam(alpha4,beta4,Y2) -
   Lam(alpha4, beta4, Y1) + Lam(alpha5, beta5, Y2)  -
    Lam(alpha5, beta5, Y1) ) )^(-d1-d2-1/theta)  )

   lglik = sum((temp))
   return(lglik)
  }
  else {
    temp = -Inf
    return(temp)
  }
}

lam = function(alpha, beta,t){
  temp = alpha*beta*t^(beta-1)
  return(temp)
}

Lam = function(alpha, beta,t){
  temp = alpha*t^(beta)
  return(temp)
}
```

C.2.7 The expanded model with power law and covariates

```
LogLikelihoodFun_exp_cov = function(par, datasett){
  # A function that computes the log likelihood in the
  # expanded model  where the conditional transition rates have
  # parametric formula
  # lambda(t) = alpha*beta*t^(beta-1)*exp(coeff * cov)
  # For the terminal event there are two categories
  # Input:
  # par: a vector containing the parameters of the conditional
  # transition rates in the following sequence alpha1, alpha2,
  # alpha3, beta1, beta2, beta3, theta, cov1, cov2, cov3, cov4,
  # cov5, where each cov is a vector with  coefficients for age
  # and sex
# datasett: a data frame with columnwise components in
# the following sequence  Y1, Y2, delta1, delta2,
# di, de, x, y, where di and de are 0 and 1 depending on
# which is the terminal event
# x and y are covariates for age and sex
# Output:
# the log likelihood

alpha1 = par[1]
alpha2 = par[2]
alpha3 = par[3]
alpha4 = par[4]
alpha5 = par[5]
beta1 = par[6]
beta2 = par[7]
beta3 = par[8]
beta4 = par[9]
beta5 = par[10]
theta = par[11]
cov1 = c(par[12], par[17])
cov2 = c(par[13], par[18])
cov3 = c(par[14], par[19])
cov4 = c(par[15], par[20])
cov5 = c(par[16], par[21])


Y1 = datasett[,1]
Y2 = datasett[,2]
d1 = datasett[,3]
d2 = datasett[,4]
di = datasett[,5]
de = datasett[,6]
x = datasett[,7]
y = datasett[,8]

  if(alpha1 > 0 & alpha2 > 0 & alpha3 > 0 & alpha4 >0 &
  alpha5 > 0){
      temp = log( (lam(alpha1, beta1, Y1))^d1 * (lam(alpha2,
      beta2, Y2))^(d2*(1-d1)*di) *(lam(alpha3, beta3, Y2))^
      (d2*(1-d1)*de)*(lam(alpha4, beta4,Y2))^(d1*d2*di)*
      (lam(alpha5, beta5,Y2))^(d1*d2*de)*exp(d1*cov1[1]*x +
      d1*cov1[2]*y  + d2*(1-d1)*di*cov2[1]*x +d2*(1-d1)*di*
      cov2[2]*y + d2*(1-d1)*de*cov3[1]*x  + d2*(1-d1)*de*
      cov3[2]*y + d1*d2*di*cov4[1]*x +d1*d2*di*cov4[2]*y +
      d1*d2*de*cov5[1]*x  +  d1*d2*de*cov5[2]*y) * (1+
      theta)^(d1*d2)*(1 + theta*(Lam(alpha1, beta1, Y1)*
      exp(cov1[1]*x + cov1[2]*y) + Lam(alpha2,beta2, Y1)*
      exp(cov2[1]*x+ cov2[2]*y) + Lam(alpha3, beta3, Y1)*
      exp(cov3[1])*x + cov3[2]*y) +Lam(alpha4,beta4,Y2)*
      exp(cov4[1]*x+cov4[2]*y) - Lam(alpha4, beta4, Y1)*
      exp(cov4[1]*x+ cov4[2]*y) + Lam(alpha5, beta5, Y2)*
      exp(t(cov5[1])*x+ cov5[2]*y)  - Lam(alpha5,beta5, Y1)*
      exp(t(cov5[1])*x+ cov5[2]*y) ) )^(-d1-d2-1/theta)  )

      lglik = sum((temp))
      return(lglik)
    }
    else {
      temp = -Inf
      return(temp)
    }
}

lam = function(alpha, beta,t){
  temp = alpha*beta*t^(beta-1)
  return(temp)
}

Lam = function(alpha, beta,t){
  temp = alpha*t^(beta)
  return(temp)
}
```