Chapter 16
Multi-state Models Used in Oncology Trials

Birgit Gaschler-Markefski, Karin Schiefele, Julia Hocke, and Frank Fleischer

Abstract Among the surrogate endpoints for overall survival (OS) in oncological trials, progression-free survival (PFS) is used as an important endpoint especially in first or second line of cancer therapies. Basic formulae for the determination of sample sizes based on time to event data can be found in the literature. Assumptions about the distributions of the survival time for OS and PFS, the accrual time and the censoring time are of key importance. Most often only uniformly distributed patient accrual and no censoring are mentioned, whereas the event time is assumed to be exponentially distributed. Considering the dependence between PFS and OS, we will investigate how a three-state model that includes states of progression/response and death can be used for a joint modelling of progression-free survival and overall survival. Sample size/power calculations are discussed for the three-state model and compared to the estimations based on exponentially distributed OS times. These sample size calculations are based on the assumption of piecewise uniformly accrual and exponentially distributed censoring time. The new three-state model approach results in a 10–30 % lower sample size and a corresponding higher power. An application to a Phase III lung cancer trial illustrates how the new approach can be successfully applied to the planning of a trial and to the monitoring of the needed events for the PFS and OS analyses.

16.1 Introduction

Oncological trials are often performed as event-driven trials, i.e. trial length and analysis time points are tied to the occurrence of a specific number of events. The most commonly used endpoint for new anticancer drug studies is overall survival

B. Gaschler-Markefski (✉) • J. Hocke • F. Fleischer
Department of Biostatistics, Boehringer Ingelheim Pharma GmbH and Co. KG, Biberach, Germany
e-mail: birgit.gaschler-markefski@boehringer-ingelheim.com; julia.hocke@boehringer-ingelheim.com; frank.fleischer@boehringer-ingelheim.com

K. Schiefele
Department of Epidemiology and Medical Biometry, University Ulm, Ulm, Germany
e-mail: karin.schiefele@uni-ulm.de

K. van Montfort et al. (eds.), *Developments in Statistical Evaluation of Clinical Trials*, DOI 10.1007/978-3-642-55345-5_16

(OS). If a patient develops progression of the tumor, then the therapy will be stopped and the patient will be switched to another (probably new) anticancer therapy. With regard to this, progression-free survival (PFS) is also used as a trial endpoint especially in early stages of cancer therapy. We will analyse OS mathematically by incorporation of the PFS information via a multi-state model. Multi-state models are probabilistic models which allow for studying transitions of a subject (in this context a person or patient) between different states over the course of time. In this chapter, an introduction to the basic concepts of multi-state modelling will be given and models commonly used in medical contexts, especially in oncology, will be presented.

In oncology as well as other indications like stroke or asthma a time to event outcome is often used as primary endpoint. For operational aspects it may be important to plan the time points of the final analysis and possible interim analyses. The time points of the interim analyses and final analysis in time to event studies are in most cases driven by the needed number of events (*landmark event number*). Therefore, a precise monitoring and prediction of the time point for the landmark event number is needed. The estimation of this time point with respect to OS can be derived based on different assumptions on the distribution of the lost-to-follow-up and the overall survival. For estimation of the time to occurrence of the landmark event number in this article an illness-death model, i.e. a three-state model for OS, is applied instead of the frequently used but oversimplifying assumption of exponentially distributed OS. An application to a phase III lung cancer trial illustrates how the new approach can be successfully applied to monitor event numbers for the OS analyses.

This chapter is structured as following: Different kinds of relevant multi-state models will be defined and their application to different contexts given. Of special interest is the three-state model for estimation of the time point of the landmark event number. Therefore, in the second part, after introducing the model assumptions as well as deriving relevant distributions, the expected number of events depending on the current time point and the planned accrual period will be derived. Based on this, the predicted landmark event time may be derived. We will compare the three-state model with an alternative one, which is restricted to exponentially distributed OS and does not account for progression. How the choice of the model influences sample size and power calculations is shown in an example. The performance of our new approach is demonstrated on real data of a non-small cell lung cancer trial.

16.2 Background Information

This section is based on models for the analysis of data with the primary endpoint being the time until occurrence of a certain event, which is also called *failure*. In the following an overview about common multi-state models will be given, whereas the kind of model is defined by the types of states it is consisting of.

16.2.1 Overview of Multi-state Models

The nonnegative random variable T corresponds to the period of time lasting from the initial time t_0 (e.g. time point of birth, randomisation, etc., mostly $t_0 = 0$) to the occurrence of the event of interest. In accordance to common terminology T is assumed to be continuous on $\mathbb{R}_+$. For analysis of discrete failure time distributions see for example [20].

Definition 16.1. A right-continuous piecewise constant stochastic process $X(t)$, $t \in [0, \infty)$ with a finite state space $S = \{1, \ldots, n\}$, $n \in \mathbb{N}$, is called a *multi-state model* (MSM).

The value of the process corresponds to the state occupied at time t and the *initial distribution* of the stochastic process is noted by $\pi_s(0) = \mathbb{P}(X(0) = s)$ for $s \in S$ (cf. [25]). The shift from one state to another is referred to as a *transition* or an *event*.

The probability for being in state $j \in S$ at time $t \in \mathbb{R}_+$ given that the process started in $i \in S$ at $u \in \mathbb{R}_+$, $u < t$, is called the *transition probability* and is noted by

$$p_{i,j}(t, u) = \mathbb{P}(X(u) = j \,|\, X(t) = i, \mathfrak{H}_t)\,, \tag{16.1}$$

whereas $\mathfrak{H}_t$ denotes the *history* of the process $X(.)$ (a σ-algebra in mathematical terms). $\mathfrak{H}_t$ consists of all the information of the process from the initial time (mostly time point 0) until t, i.e. all of the previous states and related times of transition in the interval $[0, t]$. Based on (16.1), the *state probabilities* $\pi_j(t) = \mathbb{P}(X(t) = j)$ are

$$\pi_j(t) = \sum_{i \in S} \pi_i(0)\, p_{i,j}(0, t) \tag{16.2}$$

for $j \in S$ and $t \in \mathbb{R}_+$. The *transition intensity* (also called transition rate, hazard function or (age-specific) failure rate) is defined by

$$\alpha_{i,j}(t) = \lim_{\Delta t \searrow 0} \frac{p_{i,j}(t, t + \Delta t)}{\Delta t}\,. \tag{16.3}$$

The $\alpha_{i,j}(t)$ gives the instantaneous event (or failure) rate at time t, provided the individual has been event-free until t. Consequently, the product $\alpha_{i,j}(t)\Delta t$ corresponds to the approximate probability of an event in $[t, t + \Delta t)$, given there has been no event until t (cf. [23]). A state $i \in S$ is called *absorbing* when it is not possible to leave this state once it has been reached and therefore, it holds $\alpha_{i,j}(t) = 0$ for all $t \in \mathbb{R}_+$ and $j \in S$. The time point when the process has left state $i \in S$ and first reaches $j \in S$ is called *transition time*.

Different kinds of models are defined by dependency of the transition intensity on time (cf. [25]):

1. **Time homogeneous models** have transition rates being constant over time, i.e. $p_{i,j}(t,u)$ depends only on $u-t$ and so it holds $p_{i,j}(t,u) = p_{i,j}(0,u-t)$.
2. **Markov models** have transition intensities only depending on the current state and neither on more of the previous states nor on future states, i.e. for $i, j \in S$ and $t, u \in \mathbb{R}_+$ with $0 \leq t < u$

$$\mathbb{P}(X(u) = j \,|\, X(t) = i, \mathfrak{H}_t) = \mathbb{P}(X(u) = j \,|\, X(t) = i)\,. \tag{16.4}$$

3. **Semi-(homogeneous) Markov models** have transition intensities depending on the current state $i \in S$ as well as on the time spent in state i.

In the following, only time-homogeneous Markov models will be analysed. For further description and examples on semi-(homogeneous) Markov models see e.g. [4] or [34].

A more detailed introduction to the theory of stochastic processes and multi-state models may be found in [3] (Chapter I). Also a good overview about multi-state models is given in [1, 25] and [18].

16.2.2 Types of Models

Uni-directional (or progressive) models allow for forward transitions only; once a state has been left, it can not be returned to it again. On the other hand in bi-directional (or alternating) models, the process can return to each state provided that it does not enter an absorbing state. Alternating models are relevant for e.g. reversible diseases but they would not be considered in detail in this chapter.

16.2.2.1 *k*-State Model

The k-state model is characterized by $k-1$ transient but uni-directional passable states ($k \in \mathbb{N}, k \geq 2$) and one absorbing state. Commonly, the first of the transient states is the starting point and the absorbing state is reachable from each of the transient states. Each of the following kinds of k-state models is Markovian. A method for testing the Markov property for example in a three-state progressive model is presented in [30].

Mortality Model

The simplest kind of the k-state model is the *mortality model* (*two-state model*) consisting of only two states (cf. Fig. 16.1). The process starts in '0' (alive) and stops after reaching the absorbing state '1' (dead). It holds $\alpha_{1,0}(t) = 0$ for all $t \in \mathbb{R}_+$ and the initial distribution is $\pi_0(0) = 1$. For example, Birnbaumer et al. apply this model to the kinetics of an enzyme [7].

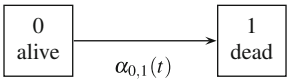


Fig. 16.1 Mortality model

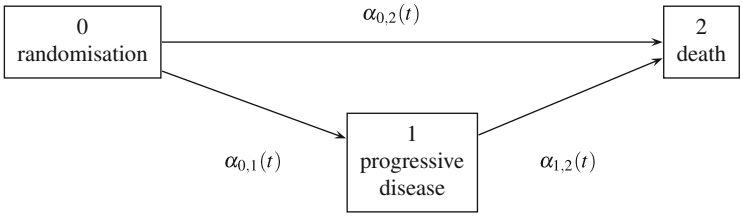


Fig. 16.2 Three-state model

Disability Model

The *disability model* (*three-state model*) is the specific multi-state model regarded in more detail in the subsequent sections. It consists of one absorbing and two transient states. Common applications of this model are state sequences like 'healthy – diseased – death' or as illustrated in Fig. 16.2 'disease – progressive disease (PD) – death'. The first mentioned setting enables inferences on the incidence of the regarded disease as well as on health rate whereas the decision if death rates of healthy subjects and patients differ may be problematic (cf. [25] p. 2). Andersen [2] applied the three-state model to the setting 'illness – comorbity – death'.

Obviously, for $\alpha_{0,1} = 0$ the disability model corresponds to the mortality model illustrated in Fig. 16.1. The transition probabilities introduced in (16.1) are for the three-state model given by (cf. [1, 25])

$$p_{0,0}(s,t) = \exp\left\{-\int_s^t \alpha_{0,1}(u) + \alpha_{0,2}(u)\,\mathrm{d}u\right\}, \tag{16.5}$$

$$p_{1,1}(s,t) = \exp\left\{-\int_s^t \alpha_{1,2}(u)\,\mathrm{d}u\right\}, \tag{16.6}$$

$$p_{0,1}(s,t) = \int_s^t p_{0,0}(s,u-)\alpha_{0,1}(u)\,p_{1,1}(u,t)\,\mathrm{d}u\,, \tag{16.7}$$

$$p_{2,2}(s,t) = 1\,, \tag{16.8}$$

$$p_{1,2}(s,t) = \int_s^t p_{1,1}(s,u-)\alpha_{1,2}(u)\,\mathrm{d}u\,, \tag{16.9}$$

$$p_{0,2}(s,t) = \int_s^t p_{0,0}(s,u-)\underbrace{[\alpha_{0,2}(u) + \alpha_{0,1}(u)\,p_{1,2}(u,t)]}_{=\alpha^*_{0,2}(u,t)}\,\mathrm{d}u\,. \tag{16.10}$$

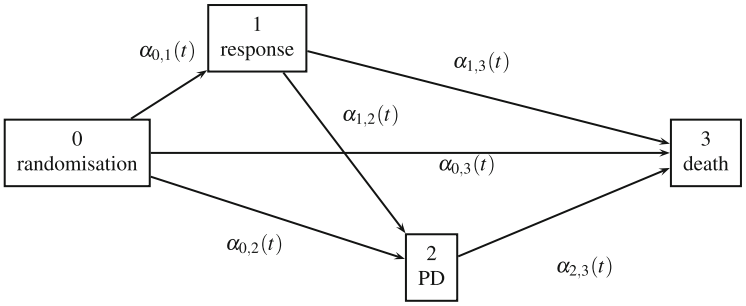


Fig. 16.3 Four-state model

The probability to stay in state 0 from time s until t is equal to the probability that the (random) time point of leaving this state is after t. It is well known that for a random variable T with hazard rate $h(.)$ it holds $\mathbb{P}(T > t) = \exp\left\{-\int_0^t h(u)\mathrm{d}u\right\}$. According to Fig. 16.2, leaving state 0 corresponds to switching into state 1 or 2 and since these are exclusive events, the hazard of the time point leaving state 0 is given by the sum of the single hazard rates. So Eq. (16.5) is verified, (16.6) can be shown analogously. Since $p_{0,1}(s,t)$ corresponds to staying in state 0 until an infinitesimal time unit before u, with u an arbitrary time between s and t, switching to state 1 at u and staying there until t, (16.7) is clear.

The overall transition rate $\alpha^*_{0,2}(u,t)$ corresponds in case of discrete time to the probability $\mathbb{P}(X(t) = 2|X(u) = 0)$ and is for continuous time equal to $\alpha_{0,2}(u) + \alpha_{0,1}(u)\, p_{1,2}(u,t)$.

In some settings it is necessary to consider also the state 'response', leading to a four-state model as shown in Fig. 16.3. Since patients having suffered progressive disease are assumed not being able to respond to the treatment without adjustment of dose/treatment, the state switches between 'progression' and 'response' are only one-directional.

In oncological trials, in particular in the metastatic setting, commonly the treatment is changed after occurrence of progressive disease in order to stop further progression. This new or adopted therapy is called second line treatment or kth line treatment in case of further previous switches. Modelling this proceeding leads to the k-state model (cf. Fig. 16.4).

16.2.2.2 Further Models

The *recurrent events model* consists of k transient states and optionally an absorbing state at the end of the line, whereas the transient ones only can be passed one after another. This model is applied if the event of interest occurs repeatedly, e.g. hospitalization, birth of a child, infections, recurrence of cancer, etc. A broad overview about the analysis of recurrent events is given in the book of Nelson [24], for further reading see also [21] or [8].

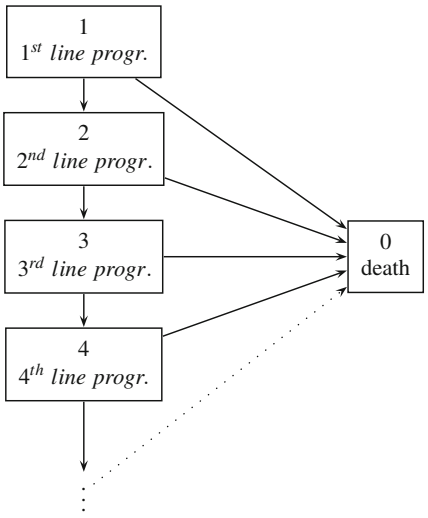


Fig. 16.4 k-state model

Adding further mutually exclusive absorbing states to the mortality model (i.e. death caused by different reasons) is called *competing risks model*. An introduction to the theory of those models is for example given in Beyersmann et al. [6] as well as in [28] and [13]. R. Chappell discusses in his manuscript two different methods for analysing competing risks models [9]. When switching to an absorbing state censors a non-terminal event, we are faced with *semi-competing risks models* which have been studied in [15] or [26]. Some authors (cf. [1], Section 3.6) call those models *bone marrow transplantion* model, since this setting is the common application. The *bivariate model* is used for modelling bivariate failure times, e.g. the survival of twins. For a more detailed description of this see for example [18], Section 5.2.

16.2.3 Recent Research in Multi-state Modelling

In recent research there are numerous applications of multi-stage modelling in the medical context given. Especially for models of chronic diseases this approach is frequently used. A three-state model for cognitive aging and suffering from dementia, with a kind of 'sub-state' (the pre-diagnosis) between 'healthy' and 'ill' and an increased transition rate after this additional state, is given by Dantan et al. [12]. They used a mixed-model approach and regarded non-informative censoring. An informative censoring mechanism is given in the model of Sweeting et al., which is a type of hidden Markov model for the analysis of disease progression in hepatitis C [31]. Lan and Datta compare a semi-Markov five-state model to a Markovian four-state model, both with assumption of log-normal as well as Weibull

distributed transition times and an uniformly or rather Weibull distributed censoring mechanism, in the context of measurement of sexual development of juvenile in puberty [22]. A four-state model with Weibull distributed transition rates for survival of dental fillings was developed by Joly et al. [19].

The most prominent area for multi-state modelling is the analysis of survival time and time until non-fatal events in oncology. There are numerous extensions and adjustments of the above basic modelling approaches. Only a few examples will be given. Porta et al. [27] combine a three-state model, including the possibility of disease recurrence, with a competing risk model and apply their dynamic model to patient data on bladder cancer. In some cases, the patient history has an effect on the transition rates and consequently the Markov property is no longer given. Putter and van Houwelingen model this by introduction of frailties (i.e. unobservable random interaction of survival times of different individuals). They apply this in the context of a three-state model, a competing risks model, a recurrent event model as well as a recurrent event model combined with mutually exclusive endpoints to breast cancer patient data [29]. Different kinds of multi-state Markov models with consideration of several progression stages are given in [35] and also applied to breast cancer data.

16.2.4 Questions to Be Solved/Data to Be Collected

Patients in oncological trials will typically receive several lines of treatment because of treatment adjustment after suffering progressive disease. For the sake of simplicity, in the following only a three-state model is investigated, i.e. each of the patients considered receives at most one change of treatment regime after progression. There are two endpoints being of interest in oncological trials, the primary endpoint is progression-free survival (PFS) and the key secondary one is overall survival (OS), both visualized gray-colored in Fig. 16.5. We are primarily

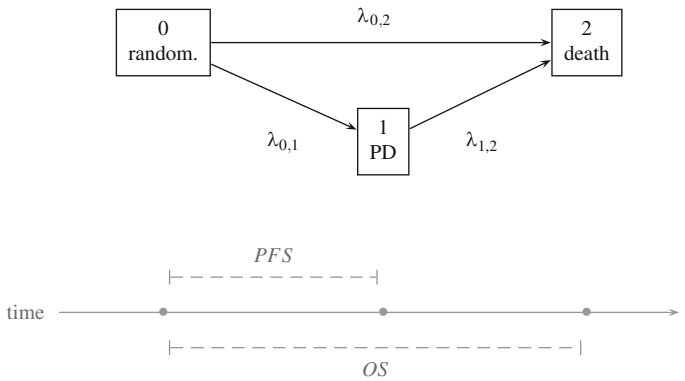


Fig. 16.5 Three-state model with constant transition rate

interested in information on overall survival. Instead of modelling OS via a single random variable, we can also incorporate the information on PFS by use of a three-state model for OS. A careful and precise definition of tumor progression is crucial [16] for accurate determination of PFS. Since there are no standard regulatory criteria, the RECIST criteria [14, 33] for solid tumours or other criteria can be used, e.g. for specific hematologic indications see [10] or [11].

Definition 16.2. The time from randomisation until death from any cause is called *overall survival* (OS).

Commonly, oncological trials are performed as event-driven trials, which means the trial length as well as the analysis time points are related to the occurrence of a specific number of events. So the study duration is a random quantity and the estimation of the time point t^* when the required number of events is observed is in question. At the begin of the study the estimated duration will be calculated and this value will be updated during the course of the trial. Furthermore, the time t^* of occurrence of the landmark event number is also relevant for planning of any interim analysis.

16.3 Statistical Methods

16.3.1 Model Assumptions

In the following, we will concentrate on the three-state model as given in Figs. 16.2 and 16.5.

16.3.1.1 Modelling of PFS and OS

For simplicity reasons, the transition rates (as defined in Eq. (16.3)) are assumed to be constant over time:

$$\begin{aligned} \alpha_{0,1}(t) &= \lambda_{0,1}\,, \\ \alpha_{0,2}(t) &= \lambda_{0,2}\,, \\ \alpha_{1,2}(t) &= \lambda_{1,2}\,, \end{aligned} \tag{16.11}$$

with $\lambda_{i,j} \in \mathbb{R}_+$ for $i = 0, 1,\ j = 1, 2$. From Eq. (16.11) follows that the random time to progression (TTP), i.e. the period between randomization and occurrence of progression, is exponentially distributed with parameter $\lambda_{0,1}$. Furthermore, the random time between progression and death as well as between randomization and dying directly is also exponentially distributed with parameter $\lambda_{1,2}$ and $\lambda_{0,2}$, respectively.

According to the definition of PFS, the PFS time corresponds to the waiting time of the stochastic process in the initial state 0, i.e. the PFS time is given by $T_0 = \min\{t \in \mathbb{R}_+ : X(t) \neq 0\}$. Based on this, the PFS time is exponentially distributed with parameter $\lambda_{0,1} + \lambda_{0,2}$.

In the regarded context, *death* is termed event. Let $f(t)$ and $g(t)$ be the density function of the event time and the lost to follow up time, respectively. The event times as well as the censoring times are assumed to be stochastically independent and identically distributed for all individuals $i = 1, \ldots, N$. Because of this, the subscript i may be suppressed for the censoring and event times in order to shorten expressions. The censoring process is assumed to follow an exponential distribution with parameter θ, i.e. $g(t) = \theta\, \mathrm{e}^{-\theta t}$ for $t \in \mathbb{R}_+$. Note that the quantities derived in the following may also be given in case that no censoring is assumed. Without consideration of censoring it is $\theta = 0$.

Since OS is the event of interest, the overall survival time will be denoted by the random variable T. The distribution of T is depending on the present state of the process, so the distribution function of OS is for $t \in \mathbb{R}_+$ given by

$$\begin{aligned} F_{T,C}(t) &= \mathbb{P}(T \leq t, C > T) = \mathbb{P}(C > T | T \leq t) \cdot \mathbb{P}(T \leq t) \\ &= \frac{\lambda_{0,1}\lambda_{1,2}}{(\lambda_{0,1} + \lambda_{0,2} - \lambda_{1,2})(\lambda_{1,2} + \theta)} \left(1 - \mathrm{e}^{-(\lambda_{1,2}+\theta)t}\right) \\ &\quad - \frac{(\lambda_{1,2} - \lambda_{0,2})(\lambda_{0,1} + \lambda_{0,2})}{(\lambda_{0,1} + \lambda_{0,2} - \lambda_{1,2})(\lambda_{0,1} + \lambda_{0,2} + \theta)} \left(1 - \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)t}\right). \end{aligned} \tag{16.12}$$

In case that there is no censoring regarded, the previous distribution function simplifies to

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leq t) \\ &= 1 - \frac{\lambda_{0,1}}{\lambda_{0,1} + \lambda_{0,2} - \lambda_{1,2}} \mathrm{e}^{-\lambda_{1,2}t} + \frac{\lambda_{1,2} - \lambda_{0,2}}{\lambda_{0,1} + \lambda_{0,2} - \lambda_{1,2}} \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t}. \end{aligned} \tag{16.13}$$

A more detailed derivation of the above equations may be found in the appendix and in Fleischer et al. [16]. For determination of corr(*PFS*, *OS*) see Fleischer et al. [16]. Heng et al. applied these results and showed that the PFS time can be used as an intermediate endpoint for OS [17].

For a patient being already progressive, the event time is the waiting time in state 1 and therefore the distribution of overall survival in this case is

$$F_{T,C}(t) = \frac{\lambda_{1,2}}{\lambda_{1,2} + \theta} \left(1 - \mathrm{e}^{-(\lambda_{1,2}+\theta)t}\right). \tag{16.14}$$

16.3.1.2 Modelling the Accrual Process

It is assumed that all patients enrolled during the accrual period will also be randomized, i.e. screening failures are not considered. The following derivations will be done for an one-arm trial, whereas generalizations to multi-arm trials work in an analogous manner (cf. [5]).

There are two different approaches for modelling the accrual process, the common one is a Poisson-process. Especially in case of numerous randomized patients, it is possible to loose restriction on randomized accrual by assumption of a fixed accrual rate $r \in \mathbb{R}_+$ over the whole time period.

At start of the trial, we assume a linear randomization with rate $r \in \mathbb{R}_+$. Therefore, the number of patients randomized until the current calendar time t_c is given by $N(t_c) = r \cdot t_c$. In general, the observed randomization rate at time $t > t_c$ will be different from r. Henceforth, from current time t_c randomization of the remaining $N - N(t_c)$ patients is assumed with constant rate $r(t_c)$, whereas

$$r(t_c) = \begin{cases} 0, & \text{if } N(t_c) \geq N\,, \\ \frac{N - r \cdot t_c}{a(t_c) - t_c}, & \text{else}\,, \end{cases} \tag{16.15}$$

for $t_c \in \mathbb{R}_+$ and $a(t_c)$ denoting the end of the randomization period. With $u \geq t_c$ a future time-point, the density of the randomization rate for the remaining randomization time is

$$r(t_c, u) = \frac{N - r \cdot t_c}{a(t_c) - t_c}\, \mathbb{I}_{(0, a(t_c) - t_c)}(u)\,, \tag{16.16}$$

because of assumption of uniformly accrual in the remaining time interval.

16.3.2 Prediction of OS Events

In the following, we will derive a closed formula for the expected number of events at a future time point t, depending on the current time point t_c. Based on this, the expected time point of the landmark number of events can be calculated. Let the number of events (i.e. deaths) observed until a certain time $t \in \mathbb{R}_+$ be given by the random variable $D(t)$. Then, $\mathbb{E}[D(t)|N, \mathfrak{H}_{t_c}]$ is the conditional expectation of the number of events that will be observed by calendar time $t > t_c$, given the data up to current calendar time t_c. The value in question is the predicted calendar time t^* when the required number of events $\hat{d}$ is expected, i.e. $\mathbb{E}[D(t^*)|N, \mathfrak{H}_{t_c}] = \hat{d}$.

The expected number of events is given by

$$\mathbb{E}[D(t)|N, \mathfrak{H}_{t_c}] = d(t_c) + \mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}] + \mathbb{E}[D_{NR}(t)|N, \mathfrak{H}_{t_c}]\,, \tag{16.17}$$

with $d(t_c) \in \mathbb{N}_0$ the number of events until the current time t_c, which is not a random variable but rather an observed quantity. $\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}]$ is the number of newly expected events between t_c and t of patients being already randomized at t_c given the data up to t_c and $\mathbb{E}[D_{NR}(t)|N, \mathfrak{H}_{t_c}]$ denotes the analogous quantity for patients randomized between t_c and t.

16.3.2.1 Calculation of $\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}]$

The conditional expectation $\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}]$ of the patients already randomized, alive and on study at time t_c who will have been observed to die by time t, has to be distinguished between patients who have already progressed until time t_c or not. Let $Y_i(t) = 0$ if patient i has not progressed until t and is under observation and at risk for an event at time t and $Y_i(t) = 1$ if the patient has already suffered progressive disease. Therefore, $\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}]$ is

$$\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}] = \mathbb{E}[D_R(t), Y(t_c) = 0|N, \mathfrak{H}_{t_c}] + \mathbb{E}[D_R(t), Y(t_c) = 1|N, \mathfrak{H}_{t_c}], \tag{16.18}$$

with $\mathbb{E}[D_R(t), Y(t_c) = 0|N, \mathfrak{H}_{t_c}]$ the expected number of events of patients not progressive until t_c and $\mathbb{E}[D_R(t), Y(t_c) = 1|N, \mathfrak{H}_{t_c}]$ the analogous quantity of patients already progressive at time t_c.

Let $E_i, i = 1, \ldots, N$ denote the random variable for the randomization time of the ith patient and let ϵ_i denote the observed randomization time of the ith patient. It is assumed that the randomization time E_i of every individual is stochastically independent from the associated event and censoring times. The randomization time of each individual is measured from $t = 0$, the calendar date when the first patient is randomized. The individual survival times (overall survival) and censoring times are measured from the calendar date of a patients randomization. The probability that the ith patient is at risk between $t_c - \epsilon_i$ and $t - \epsilon_i$, i.e. the probability that the ith patient has the event time within the time interval $(t_c - \epsilon_i, t - \epsilon_i)$ and does not get censored before the event, given that the patient survived uncensored at $t_c - \epsilon_i$, is denoted by $\mathbf{P}_i^{f,g}(t_c, t)$. It is

$$\begin{aligned}
\mathbf{P}_i^{f,g}(t_c, t) &= \mathbb{P}(T < C, T \in (t_c - \epsilon_i, t - \epsilon_i)|T > t_c - \epsilon_i, C > t_c - \epsilon_i) \\
&= \frac{\int_{t_c-\epsilon_i}^{t-\epsilon_i} f(u) \overbrace{\mathbb{P}(T < C|T = u)}^{=\mathbb{P}(C>u)} \, \mathrm{d}u}{(1 - F(t_c - \epsilon_i))(1 - G(t_c - \epsilon_i))} \\
&= \frac{F(t - \epsilon_i) - F(t_c - \epsilon_i) - \int_{t_c-\epsilon_i}^{t-\epsilon_i} f(u)G(u) \, \mathrm{d}u}{(1 - F(t_c - \epsilon_i))(1 - G(t_c - \epsilon_i))}.
\end{aligned}$$

The dependency of $\mathbf{P}_i^{f,g}(t_c, t)$ on the distribution of the event and censoring times is symbolized by the indexes f and g, the corresponding density functions. By the assumption of memoryless distributions for the event and the censoring time (i.e. for $F(.)$ and $G(.)$) it holds

$$\begin{aligned}\mathbf{P}_i^{f,g}(t_c, t) &= \mathbb{P}(T < C, T \in (0, t - t_c)) \\ &= \int_0^{t-t_c} f(u) \underbrace{\mathbb{P}(C > u)}_{=1-G(u)} \, \mathrm{d}u \\ &= F(t - t_c) - \int_0^{t-t_c} f(u)G(u) \, \mathrm{d}u \,,\end{aligned} \tag{16.19}$$

whereas the first transformation uses the definition of memoryless distributions. Obviously, in this case the risk probability of patient i is independent of the individual randomization time ϵ_i, $i = 1, \ldots, N$. In the following, we will provide that event times and censoring times follow memoryless distributions.

For the expectation $\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}]$ we get

$$\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}] = \sum_{i=1}^{N(t_c)} (1 - Y_i(t_c)) \cdot \mathbf{P}_i^{f_0,g}(t_c, t) + \sum_{i=1}^{N(t_c)} Y_i(t_c) \cdot \mathbf{P}_i^{f_1,g}(t_c, t) \,.$$

Since $\mathbf{P}_i^{f,g}(t_c, t)$ is independent of i (cf. (16.19)), we obtain

$$\begin{aligned}\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}] &= N_0(t_c) \cdot \mathbf{P}_i^{f_0,g}(t_c, t) \cdot \mathbb{I}_{Y(t_c)=0} \\ &\quad + (N(t_c) - N_0(t_c)) \cdot \mathbf{P}_i^{f_1,g}(t_c, t) \cdot \mathbb{I}_{Y(t_c)=1} \,,\end{aligned} \tag{16.20}$$

with $N_0(t_c)$ the number of patients not yet progressive and $N(t_c) - N_0(t_c)$ the number of randomized patients suffering progression until t_c. In the above equations, the index of the density f symbolises the corresponding distribution function of the event time, i.e. $f_0(t)$ denotes the density function of OS of a randomized patient (cf. (16.12)) and $f_1(t)$ is the density of OS time for a patient already progressive (cf. (16.14)).

The probability of dying between $t_c - \epsilon_i$ and $t - \epsilon_i$ given that PFS $> t_c - \epsilon_i$ equals the probability of dying before $t - t_c$, irrespective of the randomization time. By plugging $\mathbf{P}_i^{f_0,g}(t_c, t)$ into formula (16.20) we get

$$\begin{aligned}\mathbb{E}[D_R(t)|Y(t_c) = 0, N, \mathfrak{H}_{t_c}] &= \frac{N_0(t_c)}{\lambda_{0,1} + \lambda_{0,2} - \lambda_{1,2}} \left[\frac{\lambda_{0,1}\lambda_{1,2}}{\lambda_{1,2} + \theta} \left(1 - \mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}\right) \right. \\ &\quad \left. - \frac{(\lambda_{1,2} - \lambda_{0,2})(\lambda_{0,1} + \lambda_{0,2})}{\lambda_{0,1} + \lambda_{0,2} + \theta} \left(1 - \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)(t-t_c)}\right) \right] .\end{aligned} \tag{16.21}$$

If the patient has already progressed his further survival follows the distribution given in (16.14) and therefore

$$\mathbb{E}[D_R(t)|Y(t_c)=1,N,\mathfrak{H}_{t_c}] = (N(t_c)-N_0(t_c))\,\frac{\lambda_{1,2}[1-\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}]}{\lambda_{1,2}+\theta}\,. \tag{16.22}$$

Altogether, the expected number of events of the patients already randomized is given by

$$\begin{aligned}\mathbb{E}[D_R(t)|N,\mathfrak{H}_{t_c}] = {} & \frac{N_0(t_c)}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\left[\frac{\lambda_{0,1}\lambda_{1,2}}{\lambda_{1,2}+\theta}\left(1-\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}\right)\right.\\ & \left.-\frac{(\lambda_{1,2}-\lambda_{0,2})(\lambda_{0,1}+\lambda_{0,2})}{\lambda_{0,1}+\lambda_{0,2}+\theta}\left(1-\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)(t-t_c)}\right)\right]\\ & +\ (N(t_c)-N_0(t_c))\,\frac{\lambda_{1,2}[1-\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}]}{\lambda_{1,2}+\theta}\,.\end{aligned} \tag{16.23}$$

So the expected number of events in the subset of patients being already randomized, depends on the distribution parameters of the event times as well as on the number of patients suffering progression until t_c.

16.3.2.2 Calculation of $\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}]$

For determination of the expected number of events of patients not yet randomized, it has to be distinguished between three different scenarios.

Scenario 1 The randomization is finished, t_c is after end of randomization period $a(t_c)$. Thus, no more patients will be recruited after t_c and for $0 \le a(t_c) \le t_c < t$ it is $\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}] = 0$ (Fig. 16.6).

Scenario 2 The randomization is not yet finished, t_c is before end of randomization and the planned time of analysis t is after $a(t_c)$.The expected number of events for $0 \le t_c < a(t_c) \le t$ is

$$\begin{aligned}\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}] &= \int_0^{a(t_c)-t_c} r(t_c,u)\mathbb{P}(T<C, T\in(0,t-t_c-u))\,\mathrm{d}u\\ &= \int_0^{a(t_c)-t_c} r(t_c,u)\left[\int_0^{t-t_c-u} f(s)(1-G(s))\,\mathrm{d}s\right]\mathrm{d}u\,,\end{aligned} \tag{16.24}$$

with $r(t_c,u) = \frac{N-r\cdot t_c}{a(t_c)-t_c}$, $G(s) = \mathrm{e}^{-\theta s}$ and $f(s)$ the density function of OS which may be derived from (16.12) (Fig. 16.7).

Fig. 16.6 Scenario 1

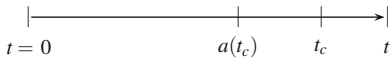


Fig. 16.7 Scenario 2

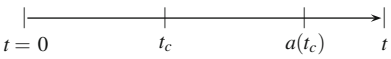
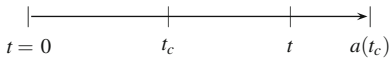


Fig. 16.8 Scenario 3



Scenario 3 The randomization is not yet finished. The planned time for interim analysis t is after t_c but before end of randomization (Fig. 16.8).

$$\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}] = \int_0^{t-t_c} r(t_c,t)\left[\int_0^{t-t_c-u} f(s)(1-G(s))\,\mathrm{d}s\right]\mathrm{d}u,$$

with the analogous variables as given in scenario 2.

With regard on the definition of randomization rate (cf. (16.16)), assumption of linear randomization and the above, $\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}]$ may be given in closed form:

$$\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}] = \begin{cases} 0, & \text{if } a(t_c) \le t_c < t\,, \\ \frac{N-N(t_c)}{a(t_c)-t_c}\int_0^{a(t_c)-t_c} \mathbf{P}_i^{f,g}(u,t-t_c)\,\mathrm{d}u, & \text{if } t_c < a(t_c) \le t\,, \\ \frac{N-N(t_c)}{a(t_c)-t_c}\int_0^{t-t_c} \mathbf{P}_i^{f,g}(u,t-t_c)\,\mathrm{d}u, & \text{if } t_c < t \le a(t_c)\,. \end{cases}$$

Insertion of the distribution function of the event and censoring times gives finally

$$\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}] =$$

$$= \begin{cases} 0, & \text{if } 0 \le a(t_c) \le t_c < t\,, \\ (N-N(t_c))\left[\frac{\lambda_{0,1}\lambda_{1,2}}{(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})(\lambda_{1,2}+\theta)}\left(1-\frac{\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-a(t_c))}-\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}}{(\lambda_{1,2}+\theta)(a(t_c)-t_c)}\right)-\right. \\ \left.-\frac{(\lambda_{1,2}-\lambda_{0,2})(\lambda_{0,1}+\lambda_{0,2})}{(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})(\lambda_{0,1}+\lambda_{0,2}+\theta)}\left(1-\frac{\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)(t-a(t_c))}-\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)(t-t_c)}}{(\lambda_{0,1}+\lambda_{0,2}+\theta)(a(t_c)-t_c)}\right)\right], \\ \qquad\qquad \text{if } 0 \le t_c < a(t_c) \le t\,, \\ \frac{N-N(t_c)}{a(t_c)-t_c}\left[\frac{\lambda_{0,1}\lambda_{1,2}}{(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})(\lambda_{1,2}+\theta)}\left(t-t_c-\frac{1-\mathrm{e}^{-(\lambda_{1,2}+\theta)(t-t_c)}}{\lambda_{1,2}+\theta}\right)-\right. \\ \left.-\frac{(\lambda_{1,2}-\lambda_{0,2})(\lambda_{0,1}+\lambda_{0,2})}{(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})(\lambda_{0,1}+\lambda_{0,2}+\theta)}\left(t-t_c-\frac{1-\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}+\theta)(t-t_c)}}{\lambda_{0,1}+\lambda_{0,2}+\theta}\right)\right], \\ \qquad\qquad \text{if } 0 \le t_c < t \le a(t_c)\,. \end{cases} \tag{16.25}$$

16.3.3 An Alternative Model

As mentioned in Sect. 16.2.1, if in the three-state model (cf. Fig. 16.5) the transition rate $\alpha_{0,1}(t)$ is equal to 0, we are faced with the mortality model of Fig. 16.1. Since $\alpha_{0,2}(t)$ is assumed to be a constant, $\lambda_{0,2} \in \mathbb{R}_+$, the transition time from state 0 to state 2 (death) is exponentially distributed. We will call this reduced model the *exponential model*.

The quantities derived above can also be given for the *exponential model* by assumption of $\lambda_{0,1} = \lambda_{1,2} = 0$ in Fig. 16.5. So the distribution function of overall survival is

$$F_{T,C}(t) = \frac{\lambda_{0,2}}{\lambda_{0,2} + \theta} \left(1 - \mathrm{e}^{-(\lambda_{0,2}+\theta)(t)}\right), \tag{16.26}$$

with θ the distribution parameter of censoring time. This distribution function reduces to those of an exponentially distributed variable with parameter $\lambda_{0,2}$, when there is no censoring considered. Furthermore, from the previous subsection follows that it is

$$\mathbb{E}[D_R(t)|N, \mathfrak{H}_{t_c}] = N(t_c) \cdot \frac{\lambda_{0,2}}{\lambda_{0,2} + \theta} \left(1 - \mathrm{e}^{-(\lambda_{0,2}+\theta)(t-t_c)}\right) \tag{16.27}$$

and

$$\mathbb{E}[D_{NR}(t)|N, \mathfrak{H}_{t_c}]$$

$$= \begin{cases} 0, & \text{if } 0 \le a(t_c) \le t_c < t\,, \\ \frac{(N-N(t_c))\lambda_{0,2}}{\lambda_{0,2}+\theta} \left[1 - \frac{1}{(\lambda_{0,2}+\theta)(a(t_c)-t_c)} \left(\mathrm{e}^{-(\lambda_{0,2}+\theta)(t_c-a(t_c))} - \mathrm{e}^{-(\lambda_{0,2}+\theta)(t-t_c)}\right)\right], & \text{if } 0 \le t_c < a(t_c) \le t\,, \\ \frac{(N-N(t_c))\lambda_{0,2}}{(\lambda_{0,2}+\theta)(a(t_c)-t_c)} \left[t - t_c - \frac{1}{\lambda_{0,2}+\theta} \left(1 - \mathrm{e}^{-(\lambda_{0,2}+\theta)(t-t_c)}\right)\right], & \text{if } 0 \le t_c < t \le a(t_c)\,. \end{cases}$$

16.3.4 Landmark Event Time

According to the formula of Schoenfeld (cf. [32]) the required number of events for a two-sided test (with significance level α and power β) may be calculated via

$$\hat{d} \approx \frac{(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2}{\ln^2(HR)\pi_1\pi_2}, \tag{16.28}$$

with z_{1-i} the ith quantile of the standard-normal distribution, *HR* the hazard ratio of $\alpha^*_{0,2}(u,t)$ between treatment groups and π_j the proportion of patients in treatment group j.

The value in question is the predicted calendar time t^* when for a given sample size N the required number of events is expected, i.e. $\mathbb{E}[D(t^*)|N,\mathfrak{H}_{t_c}] = \hat{d}$. The conditional expectation of events until time $t > t_c$, given the data up to current calendar time t_c is

$$\mathbb{E}[D(t)|N,\mathfrak{H}_{t_c}] = d(t_c) + \mathbb{E}[D_R(t)|N,\mathfrak{H}_{t_c}] + \mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}]\,,$$

with $d(t_c)$ the observed number of events until t_c, $\mathbb{E}[D_R(t)|N,\mathfrak{H}_{t_c}]$ as given in (16.23) and $\mathbb{E}[D_{NR}(t)|N,\mathfrak{H}_{t_c}]$ given in (16.25).

16.3.5 Sample-Size Calculations and Examples

On the other hand, for a fixed time point t (e.g. t the planned study duration) and $\hat{d}$ the required number of events until t, the required sample size can be calculated via

$$N = \frac{\hat{d}}{F_T(t)}\,. \tag{16.29}$$

This is based on the expected number of events corresponding to the overall number of patients randomized until t times the event probability at t.

16.3.5.1 Example 1

Suppose the treatment effect gives a hazard ratio of 0.75 for overall survival and of 0.67 for progression-free survival. The median OS time in the treatment and placebo group is 12 and 9 months, whereas the median PFS time in treatment and placebo group corresponds to 6 and 4 months, respectively. We assume an uniform accrual rate of 40 patients per month and a 1:1 randomization between treatment and placebo group. The significance level is 0.025 (one-sided) and the power is 80 %. The maximum expected observation time is 23 months.

By use of the exponential model for OS, 600 patients are needed for getting a power of 80 %. Based on this sample size, 380 events are expected at observation time $t = 23$ months. Using the three-state model, the above assumptions correspond to hazard ratios in the treatment and control group of $\lambda_{0,1} = 0.078$ and 0.116, $\lambda_{0,2} = 0.038$ and 0.057 as well as $\lambda_{1,2} = 0.105$ and 0.114, respectively. The expected number of events after 23 months is 390, based on a sample size of 600. The power in this scenario is 89 % due to the higher number of events. To get a power of 80 % when modelling overall survival via the three-state model, a sample size of only 480

patients is needed. On the other hand, 380 events will occur at observation time of 21.5 months, which saves 1.5 months of study duration.

16.3.5.2 Example 2

Our second example is based on data of a second line non-small cell lung cancer trial with 1,000 patients randomized in total. The last data monitoring committee DMC meeting has to occur after the 800th death event. A non-uniform accrual process is observed for this trial. Eighteen months after start of randomization, the event monitoring for this study is calculated by use of the exponential model as well as the three-state model. Based on the time from randomization, the time from randomization until the 800th death event is estimated for both models. We get stable estimates for both models after about 150 randomized patients and about 40 PFS events and 15 death events observed. The exponential model gives an estimation of 34 months and the three-state model an estimation of 29 months. Since the target number of 800 death events has not been observed so far, we run a simulation using the assumptions of the previous example to investigate until when the both models will estimate the time to the 380th death event and what the expected difference between the estimations is. The target of the 380th death event occurred at 21.5 months (please compare Fig. 16.9).

As seen in Fig. 16.9, 240 patients were randomized, 37 death events, and 80 PFS events were occurred after 6 months from randomization. The exponential model gave an estimation of 25 months, and the three-state model showed an estimation of 21.6 months (please compare Fig. 16.10). This is again a time difference of 4 months. Half of the required death events (190 OS events) were occurred after 14 months from randomization. Most of all patients were randomized and about half of them had a PFS event. Then the exponential model gave a more exact estimation of about 23.2 months, which is still 1.5 months more than the three-state model.

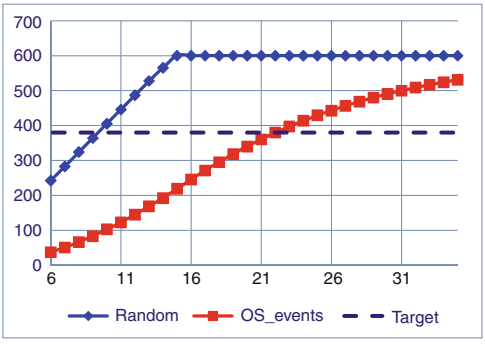


Fig. 16.9 The observed cumulative number of events over the time from randomization. (Results from the simulation example.) The number of patients randomised in (*diamonds - upper line*), the observed number of OS events (*squares - lower line*)

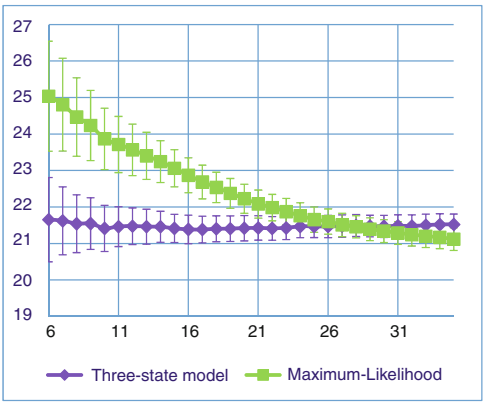


Fig. 16.10 Mean and standard error of the estimations of the time to 380th OS event from all simulations. The *squares* and the *diamonds* represent the 3-state model and the exponential model, respectively

16.3.5.3 Software Available

There is several software available. Multi-state models need specialised software, most of which are written in FORTRAN, R or SAS. The library survival available as part of S-plus and R statistical packages can be used to implement these methods. An R package `msm` was developed in 2002. In addition, an user-friendly R library, `tdc.msm`, was generated for the analysis of multi-state survival data. Technical description of this is provided in the independent article Meira-Machado et al. [25].

Appendix

Derivation of $F_T(.)$ for the disability model by assumption of exponentially distributed state times:

$$
\begin{aligned}
F_T(t) &= \mathbb{P}(T \leq t) \\
&= \int_0^t p_{0,0}(0,u)\alpha_{0,2}^*(u,t)\,\mathrm{d}u \\
&= \int_0^t \exp\left\{-\int_0^u (\alpha_{0,1}(v)+\alpha_{0,2}(v))\,\mathrm{d}v\right\}\left[\lambda_{0,2}+\lambda_{0,1}\left(1-\mathrm{e}^{-\lambda_{1,2}(t-u)}\right)\right]\mathrm{d}u \\
&= \int_0^t \exp\left\{-\int_0^u (\lambda_{0,1}+\lambda_{0,2})\,\mathrm{d}v\right\}\left[\lambda_{0,2}+\lambda_{0,1}\left(1-\mathrm{e}^{-\lambda_{1,2}(t-u)}\right)\right]\mathrm{d}u \\
&= \int_0^t \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})u}\left[\lambda_{0,2}+\lambda_{0,1}\left(1-\mathrm{e}^{-\lambda_{1,2}(t-u)}\right)\right]\mathrm{d}u
\end{aligned}
$$

$$
\begin{aligned}
&= \lambda_{0,2} \int_0^t \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})u}\,\mathrm{d}u + \lambda_{0,1} \int_0^t \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})u}\,\mathrm{d}u \\
&\quad - \lambda_{0,1}\mathrm{e}^{-\lambda_{1,2}t} \int_0^t \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})u}\,\mathrm{d}u \\
&= \frac{-\lambda_{0,2}}{\lambda_{0,1}+\lambda_{0,2}} \left(\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} - 1\right) + \frac{-\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}} \left(\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} - 1\right) \\
&\quad + \frac{\lambda_{0,1}\mathrm{e}^{-\lambda_{1,2}t}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}} \left(\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2})t} - 1\right) \\
&= \underbrace{\frac{\lambda_{0,2}}{\lambda_{0,1}+\lambda_{0,2}} + \frac{\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}}}_{=1} - \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} \underbrace{\left(\frac{\lambda_{0,2}+\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}}\right)}_{=1} \\
&\quad - \frac{\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\mathrm{e}^{-\lambda_{1,2}t} + \frac{\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} \\
&= 1 + \mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} \left(\frac{\lambda_{0,1}-\lambda_{0,1}-\lambda_{0,2}+\lambda_{1,2}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\right) - \frac{\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\mathrm{e}^{-\lambda_{1,2}t} \\
&= 1 + \frac{\lambda_{1,2}-\lambda_{0,2}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\mathrm{e}^{-(\lambda_{0,1}+\lambda_{0,2})t} - \frac{\lambda_{0,1}}{\lambda_{0,1}+\lambda_{0,2}-\lambda_{1,2}}\mathrm{e}^{-\lambda_{1,2}t}\,.
\end{aligned}
$$

A slightly different derivation of this formulae is given in Fleischer et al. [16].

References

1. Andersen, P. K., Keiding N.: Multi-state models for event history analysis. *Statistical Methods in Medical Research* **11**, 91–115 (2002)
2. Andersen, P. K.: Multistate Models in Survival Analysis: A Study of Nephropathy and Mortality in Diabetes. *Statistics in Medicine* **7**, 661–670 (1988)
3. Andersen, P. K., Borgan, Ø., Gill, R. D., Keiding, N.: *Statistical models based on counting processes*. Springer, New York (1993)
4. Andersen, P. K., Esbjerg S., Sorensen T. I. A.: Multistate models for bleeding episodes and mortality in liver cirrhosis. *Statistics in Medicine* **19**, 587–599 (2000)
5. Bagiella, E., Heitjan, D. F.: Predicting analysis times in randomized clinical trials. *Statistics in Medicine* **20**, 2055–2063 (2001)
6. Beyersmann, J., Schumacher, M., Allignol, A.: *Competing Risks and Multistate Models with R*. Springer (2012)
7. Birnbaumer, L., Bearer, C. F., Iyengar, R.: A Two-state Model of an Enzyme with an Allosteric Regulatory Site Capable of Metabolizing the Regulatory Ligand. *The Journal of Biological Chemistry* **255**(8), 3552–3557 (1980)
8. Box-Steffensmeier, J. M., De Boef, S.: Repeated events survival models: the conditional frailty model. *Statistics in Medicine* **25**(20), 3518–3533 (2006)
9. Chappell, R.: Competing risk analyses: how are they different and why should you care?. *Clinical Cancer Research* **18**(8), 2127–2129 (2012)

10. Cheson, B. D., Bennett, J. M., Kopecky, K. J., Büchner, T., Willman, C. L., Estey, E. H., Schiffer, C. A., Doehner, H., Tallman, M. S., Lister, A., Lo-Coco, F., Willemze, R., Biondi, A., Hiddemann, W., Larson, R. A., Löwenberg, B., Sanz, M. A., Head, D. R., Ohno, R., Bloomfield, C. D.: Revised Recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. *Journal of Clinical Oncology* **21**(24), 4642–4649 (2003)
11. Cheson, B. D., Pfistner, B., Juweid, M. E., Gascoyne, R. D., Specht, L., Horning, S. J., Coiffier, B., Fisher, R. I., Hagenbeek, A., Zucca, E., Rosen, S. T., Stroobants, S., Lister, T. A., Hoppe, R. T., Dreyling, M., Tobinai, K., Vose, J. M., Connors, J. M., Federico, M., Diehl, V.: Revised Response Criteria for Malignant Lymphoma. *Journal of Clinical Oncology* **25**(5), 579–586 (2007)
12. Dantan, E., Poly, P., Dartigues, J.-F., Jacqmin-Gadda, H.: Joint model with latent state for longitudinal and multistate data. *Biostatistics* **12**(4), 723–736 (2011)
13. Deshpande, J. V., Purohit, S. G.: Survival, hazard and competing risks. *Current Science* **80**(9), 1191–1202 (2001)
14. Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., Verweij, J.: New response evaluation criteria in solid tumorus: revised RECIST guidelines (version 1.1). *European Journal of Cancer* **45**, 228–247 (2009)
15. Fine, J. P., Jiang, H., Chappell, R.: On semi-competing risks. *Biometrika* **88**, 907–919 (2001)
16. Fleischer, F., Gaschler-Markefski, B., Bluhmki, E.: A statistical model for the dependence between progression-free survival and overall survival. *Statistics in Medicine* **28**, 2669–2686 (2009)
17. Heng, D. Y. C., Xie, W., Bjarnason, G. A., Vaishampayan, U., Tan, M.-H., Knox, J., Donskov, F., Wood, L., Kollmannsberger, C., Rini, B. I., Choueiri, T. K.: Progression-Free Survival as a Predictor of Overall Survival in Metastatic Renal Cell Carcinoma Treated With Contemporary Targeted Therapy. *Cancer* **117**(12), 2637–2642 (2011)
18. Hougaard, P.: Multi-state Models: A Review. *Lifetime Data Analysis* **5**, 239–264 (1999)
19. Joly, P., Gerds, T. A., Qvist, V., Commenges, D., Keiding, N.: Estimating survival of dental fillings on the basis of interval-censored data and multi-state models. *Statistics in Medicine* **31**, 1139–1149 (2012)
20. Kalbfleisch, J. D., Prentice, R. L.: *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, Inc. (1980)
21. Kvist, K., Andersen, P. K., Angst, J., Kessing, L. V.: Repeated events and total time on test. *Statistics in Medicine* **27**, 3817–3822 (2008)
22. Lan, L., Datta, S.: Non-parametric estimation of state occupation, entry and exit times with multistate current status. *Statistical Methods in Medical Research* **19**, 147–165 (2010)
23. Lawless, J. F.: *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc. (1982)
24. Nelson, W. B., Prentice, R. L.: *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*. ASA-SIAM Series on Statistics and Applied Probability, Philadelphia. (2003)
25. Meira-Machado, L., de Uña-Álvarez, J., Carso-Suárez, C., Andersen, P. K.: Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research* **18**(2), 195–222 (2009)
26. Peng, L., Fin, J. P.: Regression modeling of semicompeting risks data. *Biometrics* **63**, 96–108 (2006)
27. Porta, N., Calle, M. L., Malats, N., Gómez, G.: A dynamic model for the risk of bladder cancer progression. *Statistics in Medicine* **31**, 287–300 (2012)
28. Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., Breslow, N. E.: The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics* **34**, 541–554 (1978)

29. Putter, H., van Houwelingen, H. C.: Frailties in multi-state models: Are they identifiable? Do we need them?. *Statistical Methods in Medical Research*. online version **0**, 1–18 (2011)
30. Rodríguez-Girondo, M., de Uña-Álvarez, J.: Testing Markovianity in the three-state progressive model via future-past association. *Biometrical Journal* **54**(2), 163–180 (2012)
31. Sweeting, M. J., Farewell, V. T., De Angelis, D.: Multi-state Markov models for disease progression in the presence of informative examination times: An application to hepatitis C. *Statistics in Medicine* **29**, 1161–1174 (2010)
32. Schoenfeld, D. A.: Sample-Size Formula for the Proportional-Hazards Regression Model. *Biometrics* **39**(2), 499–503 (1983)
33. Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., Gwyther, S. G.: New Guidelines to Evaluate the Response to Treatment in Solid Tumors. *Journal of the National Cancer Institute* **92**(3), 205–216 (2000)
34. Tunes-da-Silva, G., Pedroso-de-Lima, A. C., Sen, P. K.: A semi-Markov multistate model for estimation of the mean quality-adjusted survival for non-progressive processes. *Lifetime Data Analysis* **15**, 216–240 (2009)
35. Uhry, Z., Hédelin, G., Colonna, M., Asselain, B., Arveux, P., Rogel, A., Exbrayat, C., Guldenfels, C., Courtial, I., Soler-Michel, P., Molinié, F., Eilstein, D., Duffy, S. W.: Multi-state Markov models in cancer screening evaluation: a brief review and case study. *Statistical Methods in Medical Research* **19**, 463–486 (2010)