Lifetime Data Anal (2013) 19:33–58
DOI 10.1007/s10985-012-9230-8

Applying competing risks regression models: an overview

Bernhard Haller · Georg Schmidt · Kurt Ulm

Received: 7 February 2012 / Accepted: 10 September 2012 / Published online: 26 September 2012

Abstract In many clinical research applications the time to occurrence of one event of interest, that may be obscured by another—so called competing—event, is investigated. Specific interventions can only have an effect on the endpoint they address or research questions might focus on risk factors for a certain outcome. Different approaches for the analysis of time-to-event data in the presence of competing risks were introduced in the last decades including some new methodologies, which are not yet frequently used in the analysis of competing risks data. Cause-specific hazard regression, subdistribution hazard regression, mixture models, vertical modelling and the analysis of time-to-event data based on pseudo-observations are described in this article and are applied to a dataset of a cohort study intended to establish risk stratification for cardiac death after myocardial infarction. Data analysts are encouraged to use the appropriate methods for their specific research questions by comparing different regression approaches in the competing risks setting regarding assumptions, methodology and interpretation of the results. Notes on application of the mentioned methods using the statistical software R are presented and extensions to the presented standard methods proposed in statistical literature are mentioned.

Keywords Competing risks · Cause-specific hazard · Subdistribution hazard · Mixture model · Vertical modelling · Pseudo-observation approach

B. Haller (✉) · K. Ulm
Institut für Medizinische Statistik und Epidemiologie der Technischen Universität München, Ismaninger Straße 22, 81675 Munich, Germany
e-mail: bernhard.haller@tum.de

G. Schmidt
1. Medizinische Klinik und Poliklinik der Technischen Universität München, Munich, Germany

1 Introduction

In the presence of competing risks, i.e. when two or more mutually exclusive events may possibly occur, a joint distribution for the time to different types of event cannot be estimated without making strong unverifiable assumptions (Tsiatis 1975). It is well known that the application of standard survival models and methods is not adequate in that situation. In the last three decades different methods for the analysis of failure time data in the presence of competing risks have been introduced. Prentice et al. (1978) proposed the use of standard survival models like Cox regression on the cause-specific hazard. In the cause-specific hazard model the effect of the investigated covariates on the competing event(s) is ignored, so there is no direct connection between the regression coefficients and the incidence of events. Larson and Dinse (1985) published an approach to express the joint distribution of event times and types of event as the product of the marginal distribution of event types and the conditional distribution of event times given the type of event. The authors proposed to use a logistic regression model to assess the influence of the covariates of interest on the type of event and piece-wise exponential regression models to asses their effect on failure time given the type of event. More flexible distributions for survival times were introduced in recent years as e.g. the generalized three-parameter gamma distribution (Lau et al. 2008). Fine and Gray (1999) introduced a regression approach focusing on the so called subdistribution hazard. In the Fine and Gray model the regression coefficients are monotonously linked to the cumulative incidence function and the occurrence of competing events has an influence on the coefficients. Modified standard survival models can be fit to estimate the influence of the investigated covariates on the subdistribution hazard. Andersen et al. (2003) introduced a method to estimate covariate effects on measures of interest in the presence of censored observations based on pseudo-values. The method was adjusted later for the competing risks setting by using the cumulative incidence function as measure of interest (Klein and Andersen 2005). The generalized estimating equation approach by Liang and Zeger (1986) is used to estimate the influence of covariates on the cumulative incidence function. Nicolaie et al. (2010) proposed another way to factorize the joint distribution of event times and types of event by expressing the joint distribution as product of the marginal distribution of the event times and the conditional distribution of the event types given the time of event. The so called vertical modelling approach gives an estimate for the relative hazard, showing the pattern of events in the course of time.

Due to different measures used for regression modelling and different approaches available for the analysis of time-to-event data with mutually exclusive types of event, analysis and interpretation of competing risks data is not straightforward and many sources of error are present in that situation. In this article an overview over current methodologies for the regression analysis of competing risks data is provided. In Sect. 2 terms and important measures used in the competing risks framework will be introduced. The above mentioned methods for the regression analysis of competing risks data will be described in more detail in Sect. 3 providing an overview over fundamental theory for all approaches as well as possible strategies for model applications using the statistical software R (R Development Core Team 2011). In Sect. 4 all methods will be applied to a cohort study including 2,341 patients who survived myocardial

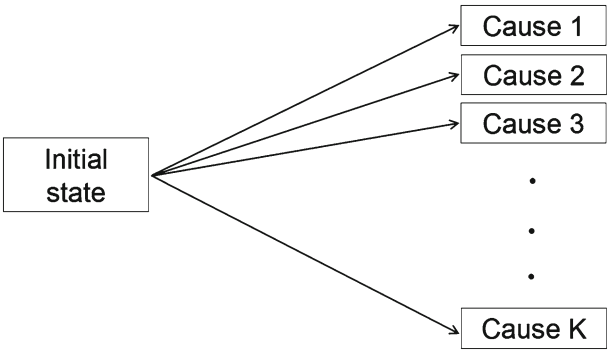

Fig. 1 Competing risks model: one initial state and K mutually exclusive types of failure

infarction (MI) (Bauer et al. 2009). Aim of the analysis is to establish a prespecified risk stratification for cardiac death with death from other causes as competing event. A discussion of the results and a comparison of the methods regarding interpretation and applicability will be provided in Sect. 5. A sketch of the R code used for data analysis is shown in the Appendix.

2 Competing risks

The problem of competing risks occurs, when the time from one certain starting point to an event of interest may not be observable, because of the incidence of another, so-called competing event (Fig. 1). Competing risks problems occur in different fields as medical statistics, engineering or social sciences with a rising awareness of the pitfalls present in that situation and a wider use of adequate models. For example in a cancer study investigating the time from treatment initiation to tumour-related death, deaths from other causes are competing events. In most competing risks applications one certain event of interest and one or more other possible events can be specified. In this situation the events not of major interest can be summarized to one category of competing events. It is also possible that more than one event is of special interest and the different types of event are treated equivalently. A large amount of literature dealing with competing risks can be found in textbooks (Beyersmann et al. 2012; Pintilie 2006; Crowder 2001; Kalbfleisch and Prentice 2002) or in introductory articles (e.g. Putter et al. 2007; Klein 2010; Lau et al. 2009; Bakoyannis and Touloumi 2011) mainly describing regression models based on cause-specific and subdistribution hazards.

It is well known that the analysis of times to a certain event k, conducted by estimation of the survivor function using standard Kaplan–Meier method (Kaplan and Meier 1958), is not adequate in the presence of competing risks (see e.g. Putter et al. 2007; Andersen and Keiding 2012). The cumulative incidence function $F_k(t)$, estimating the probability of failing from cause k before a given time t, is used instead to provide information for a certain population or to compare a discrete number of subgroups descriptively. The cumulative incidence function can be denoted as

$$F_k(t) = P(T \leq t, D = k), \tag{1}$$

with T and D being random variables representing the time to the first observed event and the type of event, respectively. In the absence of competing risks the cumulative incidence function equals $1 - S(t)$, where $S(t)$ is the survivor function, which can be derived by the Kaplan–Meier estimator. In the presence of competing risks, the cumulative incidence function in a population or in subgroups of interest can be estimated as

$$\hat{F}_k(t) = \sum_{i:t_i \leq t} \hat{\lambda}_k(t_i)\hat{S}(t_{i-1}). \tag{2}$$

Here $\hat{S}(t)$ is the estimator for the overall survivor function at time t including all types of event and t_i denotes the ith ordered event time. λ_k is defined as the cause-specific hazard rate which is given as

$$\lambda_k(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, D = k | T \geq t)}{\Delta t} \tag{3}$$

and can be estimated by

$$\hat{\lambda}_k(t_i) = \frac{d_{ki}}{n_i}, \tag{4}$$

where d_{ki} is the number of failures from type k at time t_i and n_i the risk set at time t_i, i.e. the number of patients who were not censored and have not failed from any cause up to time t_i.

3 Regression models in the presence of competing risks

3.1 Hazard-based regression models

3.1.1 Cause-specific hazard regression

As in common survival regression models, a measure of interest that can be used in the presence of right censored event times is needed. Prentice et al. (1978) proposed to estimate covariate effects on the cause-specific hazard rate. A semi-parametric Cox regression approach (Cox 1972) with a flexible unspecified cause-specific baseline hazard rate $\lambda_{k,0}(t)$ is proposed here

$$\lambda_k(t|\mathbf{X}) = \lambda_{k,0}(t) exp(\boldsymbol{\beta}_k^\top \mathbf{X}). \tag{5}$$

As in common Cox regression, the cause-specific hazard rates are assumed to be proportional translating to covariate effects that are constant over time. This assumption can be checked by graphical methods using Schoenfeld residuals (Schoenfeld 1982).

Modifications of the model as inclusion of time dependent covariates or time dependent effects can be considered to construct a valid Cox regression model. See for example Therneau and Grambsch (2000) for a detailed description of the Cox regression model including assumptions, methods for model checking and extensions of the classical model presented by Cox. Instead of the semi-parametric Cox regression model parametric models assuming e.g. exponentially, Weibull or gamma distributed event times can be performed (see e.g. Klein and Moeschberger 2003).

Analysis of competing risks data based on the cause-specific hazard using Cox regression can be conducted in statistical standard software packages by implementing classical Cox regression treating failures from the cause of interest as events and failures from other causes as censored observations. Results of the regression model must be interpreted carefully, since the estimated regression coefficients give the effect of the covariates on the instantaneous probability of failing from cause k given a subject experienced no event until time t. Since competing events are not considered, the effects of the covariates on the cause-specific hazard rate cannot be translated directly to an effect on the cumulative incidence function. That means a higher cause-specific hazard in group A compared to group B does not necessarily lead to a higher incidence of events of interest in group A than in group B. A nice illustration of that fact can be found in Putter et al. (2007) or in Allignol et al. (2011).

Beyersmann et al. (2009) state that cause-specific hazards "completely determine the competing risks process", so cumulative incidence functions can be estimated from separate cause-specific hazard regression models for all types of event. The cumulative incidence function for the kth out of K events is

$$F_k(t|\mathbf{X}) = \int_0^t \lambda_k(s|\mathbf{X})\, exp\Big(-\sum_{l=1}^{K} \Lambda_l(s|\mathbf{X})\Big)\mathrm{d}s, \tag{6}$$

where $\Lambda_k(t|\mathbf{X})$ denotes the cumulative cause-specific hazard rate for event k at time t for a given matrix of covariates $\mathbf{X}$, which is defined as $\Lambda_k(t|\mathbf{X}) = \int_0^t \lambda_k(s|\mathbf{X})\mathrm{d}s$ with $\lambda_k(t|\mathbf{X})$ as given by Eq. 5. The set of covariates considered may vary for different event types. Many extensions of the cause-specific hazard regression have been published, as flexible modelling of cause-specific-hazard rates (Belot et al. 2010; Scheike and Zhang 2008), simultaneous estimation of the cause-specific hazards for all event types including tests on equality of baseline hazards and covariate effects on different types of event (Lunn and McNeil 1995) or testing and estimation of time-varying effects (Sun et al. 2008).

3.1.2 Subdistribution hazard regression—the Fine and Gray model

To overcome some of the problems occurring with cause-specific hazard regression as described above, Fine and Gray (1999) developed a regression model that directly links the regression coefficients with the cumulative incidence function. In the Fine and Gray model the association between the so called subdistribution hazard introduced earlier by Gray (1988) and covariates of interest is assessed. The subdistribution hazard for event k is defined as the probability for a subject to fail from cause k in an

infinitesimal small time interval Δt, given the subject experienced no event until time t or experienced an event other than k before time t

$$\lambda_k^*(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, D = k | T \geq t \cup \{T < t, D \neq k\})}{\Delta t}. \tag{7}$$

Individuals failing before time t, but not from the cause of interest, remain in the risk set for all future failure times. If events are documented in discrete time and no censoring is present, the subdistribution hazard at time t_i can be estimated as

$$\hat{\lambda}_k^*(t_i) = \frac{d_{ki}}{n_i^*}, \tag{8}$$

where d_{ki} denotes the number of failures of type k at time t_i and n_i^* the modified risk set including all subjects who did not experience any event until time t_i and all subjects that failed before t_i from a cause other than k. An illustration of differences in risk sets for the cause-specific and the subdistribution hazard can be found in Lau et al. (2009). The estimated subdistribution hazard $\hat{\lambda}_k^*(t)$ equals the estimated cause-specific hazard $\hat{\lambda}_k(t)$ until the first competing event is observed and is smaller than the cause-specific hazard for all following time points, since the modified risk set n_i^* used for the estimation of the subdistribution hazard is larger than the cause-specific hazard risk set n_i after incidence of the first competing event. In the presence of non-administrative right-censoring individuals in the adapted risk set n_i^* are weighted using the inverse probability of censoring weighting (IPCW) approach introduced by Robins and Rotnitzky (1992). Fine and Gray used that approach to obtain a weighted score function leading to consistent regression coefficients.

As in the cause-specific hazard regression, different regression models for the subdistribution hazard can be used. Due to its wide acceptance and awareness, a Cox-type regression model is presented again

$$\lambda_k^*(t|\mathbf{X}) = \lambda_{k,0}^*(t) exp(\boldsymbol{\beta}_k^{*\top} \mathbf{X}), \tag{9}$$

where $\lambda_{k,0}^*(t)$ denotes the subdistribution baseline hazard function (i.e. the subdistribution hazard for a (fictitious) individual with all covariates set to zero). Subdistribution hazard rates are assumed to be proportional for the included covariates. The subdistribution hazard is linked directly to the cumulative incidence function in a way known from classical survival analysis with one possible endpoint

$$F_k(t|\mathbf{X}) = 1 - exp(-\Lambda_k^*(t|\mathbf{X})) = 1 - exp\left(-\int_0^t \lambda_k^*(s|\mathbf{X}) \mathrm{d}s\right). \tag{10}$$

Hence the cumulative incidence function for the event of interest can be estimated directly from the regression coefficients obtained by a Fine and Gray model without explicit consideration of the covariate effects on competing events. Care has to be

taken again in the interpretation of the results, since the regression coefficients aim on the incidence of events, not on the rate (see Sect. 3.1.3).

In the statistical software R the library *cmprsk* (Gray 2010) is provided allowing to fit proportional subdistribution hazard models. Time-fixed covariates and covariates interacting with functions of time can be included and the assumption of proportional sub-distribution hazards can be checked graphically using Schoenfeld-type residuals.

3.1.3 Comparison of cause-specific and subdistribution hazard regression

Regression models for the cause-specific hazard and the subdistribution hazard are used most often for the analysis of competing risks data in medical research. Since these approaches focus on different measures they might lead to substantially different results and interpretations. Beyersmann et al. (2007) compare competing risk analyses using cause-specific and subdistribution hazards in a real data example. They present the mathematical relationship between cause-specific and subdistribution hazards (see also Beyersmann and Schumacher 2007). In the presence of two possible types of failure, the relationship can be derived from Eqs. 6 and 10 to be

$$\lambda_1(t|\mathbf{X}) = \left(1 + \frac{F_2(t|\mathbf{X})}{S(t|\mathbf{X})}\right) \cdot \lambda_1^*(t|\mathbf{X}), \tag{11}$$

where $S(t|\mathbf{X})$ denotes the probability of being free of any event up to time t given $\mathbf{X}$.

Latouche et al. (2007) investigated the results obtained from proportional subdistribution hazard regression models given proportionality assumptions actually hold for a cause-specific hazards model and concluded that effect estimates are different in both models with the amount of difference depending on the cause-specific covariate effects on the event of interest as well as the effects on the competing event(s). Differences between cause-specific and subdistribution hazard regression were also presented and discussed in detail by Dignam and Kocherginsky (2008) based on different simulation scenarios. It is displayed that the results of a two group comparison regarding cause-specific or subdistribution hazard ratios might differ substantially in the presence of competing events.

E.g. in a scenario with one covariate having a high effect on the cause-specific hazard of a competing event, but the risk of failing from the event of interest is independent of that covariate, a Fine and Gray regression model will reveal an effect of that covariate on the subdistribution hazard, because the observed incidence of events of interest is diminished by a high number of competing events reducing the number of patients at risk. In extreme scenarios both methods might even give different signs for regressions coefficients indicating a higher cause-specific hazard, but a lower subdistribution hazard for one group compared to the other. Therefore investigators should be aware of differences between cause-specific hazard and subdistribution hazard regression to avoid misuse of the methods and misinterpretation of obtained results. Furthermore, model assumptions (proportionality of hazard rates) relate to different measures. Grambauer et al. (2010) describe that proportional cause-specific hazard rates mostly translate into non-proportional subdistribution hazard rates and vice versa, but they state that the estimated coefficient in the subdistribution hazard

model gives a consistent estimate of the so called least false parameter $\tilde{\beta}$, that can be interpreted as time-averaged subdistribution hazard ratio.

3.2 Mixture models

As summarized in the introduction, Larson and Dinse (1985) proposed to consider the joint distribution of event types and times as product of the marginal distribution of types of event $P(D)$ and the conditional distribution of the times to the accordant event given the type of event $P(T|D)$

$$P(D,T) = P(D)P(T|D), \tag{12}$$

with D denoting a random variable for the type of event and T a non-negative random variable for the time of failure. In order to estimate distribution parameters or regression coefficients, respectively, and to draw inference, the likelihood of the mixture model has to be formulated. As presented by Lau et al. (2008), in a mixture model with two possible types of failure the likelihood contributed by individual i can generally be written as

$$L_i = \big(\pi_i f_1(t_i)\big)^{I(D_i=1)} \times \big((1-\pi_i) f_2(t_i)\big)^{I(D_i=2)} \times \big(\pi_i S_1(t_i) + (1-\pi_i) S_2(t_i)\big)^{I(D_i=0)}, \tag{13}$$

where π_i is the probability of failing from cause 1 for individual i, $f_k(t)$ the value of the density function of the survival time distribution, $S_k(t)$ the value of the survivor function for failure type k at time t and $I(D_i = k)$ indicates the type of failure by giving the value 1 if $D_i = k$ and 0 else, where $D_i = 0$ denotes a censored observation.

A multinomial logistic regression model can be used to assess the influence of covariates on the probability of failing from a certain event k (see e.g. Fahrmeir and Tutz 2001)

$$P(D = k|\mathbf{X}) = \frac{exp(\mu_k + \boldsymbol{\pi}_k^\top \mathbf{X})}{\sum_{l=1}^{K} exp(\mu_l + \boldsymbol{\pi}_l^\top \mathbf{X})}. \tag{14}$$

In the case of two possible events or the case of one event of interest and summarizing the competing events to one category, the multinomial logistic regression will reduce to a binary logistic regression model, so the constant term $\boldsymbol{\mu}$ becomes a scalar and the probability for an event of type 1 for subject i given his vector of covariates $\mathbf{x_i}$ can be expressed as

$$P(D_i = 1|\mathbf{x_i}) = 1/(1 + exp(-(\mu + \boldsymbol{\pi}^\top \mathbf{x_i}))). \tag{15}$$

So a valid model can provide an estimate for the probability that an individual will fail from the event of interest based on his covariate information.

An adequate assumption for the conditional distribution of the survival times for a given type of event has to be made. Different survival time distributions were proposed in the literature like piecewise exponentially distributed survival times (Larson and Dinse 1985), described in detail by Friedman (1982) for classical survival analysis, or a three-parameter generalized gamma distribution proposed by Lau et al. (2008), which was investigated by Cox et al. (2007) for classical survival analysis. Details on survival time distributions can be found e.g. in Kalbfleisch and Prentice (2002) or in Klein and Moeschberger (2003). Ng and McLachlan (2003) and Escarela and Bowater (2008) proposed semi-parametric mixture models assuming proportional hazards for a given type of failure and fitting Cox proportional hazards models for the conditional hazard functions, so no distribution assumption has to made for the event times given the type of event.

The survivor function for a given type of failure and a given set of covariates can be denoted as

$$S_k(t|\mathbf{X}) = P(T > t|\mathbf{X}, D = k) = exp\left(-\int_0^t h_k(s)exp\left(\boldsymbol{\beta}_k^\top \mathbf{X}\right) \mathrm{d}s\right). \quad (16)$$

Here $h_k(s)$ is the null hazard function for a (fictitious) individual with all covariates set to zero. Using a piece-wise exponential model with $m = \{1, \ldots, M\}$ pieces, like in the original article by Larson and Dinse (1985), the null hazard function has the form

$$h_k(t) = exp(\alpha_{km}) \quad \text{for all } t \text{ in interval } m, \quad (17)$$

where α_{km} represents the log–null hazard for an event of type k in the mth interval. The covariate sets that are assumed to effect the probability of failing from the event of interest or the time to an event given the type of failure do not have to be identical (see e.g. Lau et al. 2011), but should be chosen based on knowledge and biological plausibility.

Formulation of the likelihood for an adequate mixture model (see Eq. 13) may be complicated, which is one of the major drawbacks of that approach. Another reason for the rare usage of mixture models may be the lack of standard software available for parameter estimation or inference. Maximum-likelihood estimates of the regression coefficients can be assessed using an expectation-maximization (EM) algorithm (McLachlan and Krishnan 1997) or by using non-linear regression methods as *nlm* in R. Standard errors and confidence intervals for the regression coefficients can be obtained by bootstrapping methods (e.g. Efron and Tibshirani 1994) or via the Hessian matrix provided by non linear regression algorithms.

Lau et al. (2011) presented a method to estimate time-dependent cause-specific and subdistribution hazard ratios and corresponding summaries over time from a mixture model that lead to similar results as the semi-parametric methods presented in Sects. 3.1.1 and 3.1.2.

3.3 Vertical modelling

In the vertical modelling approach proposed by Nicolaie et al. (2010) a different factorisation of the joint distribution of types of event and event times is used. The joint distribution is expressed as the product of the marginal distribution of event times for all types of events and the conditional distribution for types of event given the event time

$$P(T, D) = P(T)P(D|T). \tag{18}$$

For each observed event time relative hazards for the types of events can be estimated, revealing the pattern of occurrence of different events given the event times. Using this approach, it can e.g. be assessed if certain events tend to happen soon after an intervention while other events occur rather in the long term follow-up. The marginal distribution of the survival times can be estimated using the Kaplan–Meier method or, if the effects of covariates on the marginal event time distribution are of interest, by some regression model as proportional hazards or parametric survival models. The probability for occurrence of a certain event k, given some event $1, \ldots, K$ was observed at time t, is called the relative hazard π_k,

$$\pi_k(t) = P(D = k|T = t). \tag{19}$$

The relative hazard for event k given an observed event time t_i can be estimated as

$$\hat{\pi}_k(t_i) = \frac{\frac{d_{ik}}{n_i}}{\frac{d_i}{n_i}} = \frac{d_{ik}}{d_i}. \tag{20}$$

Here d_i describes the number of observed events at time t_i, d_{ik} the number of events of type k at time t_i and n_i the number of subjects at risk at time t_i. Since in most applications events are documented in continuous time, the relative hazards will give a series of zeros and ones for each type of event. So Nicolaie et al. proposed to either use some pre-specified time intervals summarizing multiple events for the estimation of relative hazards or to use multinomial regression models with spline functions like B-Splines (see e.g. Hastie 1997) for flexible estimation of time and covariate effects on the relative hazards. If events are summarized in discrete time intervals, the number and length of the intervals considered can play an important role. Choice of many small intervals allows flexible description of the relative hazards, but leads to higher uncertainty in the estimates, as less events are observed in each interval. Due to the structure of the model only subjects with observed events can be used for estimation of relative hazards, whereas all observations are used for estimation of the marginal distribution of survival times and corresponding covariate effects. As the relative hazard $\pi_k(t)$ denotes the probability for an event of type k at time t, given any event happened at the corresponding timepoint or in the time interval, respectively, the relative hazards

$\pi_1(t), \ldots, \pi_K(t)$ sum up to one for any t. A possible model for estimation of time and covariate effects on the relative hazard is

$$\pi_k(t|\mathbf{X}) = \frac{exp\left(\boldsymbol{\gamma}_k^\top \mathbf{B}(t) + \boldsymbol{\beta}_k^\top \mathbf{X}\right)}{\sum_{l=1}^{K} exp\left(\boldsymbol{\gamma}_l^\top \mathbf{B}(t) + \boldsymbol{\beta}_l^\top \mathbf{X}\right)}. \tag{21}$$

$\mathbf{B}(t)$ denotes the vector of the spline basis functions, $\boldsymbol{\gamma}_k$ the vector of regression weights of the B-spline functions for event k, $\mathbf{X}$ the matrix of covariates and $\boldsymbol{\beta}_k$ the vector of covariate regression coefficients for the kth type of event. Additionally, an interaction effect between covariates and B-spline functions can be estimated, if sufficient data are available. Graphical methods, e.g. drawing the relative hazards for all possible event types versus time, seem to be most adequate for presentation and interpretation of the results. A summary of the results should always include the relative hazards as well as the distribution of event times, so that relative hazards for each type of event and the pattern of overall events can be interpreted in consideration. Presentation of the distribution of event times by e.g. Kaplan–Meier curve(s) might avoid mis- or overinterpretation of relative hazards in time-intervals with a low number of observed events.

3.4 Competing risks regression based on pseudo-observations

Andersen et al. (2003) introduced a method for the estimation of covariate effects on state probabilities in multi-state models using pseudo-observations. Since classical survival models and competing risks models can be interpreted as special cases of multi-state models, this approach can be adjusted for the competing risks setting as demonstrated by Klein and Andersen (2005). Generally, the pseudo-observation approach can be considered to estimate the effects of covariates on any function of event times $f(T)$, if an unbiased estimator $\hat{\theta}$ exists for

$$\theta = E\big(f(T)\big). \tag{22}$$

A summary of different methods for survival analysis based on pseudo-observations is presented by Andersen and Perme (2010). Main idea of the approach is to replace censored observations, which are usually present in event-time analysis, by some useful measure, so that standard methods can be used for data analysis. The estimated pseudo-observations $\hat{\theta}_{ih}$, which are assessed via leave-one-out estimates (see e.g. Miller 1974) for some measure of interest at a predefined series of timepoints $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_H)$ can be used for that purpose

$$\hat{\theta}_{ih} = n\,\hat{\theta}(\tau_h) - (n-1)\,\hat{\theta}^{(i)}(\tau_h). \tag{23}$$

Here $\hat{\theta}(\tau_h)$ is the estimated measure of interest at time τ_h using all observations and $\hat{\theta}^{(i)}(\tau_h)$ indicates the estimated measure of interest derived from all but the ith observation. So a n×H-matrix of pseudo-observations is obtained. For regression purposes

these pseudo-observations $\hat{\theta}_{ih}$ can be used as dependent variable (Klein and Andersen 2005)

$$g(\theta_{ih}|\mathbf{x_{ih}}) = \boldsymbol{\beta}^{\top}\mathbf{x_{ih}}, \tag{24}$$

where g is a link function and $\mathbf{x_{ih}}$ is the vector of covariates of subject i at time τ_h. In order to obtain valid standard errors the generalized estimation equation approach (GEE, Liang and Zeger 1986) can be applied to account for multiple observations per subject.

In the competing risks setting the relevant measure $f(T)$ is the cumulative incidence function for event k. Pseudo-observations for each individual can be generated following Eq. 23 inserting the estimate for the cumulative incidence function at time τ_h using all observations for $\hat{\theta}(\tau_h)$ and the cumulative incidence function based on all but the ith observation for $\hat{\theta}^{(i)}(\tau_h)$. When a complementary log–log link is used in Eq. 24, the regression coefficients can be interpreted as logarithm of the subdistribution hazard ratio, if all covariates are time-independent (Klein and Andersen 2005). The analysis can be performed using the R function *geese* from the R library *geepack* (Højsgaard et al. 2005), that allows different link functions between response and linear predictor. SAS and R functions for the computation of pseudo-values for time-to-event data are provided and discussed by Klein et al. (2008).

4 Example: application to cardiac data

4.1 Description of the data

The five presented methods were all applied to a dataset collected at a cohort study in the Klinikum rechts der Isar and in the German Heart Centre Munich, both located in Munich, Germany, between January 1995 and March 2005. A total of 2,343 patients who survived an acute MI at an age of 75 years or younger were included in the study. The analysed data are presented in Bauer et al. (2006, 2009) and Barthel et al. (2003) including medical details and a more substantial description of the study cohort. Two of the patients were excluded from the analysis due to missing values, so the results presented are based on the evaluation of 2,341 individuals. Patients were planned to be followed for at least 5 years. Time from the MI to death and type of death (cardiac or non-cardiac reason) were documented. At inclusion time patients were prospectively categorized to risk groups. Patients with a left ventricular ejection fraction (LVEF) of less than 30 % and patients with an LVEF of more than 30 %, but severe autonomic failure (SAF), were specified to be of high risk for cardiac death ($n = 236$), patients with an LVEF of more than 30 % and no SAF to be of low risk ($n = 2{,}105$).

1,140 patients were followed for 5 years, so the median follow-up time was 5 years assessed by inverse Kaplan–Meier method (Schemper and Smith 1996). About 75 % of the patients were followed for at least 3 years. Patients lost to follow-up or retreating from the trial were considered as censored observations. During follow-up 181 of the 2,341 patients died, 104 of them from cardiac reasons (55 sudden cardiac deaths), 77 patients died from other causes or types of death were not specified ($n = 14$). For ease

Table 1 Estimated cumulative incidences ($\hat{F}$) for cardiac and non-cardiac death 5 years after MI with 95 % confidence intervals

	Cardiac death		Non-cardiac death	
	$\hat{F}_{card.}$ (5 years) (%)	95 % CI	$\hat{F}_{non\text{-}card.}$ (5 years) (%)	95 % CI
Overall	5.1	4.2–6.1 %	4.1	3.1–5.0 %
Low risk	2.7	1.9–3.4 %	3.3	2.4–4.2 %
High risk	27.5	21.1–33.9 %	10.9	6.3–15.7 %

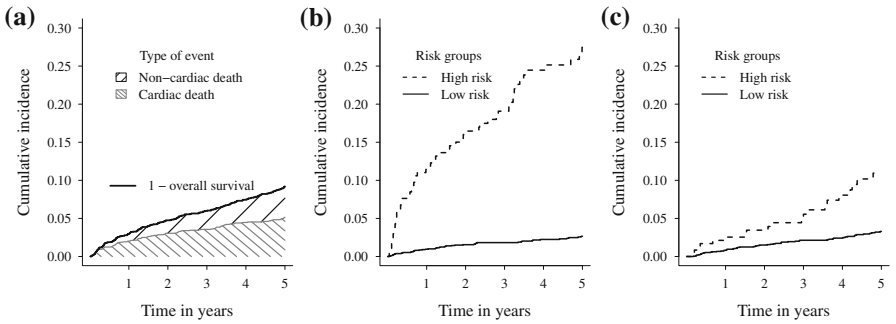

Fig. 2 Cumulative incidence functions for cardiac or non-cardiac death for the whole study population with cumulative incidences for both types of event summing up to 1 – overall survival (**a**), comparison of high (*dashed*) and low risk group (*solid line*) regarding incidences of cardiac (**b**) and non-cardiac death (**c**)

of analysis and interpretation these 77 patients were defined to have died from non-cardiac reasons. The estimated probability of dying in the first 5 years after MI was 9.2 % (95 % confidence interval (95 % CI) 7.9 to 10.5 %). Estimates of the cumulative incidence functions 5 years after MI with 95 % confidence intervals for both types of death are presented in Table 1 for the whole study population and stratified for the risk groups. In Fig. 2 non-parametric estimates of the cumulative incidence functions for the two competing types of event are presented.

Aim of the analysis is to evaluate the risk stratification. Therefore, the effects of risk group, age (dichotomized at 65 years) and diabetes on cardiac mortality were assessed using the methods described in Sect. 3. All analyses were performed using the statistical software R (R Development Core Team 2011) and its libraries *survival* (Therneau 2011), *cmprsk* (Gray 2010), *geepack* (Højsgaard et al. 2005) and *splines* (part of R).

4.2 Cause-specific hazard regression

The effect of the risk group on the cause-specific hazard adjusted for age and diabetes was analysed. Therefore, since an investigation of Schoenfeld residuals revealed no evidence against the assumption of proportionality for both types of event (not shown), a Cox regression model was fit for each of the two types of failure to estimate the effect of the three covariates on the cause-specific hazards using the R function *coxph* from the library *survival*. In order to describe the whole competing risks process and to

Table 2 Results of the cause-specific hazard regression models for both types of failure

	$\hat{\beta}$	$\exp(\hat{\beta})$	Std. error	p value
Cardiac death				
Risk group	2.36	10.53	0.20	<0.001
Diabetes	0.72	2.06	0.21	0.001
Age ≥65	0.48	1.60	0.20	0.016
Non-cardiac death				
Risk group	1.06	2.89	0.26	<0.001
Diabetes	0.70	2.01	0.25	0.005
Age ≥65	1.28	3.69	0.24	<0.001

estimate the cumulative incidence function for a given set of covariates, the influence of the covariates on all types of failures have to be assessed (see Eq. 6). For each type of failure patients experiencing the competing event were considered as censored observations. Risk group, diabetes and age had a significant effect on the cause-specific hazards for both types of event (results are shown in Table 2). A cause-specific hazard ratio between the high risk and the low risk group for cardiac death (HR_c^{cs}) of 10.53 (95 % CI 7.10 to 15.64) was observed. This indicates an about ten times higher risk of dying from a cardiac event for patients with an LVEF ≤ 30 % or with SAF compared to patients with an LVEF > 30 % and no SAF. The analysis for non-cardiac deaths revealed an increased risk for patients from the high risk group, too, but the effect was much lower ($HR_{nc}^{cs} = 2.89$, 95 % CI 1.73 to 4.85). Age had a greater influence on the cause-specific hazard for non-cardiac death with a cause-specific hazard ratio of 3.69 (95 % CI 2.29 to 5.91) compared to 1.60 (95 % CI 1.09 to 2.40) for cardiac events. For both types of event the cause-specific hazard for patients with diabetes was about twice as high as for patients without diabetes. Cumulative incidence functions were predicted from the Cox regression models following Eq. 6 for both risk groups using the mean of diabetes, i.e. the proportion of patients with diabetes (17.6 %), and the mean of the indicator variable for age, i.e. the proportion of patients being at least 65 years of age (30.2 %). The baseline hazard required for the estimation of $\Lambda_k(t)$ was estimated using the method by Breslow (see e.g. Kalbfleisch and Prentice 2002). The predicted cumulative incidence curves are shown in Fig. 5a.

4.3 Subdistribution hazard regression—Fine and Gray model

A proportional subdistribution hazards model as described in Eq. 9 was fit to assess the influence of risk group, diabetes and age on the subdistribution hazard for both types of event. The analysis was performed using the function *crr* in the R library *cmprsk*. Since the assumption of proportional hazards cannot be valid for cause-specific and subdistribution hazards (Grambauer et al. 2010) and due to conceptual problems appearing when proportional subdistribution hazard models are fitted for both types of events shown by Beyersmann et al. (2012), results from the subdistribution hazards models have to be interpreted as time-averaged effects.

Table 3 Results of the subdistribution hazard regression (Fine and Gray) models for both types of failure

	$\hat{\beta}$	$\exp(\hat{\beta})$	Std. error	p-value
Cardiac death				
Risk group	2.32	10.21	0.20	<0.001
Diabetes	0.68	1.98	0.21	0.001
Age ≥65	0.47	1.60	0.20	0.017
Non-cardiac death				
Risk group	0.84	2.31	0.28	0.002
Diabetes	0.62	1.85	0.24	0.011
Age ≥65	1.28	3.58	0.25	<0.001

Results of the two regression models investigating the influence of the covariates on the subdistribution hazard are shown in Table 3 for both types of failure. Effects on the subdistribution hazard can be directly translated to effects on the cumulative incidence function. For cardiac mortality a subdistribution hazard ratio (HR_c^{sd}) comparing the high risk to the low risk group of 10.21 (95 % CI 6.91 to 15.08) was found, indicating a much higher incidence of cardiac events for patients categorized to be of high risk. The effect of the risk group allocation was weaker for non-cardiac death (HR_{nc}^{sd} = 2.31, 95 % CI 1.39 to 3.97). Effects of diabetes were similar for both types of failure with a higher subdistribution hazard for patients suffering from diabetes, whereas age had a higher effect on the subdistribution hazard of non-cardiac deaths. Cumulative incidence functions for cardiac death comparing high and low risk group at mean of age and diabetes are shown in Fig. 5b.

4.4 Mixture model

For the analysis of the data using a mixture model the semi-parametric approach proposed by Ng and McLachlan (2003) was applied, so no assumptions for the distributions of failure times for a given type of event had to be made, but the hazard rates were assumed to be proportional. Parameter estimates were obtained via an expectation-conditional maximization (ECM) algorithm. In the ECM algorithm parameters are estimated iteratively by altering the expectation of the failure type for censored observations given the observed data and the current parameter estimates (E-step) and the maximization of the log-likelihood given the observed data and the expected failure-type probabilities for censored observations (M-step). These steps are repeated until some pre-specified convergence criterion is fulfilled (e.g. absolute or relative change of the parameter estimates). The value of the likelihood is increased by each iteration step. Different starting values were used to avoid finding a local maximum, but all computations led to the same final results. Five hundred bootstrap samples were generated to estimate standard errors of the regression coefficients. As described by Ng and McLachlan subsamples were randomly drawn with replacement from patients experiencing cardiac death, from patients failing from non-cardiac death and from censored individuals according to the numbers observed in the original dataset. Results of the analysis are presented in Table 4. The coefficients of the logistic regression,

Table 4 Regression coefficients obtained from the mixture model analysis with 95 % confidence intervals based on 500 bootstrap samples

	Event types		Event times			
	Cardiac		Cardiac		Non-cardiac	
	$\hat{\pi}$	95 % CI (bs)	$\hat{\beta}$	95 % CI (bs)	$\hat{\beta}$	95 % CI (bs)
Constant	−2.30	−3.92 to 1.65	–	–	–	–
Risk group	2.22	−1.44 to 3.97	0.88	−0.92 to 3.19	1.76	−0.21 to 2.76
Diabetes	−0.43	−2.00 to 1.82	1.17	−0.44 to 2.02	0.52	−0.49 to 2.07
Age ≥65	0.96	−1.45 to 2.66	−0.25	−1.60 to 1.02	1.54	0.71 to 2.54

modelling the expected type of failure, indicate that high risk patients were more likely to die from cardiac events (OR = exp(2.22) = 9.21, 95 % bootstrap CI 0.24 to 52.98). For a low risk patient aged at least 65 years and having no diabetes, following Eq. 15, a probability of dying from a cardiac event of 20.7 % was estimated, for a person of the same age, who is also free of diabetes, but who was identified to be of high risk, the predicted probability increases to 70.7 %. For both types of failure, patients from the high risk group tended to survive for a shorter time period, as their estimated risk for failing from the given type of event is increased (hazard ratios of exp(0.88) = 2.41 and exp(1.76) = 5.81).

4.5 Vertical modelling

In the vertical modelling approach patterns for the occurrence of events in the course of time can be investigated. Marginal survivor functions for both risk groups adjusted for age and diabetes were estimated by a Cox regression model and are presented using the mean of diabetes and the mean of age derived from the whole study population (Fig. 3a). In order to estimate relative hazards of the event types in the course of time, a logistic regression model was fitted considering all uncensored subjects. Time, risk group, diabetes and age were included as covariates, an indicator variable giving one, if the observed event was death from cardiac reasons, and zero for a death from non-cardiac reasons was used as dependent variable. Flexible cubic B-spline functions were used to estimate the effect of time smoothly. As proposed by Nicolaie et al. (2010) interaction terms between risk group and the smooth functions of time were considered to allow for different patterns in both groups. Calculations were conducted using the function *glm* with flexible B-splines incorporated in the *splines* library. Coefficients for the main effects obtained from the logistic regression model estimating the probability of occurrence of a cardiac event, given any event was observed, are presented in Table 5. As interpretation of regression coefficients is difficult due to the use of B-spline functions and interaction terms, estimated relative hazards are displayed for both types of event in Fig. 3b, c. The estimated probability for death of any type, adjusted for age and diabetes, is higher in the high risk group compared to the low risk group (Fig. 3a). For a high risk patient the probability for dying from a cardiac event, given the patient dies at a certain time t, is substantially higher than the probability

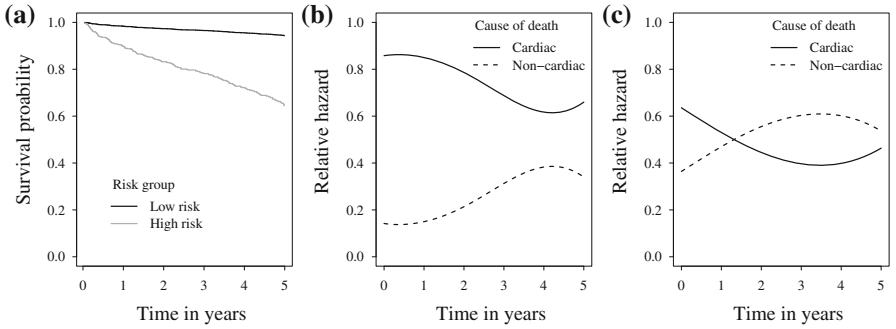

Fig. 3 Results of the vertical modelling approach: survivor functions for both risk groups adjusted for age and diabetes (**a**), relative hazards for the high risk group (**b**) and relative hazards for the low risk group (**c**)

Table 5 Results from the vertical modelling approach

	$\hat{\beta}$	Std. error	p value
Constant	0.75	0.61	0.218
Risk group	1.27	0.91	0.165
Diabetes	−0.08	0.36	0.825
Age ≥65	−0.65	0.34	0.053
B-Spline comp.	Not shown		

Regression coefficients for B-Spline components of time, and the interaction terms between the B-Spline components and risk group are not shown, as these cannot be interpreted properly

for a non-cardiac event for all timepoints (Fig. 3b), whereas both types of events seem to appear with a similar probability in the low risk group (Fig. 3c). For both types of event the probability for a cardiac event, given any event occurred at a certain time, seems to decrease slightly over time, indicating a higher relative hazard for cardiac events in the first year after MI.

4.6 Analysis based on pseudo-observations

In order to analyse the data using the approach proposed by Klein and Andersen (2005), pseudo-values for the cumulative incidence function for death from cardiac reasons were estimated. First, the cumulative incidence function for cardiac death was estimated for 21 different points in time (three month intervals equally spaced from baseline to 5 years of follow-up) for the whole data set. The procedure was repeated for all prespecified timepoints leaving out each subject once. From these estimates $2{,}341 \times 21$ pseudo-observations were calculated following Eq. 23. Examples for pseudo-observations obtained from the observed data are displayed in Fig. 4. The patient displayed in the left picture (a) died from a cardiac reason after 2.57 years, the patient in Fig. 4b was censored after 3.79 years and the patient in the right picture (c) died from a non-cardiac reason after 2.07 years. Due to the large amount of patients followed for 5 years without any critical event, the pseudo-value approach does not affect the weights of censored individuals heavily, but patients experiencing an event

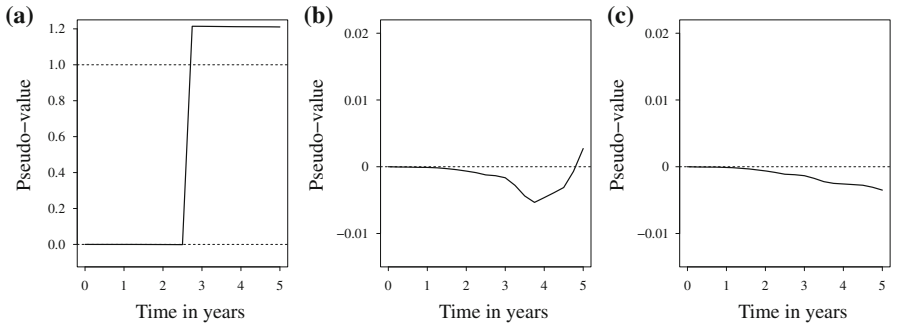

Fig. 4 Examples for pseudo-values: cardiac death after 2.57 years (**a**), censored after 3.79 years (**b**), non-cardiac death after 2.07 years (**c**); different scales are used for (**b**) and (**c**) compared to (**a**)

of interest will obtain weights larger than one for time points later than their time of cardiac death, the value depending on the event time. These pseudo-values were used as dependent variable in a GEE model to account for multiple observations in the same subjects. Age, diabetes and 20 dummy variables indicating the timepoint were included as covariates. The independence working covariance matrix was used in the GEE model. Applying the function *geese* of the R library *geepack*, the influence of the covariates of interest on the cumulative incidence function for cardiac death was estimated using a complementary log–log (cloglog) link between the response and the linear predictor, so the estimated coefficients can be interpreted as logarithms of subdistribution hazard ratios. The results of the GEE model are presented in Table 6. Effects observed in the pseudo-value approach are similar to those obtained in the Fine and Gray model and can be interpreted analogously as an effect on the subdistribution hazard translating to an effect on the cumulative incidence function. A subdistribution hazard ratio comparing the high risk to the low risk group of $\exp(2.36) = 10.59$ was estimated (95 % CI 6.88 to 16.28). As described by Andersen and Perme (2010), the standard errors obtained in the pseudo-value approach are higher than those in the Fine and Gray regression model. Regression coefficients for the different time points partly presented in Table 6 are not of major interest, but are necessary for estimation of the cumulative incidence function. The cumulative incidence function estimated via the pseudo-observation approach (Fig. 5c) is similar to the cumulative incidence functions obtained from the cause-specific hazard regression or the subdistribution hazard regression. Steps of the function are obtained for each timepoint specified for the estimation of pseudo-values.

5 Discussion

In recent years statistical researchers as well as other applicants of statistical methods have become more and more aware of problems and pitfalls present in the analysis of time-to-event data with competing risks. Nevertheless, in a recent literature review of leading medical journals Koller et al. (2012) revealed, that competing risks data were not considered adequately in the analysis of time-to-event data in numbers of medical publications, although the problem is substantially described in statistical literature.

Table 6 Regression coefficients obtained by the pseudo-value approach—coefficients are not shown for all time points (skipped coefficients are monotonously increasing)

	$\hat{\beta}$	$\exp(\hat{\beta})$	Std. error	p value
Constant	−6.81	0.00	0.35	<0.001
Risk group	2.36	10.59	0.22	<0.001
Diabetes	0.81	2.25	0.25	0.001
Age ≥65	0.53	1.70	0.26	0.043
Time = 3 months	1.04	2.83	0.30	<0.001
Time = 6 months	1.21	3.35	0.26	<0.001
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Time = 60 months	2.74	15.49	0.18	<0.001

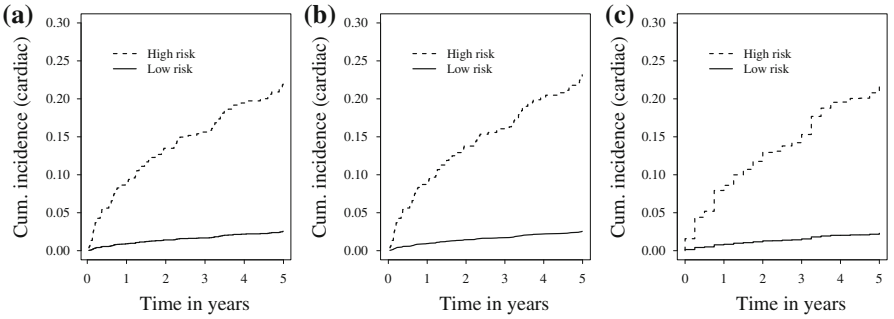

Fig. 5 Estimated cumulative incidence functions for cardiac death using the cause-specific hazard regression (**a**), the subdistribution hazard regression (**b**) and the analysis based on pseudo-observations (**c**) comparing the high risk (*dashed*) and low risk group (*solid line*)

Many articles and some textbooks published in the last two decades showed drawbacks or failures of classical time-to-event methods used in this situation. Most of these articles and text books, summarizing and describing approaches for the analysis of competing risks data, are either focussed on cause-specific hazard and subdistribution hazard regression models (e.g. Bakoyannis and Touloumi 2011; Putter et al. 2007) or additionally describe the mixture model approach (e.g. Lau et al. 2009). In this article these three methods are complemented by the vertical modelling approach presented by Nicolaie et al. (2010) and the analysis of survival data based on pseudo-observations proposed by Klein and Andersen (2005). In our article main ideas and theoretical background for these five approaches are presented and compared regarding intention of the modelling, model assumptions and interpretation of obtained results. Additionally, available literature describing the methods in more detail and extending the basic models are mentioned.

All methods were applied to a real dataset of a cohort study investigating risk stratification for patients who survived a MI. Based on the observed data the prespecified risk stratification seems to discriminate well between patients of high and low risk for cardiac death. Over ten times higher cause-specific and subdistribution hazard ratios (estimated by the Fine and Gray regression model or via the pseudo-observation

approach) for cardiac death were found for high risk patients compared to low risk patients, whereas hazard ratios for non-cardiac deaths were much smaller. This effect can also be seen in the results of the mixture model approach. Odds for dying from a cardiac reason were about nine times higher for a high risk patient than for a low risk patient. Graphical investigation of the results obtained by the vertical modelling approach revealed a higher relative hazard for cardiac deaths for patients identified as being of high risk, whereas the hazards of cardiac and non-cardiac deaths, given any event was observed, were pretty similar in the low risk group.

The well-established and commonly used methods, cause-specific hazard regression and subdistribution hazard regression, could be applied most easily, because of a variety of implemented functions provided in the statistical software package R. For the application of a mixture model, the vertical modelling approach and the analysis of survival times based on pseudo-observation some additional computation is necessary, but functions provided in standard software can be used for certain steps. A sketch of the R code used for data analysis is provided in the Appendix. In the analysed data, results of the cause-specific and the subdistribution hazard regression were very similar, which in general will not be the case, as presented e.g. by Dignam and Kocherginsky (2008) or Grambauer et al. (2010). Results of both modelling approaches have to be interpreted carefully. In the cause-specific hazard regression, the effect of the competing event(s) is not considered, so a higher cause-specific hazard does not necessarily translate into a higher cumulative incidence function, which means it does not translate into a higher proportion of observed events. In the subdistribution hazard regression, the covariates are directly linked to the cumulative incidence function of the event of interest. Hence it is possible that an observed effect of a covariate on the event of interest is caused by an effect on a competing event, which might lead to biological implausible results. Bakoyannis and Touloumi (2011) recommend to investigate cause-specific and subdistribution hazard regression simultaneously to avoid misinterpretation of the data. Putter et al. (2007) point out that the effects of the covariates on the competing event(s) have to be considered when the covariate effects on the cumulative incidence function are derived from a cause-specific hazard regression model. Andersen and Keiding (2012) question the interpretability and therefore the usefulness of the Fine and Gray approach, as in their opinion keeping individuals, who failed from a competing event, in the risk set is not justified.

Adequate estimation of a mixture model as proposed by Larson and Dinse (1985) gives estimates for the marginal probabilities for the type of event a subject might experience. A large number of samples and an appropriate follow-up time seem to be necessary to obtain valid estimates. Interpretation of the mixture model might be difficult due to the large number of parameters obtained. Since for each covariate included in the model the effect on the distribution of the event types and the event times given the type of event is estimated, this modelling approach seems to be adequate to explore the data and to find certain patterns or structures in the data, but it does not seem to be adequate to be used for hypothesis testing. The confidence intervals estimated from our data, which were obtained via the bootstrap method, were very wide, probably due to the large amount of censoring and the rather small number of events, and computations took several days due to the repeated application of the EM algorithm. Further research developing faster methods could improve the performance as well as new

technical developments as e.g. multicore computations as provided by the R package *multicore* (Urbanek 2011).

As the mixture model approach, the method of vertical modelling proposed by Nicolaie et al. (2010) appears to be a rather explorative tool estimating the relative hazards and evaluating the pattern of events happening in the course of time. The presentation of the marginal event time distribution on the one hand and the relative hazards given the event time on the other hand complicates the interpretation of the results. The approach does not seem to be adequate for a valid primary data analysis in a confirmatory clinical trial, but in certain applications the display of the relative hazards over time, describing which event is most likely to happen given any event happens, might add valid information.

The analysis of time-to-event data in the presence of competing risks based on pseudo-values using a GEE-model with complementary log–log link gives results similar to the subdistribution hazard model and can be interpreted analogously. Simulation studies showed that standard errors of this approach are larger than in the Fine and Gray model and so the pseudo-value approach was recommended not to be used for estimation when standard methods are available (Andersen and Perme 2010), but it might be very useful in situations where standard methods do not exist. Pseudo values can also be useful for checking model assumptions by investigation of pseudo residuals (Perme and Andersen 2008). The arbitrary choice of the number and placement of investigated time points appear to be drawbacks of that approach, but application of the established and well-investigated GEE-model might be used to develop some new ideas based on the pseudo-observation approach.

In a confirmatory clinical trial cause-specific hazard regression or subdistribution hazard regression seem to be most adequate for the primary analysis. Depending on the research question, one of these models should be selected a priori for the analysis of the primary endpoint. The model used for the primary analysis should be stated in the study protocol to avoid model selection based on the obtained results, which might lead to biased interpretations (see also Tai et al. 2011). Other models and effects of covariates on the competing events should be performed in the sense of sensitivity analyses to avoid misinterpretation of the results and to analyse patterns of events occurring over time in an explorative manner. Further research seems to be necessary to evaluate adequate measures for confirmatory hypothesis testing and to investigate consequences of covariate effects on the competing events for the interpretation of the primary endpoint.

Acknowledgments The cohort study conducted in the Klinikum rechts der Isar and in the German Heart Centre Munich was supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (13N/7073/7), the Kommission für Klinische Forschung, and the Deutsche Forschungsgemeinschaft (SFB 386).

Appendix: Applying competing risk regression models using R

Variables considered:

- Time: Event time or censoring time
- Status: Indicating type of event (1 = cardiac, 2 = non-cardiac, 0 = censoring)

- Group: Indicating risk group (0 = low risk group, 1 = high risk)
- Age: Indicating patient's age (0 = (age < 65), 1 = (age ≥ 65))
- Diab: Indicating diabetes (0 = no, 1 = yes)

Cause-specific hazard regression for the event of interest

```
require(survival)
COXcsh <- coxph(Surv(Time,Status==1)~Group+Age+Diab)
```

Subdistribution hazard regression

```
require(cmprsk)
COXsdh <- crr(Time,Status,cbind(Group,Age,Diab),failcode=1)
```

Mixture model

ECM algorithm as described by Ng and McLachlan (2003) for two possible types of failure and three covariates. For estimation of the cumulative hazard functions a dataset ordered by observed times is required. Expectation and conditional maximization steps have to be iterated until some predefined convergence criterion is fulfilled.

- Expectation for τ_i denoting the probability for a failure of type 1 for individual i in the $(k+1)$th iteration given kth estimates for μ, $\boldsymbol{\pi}$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and the baseline survival functions $S_{01}(t_i, \mathbf{x}_i, \boldsymbol{\beta_1}^{(k)})$ and $S_{02}(t_i, \mathbf{x}_i, \boldsymbol{\beta_2}^{(k)})$ according to Eqs. (8) and (10) from Ng and McLachlan (2003):

```
p1 <- exp(mu+Group*pi1+Age*pi2+Diab*pi3) /
 (1+exp(mu+Group*pi1+Age*pi2+Diab*pi3))
p2 <- 1-p1
tau <- p1*S01^exp(Group*b1_1+Age*b1_2+Diab*b1_3) /
 (p1*S01^exp(Group*b1_1+Age*b1_2+Diab*b1_3) +
 p2*S02^exp(Group*b2_1+Age*b2_2+Diab*b2_3))
```

- Function for $(k+1)$th estimation of μ and $\boldsymbol{\pi}$. *Q0* denotes the logistic regression component of the expectation of the complete-data log-likelihood given the current parameter estimates.

```
require(rootSolve)
Q0 <- function(MU) {
mu.opt <- MU[1]
pi1.opt <- MU[2]
pi2.opt <- MU[3]
pi3.opt <- MU[4]
P1 <- exp(mu.opt+pi1.opt*Group+pi2.opt*Age+pi3.opt*Diab)/
 (1+exp(mu.opt+pi1.opt*Group+pi2.opt*Age+pi3.opt*Diab))
fct.mu<-sum((as.numeric(Status==1)+(Status==0)*tau-P1))
fct.p1<-sum((as.numeric(Status==1)+(Status==0)*tau-P1)*Group)
fct.p2<-sum((as.numeric(Status==1)+(Status==0)*tau-P1)*Age)
fct.p3<-sum((as.numeric(Status==1)+(Status==0)*tau-P1)*Diab)
return(c(fct.mu,fct.p1,fct.p2,fct.p3)) }

opt <- multiroot(Q0,c(0,0,0,0))
mu.new <- opt$root[1]
```

```
pi1.new <- opt$root[2]
pi2.new <- opt$root[3]
pi3.new <- opt$root[4]
```

– Calculation of the cumulative baseline hazard function and the baseline survivor function for event type 1 in the $(k+1)$th iteration according to Eq. 12 from Ng and McLachlan (2003). Measures for $k = 2$ can be estimated analogously:

```
h01 <- c()

# Estimation of baseline hazard rate
n <- length(Time)
for(i in 1:n)
 h01[i] <- 1 / sum((((Status[i:n]==1) +
 (Status[i:n]==0)*tau[i:n])* exp(Group[i:n]*b1_1+
 Age[i:n]*b1_2+Diab[i:n]*b1_3)))*(Status[i]==1)

# Replace empty components at the end with zeros
h01[which(is.na(h01))] <- 0
# Calculate cumulative baseline hazard function
H01 <- cumsum(h01)
# Calculate baseline survival function
S01 <- exp(-H01)
```

– Conditional maximization step to obtain the $(k+1)$th estimate for $\boldsymbol{\beta}_1$. According to Eq. 9 from Ng and McLachlan (2003) maximization can be conducted separately for all types of event. $\boldsymbol{\beta}_2^{(k+1)}$ can be obtained analogously.

```
# Event type 1:
Q1 <- function(b1.opt)
{ b1.opt_1 <- b1.opt[1]
b1.opt_2 <- b1.opt[2]
b1.opt_3 <- b1.opt[3]
eta1 <- Group*b1.opt_1 +
Age*b1.opt_2 + Diab*b1.opt_3
fct1 <- sum(((Status==1)-((Status==1)+(Status==0)*tau) *
 H01*exp(Group*b1.opt_1+Age*b1.opt_2+Diab*b1.opt_3))*Group)
fct2 <- sum(((Status==1)-((Status==1)+(Status==0)*tau) *
 H01*exp(Group*b1.opt_1+Age*b1.opt_2+Diab*b1.opt_3))*Age)
fct3 <- sum(((Status==1)-((Status==1)+(Status==0)*tau) *
 H01*exp(Group*b1.opt_1+Age*b1.opt_2+Diab*b1.opt_3))*Diab)
return(c(fct1,fct2,fct3))}
b1 <- multiroot(Q1,c(0,0,0))$root
b1_1.new <- b1[1]
b1_2.new <- b1[2]
b1_3.new <- b1[3]
```

Vertical modelling

– Cox regression for marginal event-time distribution considering covariates

```
coxph(Surv(Time,Status>=1) ~ Group + Age + Diab)
```

– Estimation of relative hazards from a logistic regression model including B-splines for flexible influence of *Time* and interaction between *Group* and *Time*. Only individuals with an observed event can be considered. Estimates of relative hazards

for cardiac and non-cardiac death for the high risk group are calculated from the regression coefficients.

```
library(splines)
GLM <- glm(Status==1 ~ Group * bs(Time) + Diab +
Age, family=binomial(link='logit'),subset=Status>0)
# Relative hazard for cardiac death in the high risk group
rel.haz.highrisk.cardiac <-
predict(GLM,type='response',newdata=data.frame(Group=1,
Time=seq(0,5,length=300),Diab=mean(Diab),Age=mean(Age)))
# Relative hazard for non-card. death in the high risk group
rel.haz.highrisk.noncardiac <- 1 - rel.haz.highrisk.cardiac
```

Pseudo observations

R code for generation of pseudo observations and estimation of covariate effects applying a GEE model can be found in Klein et al. (2008).

References

Allignol A, Schumacher M, Wanner C, Drechsler C, Beyersmann J (2011) Understanding competing risks: a simulation point of view. BMC Med Res Methodol 11:86. doi:10.1186/1471-2288-11-86

Andersen PK, Keiding N (2012) Interpretability and importance of functionals in competing risks and multistate models. Stat Med 31:1074–1088. doi:10.1002/sim.4385

Andersen PK, Perme MP (2010) Pseudo-observations in survival analysis. Stat Methods Med Res 19:71–99. doi:10.1177/0962280209105020

Andersen PK, Klein JP, Rosthøj S (2003) Generalised linear models for correlated pseudo-observations, with applications to multi-state models. Biometrika 90:15–27. doi:10.1093/biomet/90.1.15

Bakoyannis G, Touloumi G (2011) Practical methods for competing risks data: a review. Stat Methods Med Res. doi:10.1177/0962280210394479

Barthel P, Schneider R, Bauer A, Ulm K, Schmitt C, Schömig A, Schmidt G (2003) Risk stratification after acute myocardial infarction by heart rate turbulence. Circulation 108:1221–1226. doi:10.1161/01.CIR.0000088783.34082.89

Bauer A, Kantelhardt JW, Barthel P, Schneider R, Makikallio T, Ulm K, Hnatkova K, Schömig A, Huikuri H, Bunde A, Malik M, Schmidt G (2006) Declaration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. Lancet 367:1674–1681. doi:10.1016/S0140-6736(06)68735-7

Bauer A, Barthel P, Schneider R, Ulm K, Müller A, Joeining A, Stich R, Kiviniemi A, Hnatkova K, Huikuri H, Schömig A, Malik M, Schmidt G (2009) Improved stratification of autonomic regulation for risk prediction in post-infarction patients with preserved left ventricular function (isar-risk). Eur Heart J 30:576–583. doi:10.1093/eurheartj/ehn540

Belot A, Abrahamowicz M, Remontet L, Giorgi R (2010) Flexible modeling of competing risks in survival analysis. Stat Med 29(23):2453–2468. doi:10.1002/sim.4005

Beyersmann J, Schumacher M (2007) Letter to the editor: Misspecified regression model for the subdistribution hazard of a competing risk. Stat Med 26:1649–1652. doi:10.1002/sim.2727

Beyersmann J, Dettenkofer M, Bertz H, Schumacher M (2007) A competing risks analysis of bloodstream infection after stem-cell transplantation using subdistribution hazards and cause-specific hazards. Stat Med 26:5360–5369. doi:10.1002/sim.3006

Beyersmann J, Latouche A, Buchholz A, Schumacher M (2009) Simulating competing risks data in survival analysis. Stat Med 28(6):956–971. doi:10.1002/sim.3516

Beyersmann J, Schumacher M, Allignol A (2012) Competing risks and multistate models with R. Springer, New York

Cox DR (1972) Regression models and life-tables. J R Stat Soc Ser B 34:187–220. doi:10.2307/2985181

Cox C, Chu H, Schneider MF, Muñoz A (2007) Tutorial in biostatistics: parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. Stat Med 26:4352–4374. doi:10.1002/sim.2836

Crowder MJ (2001) Classical competing risks. Chapman & Hall/CRC, Boca Raton

Dignam JJ, Kocherginsky MN (2008) Choice and interpretation of statistical tests used when competing risks are present. J Clin Oncol 26(24):4027–4034. doi:10.1200/JCO.2007.12.9866

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. Chapman and Hall/CRC, London

Escarela G, Bowater RJ (2008) Fitting a semi-parametric mixture model for competing risks in survival data. Commun Stat 37:277–293. doi:10.1080/03610920701649134

Fahrmeir L, Tutz G (2001) Multivariate statistical modelling based on generalized linear models. Springer, New York

Fine JP, Gray RJ (1999) A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 94:496–509. doi:10.2307/2670170

Friedman M (1982) Piecewise exponential models for survival data with covariates. Ann Stat 10:101–113

Grambauer N, Schumacher M, Beyersmann J (2010) Proportional subdistribution hazards modeling offers a summary analysis, even if misspecified. Stat Med 29:875–884. doi:10.1002/sim.3786

Gray R (1988) A class of k-sample tests for comparing the cumulative incidence function in the presence of a competing risk. Ann Stat 16:1141–1154. doi:10.2307/2241622

Gray B (2010) cmprsk: Subdistribution analysis of competing risks. URL http://CRAN.R-project.org/package=cmprsk, R package version 2.2-1

Hastie TJ (1997) Generalized additive models. Chapman & Hall/CRC, New York

Højsgaard S, Halekoh U, Yan J (2005) The R package geepack for generalized estimating equations. J Stat Softw 15:1–11. http://CRAN.R-project.org/package=survival, R package version 2.36-5

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data. Wiley, Hoboken

Kaplan EL, Meier P (1958) Non-parametric estimation from incomplete observations. J Am Stat Assoc 53:457–481

Klein JP (2010) Competing risks. WIREs Comput Stat 2:333–339. doi:10.1002/wics.83

Klein JP, Andersen PK (2005) Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. Biometrics 61:223–229. doi:10.1111/j.0006-341X.2005.031209.x

Klein JP, Moeschberger ML (2003) Survival analysis—techniques for censored and truncated data. Springer, New York

Klein J, Gerster M, Andersen P, Tarima S, Perme M (2008) SAS and R functions to compute pseudo-values for censored data regression. Comput Methods Prog Biomed 89(3):289–300. doi:10.1016/j.cmpb.2007.11.017

Koller MT, Raatz H, Steyerberg EW, Wolbers M (2012) Competing risks and the clinical community: irrelevance or ignorance. Stat Med 31:1089–1097

Larson MG, Dinse GE (1985) A mixture model for the regression analysis of competing risks data. J R Stat Soc Ser C 34:201–211

Latouche A, Boisson V, Chevret S, Porcher R (2007) Misspecified regression model for the subdistribution hazard of a competing risk. Stat Med 26(5):965–974. doi:10.1002/sim.2600

Lau B, Cole SR, Moore SR, Gange SJ (2008) Evaluating competing adverse and beneficial outcomes using a mixture model. Stat Med 27:4313–4327. doi:10.1002/sim.3293

Lau B, Cole SR, Gange SJ (2009) Competing risk regression models for epidemiologic data. Am J Epidemiol 170:244–256. doi:10.1093/aje/kwp107

Lau B, Cole S, Gange S (2011) Parametric mixture models to evaluate and summarize hazard ratios in the presence of competing risks with time-dependent hazards and delayed entry. Stat Med 30:654–665. doi:10.1002/sim.4123

Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22. doi:10.1093/biomet/73.1.13

Lunn M, McNeil D (1995) Applying Cox regression to competing risks. Biometrics 51:524–532. doi:10.2307/2532940

McLachlan GJ, Krishnan T (1997) The EM algorithm and extensions. Wiley, New York

Miller RG (1974) The Jackknife—a review. Biometrika 6(1):1–15. doi:0.1093/-biomet/61.1.1

Ng GK, McLachlan GJ (2003) An em-based semi-parametric mixture model approach to the regression analysis of competing-risks data. Stat Med 22:1097–1111. doi:10.1002/sim.1371

Nicolaie MA, van Houwelingen HC, Putter H (2010) Vertical modeling: a pattern mixture approach for competing risks modeling. Stat Med 29:1190–1205. doi:10.1002/sim.3844

Perme MP, Andersen PK (2008) Checking hazard regression models using pseudo-observations. Stat Med 27:5309–5328. doi:10.1002/sim.3401
Pintilie M (2006) Competing risks: a practical perspective. Wiley, Chichester
Prentice R, Kalbfleisch J, Peterson A, Flournoy N, Farewell V, Breslow N (1978) The analysis of failure times in the presence of competing risks. Biometrics 34:541–554. doi:10.2307/2530374
Putter H, Fiocco M, Geskus RB (2007) Tutorial in biostatistics: competing risks and multi-state models. Stat Med 26(11):2389–2430. doi:10.1002/sim.2712
R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/, ISBN 3-900051-07-0
Robins J, Rotnitzky A (1992) Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell N, Dietz K, Farewell V (eds) AIDS epidemiology—methodological issues. Birkhäuser, Boston pp 24–33
Scheike TH, Zhang MJJ (2008) Flexible competing risks regression modeling and goodness-of-fit. Lifetime Data Anal 14:464–483. doi:10.1007/s10985-008-9094-0
Schemper M, Smith T (1996) A note on quantifying follow-up in studies of failure time. Lancet 17:343–346
Schoenfeld D (1982) Partial residuals for the proportional hazards regression model. Biometrika 69(1):239–241. doi:10.1093/biomet/69.1.239
Sun Y, Hyun S, Gilbert P (2008) testing and estimation of time-varying cause-specific hazard ratios with covariate adjustment. Biometrics 64:1070–1079. doi:10.1111/j.1541-0420.2008.01012.x
Tai BC, Wee J, Machin D (2011) Analysis and design of randomised clinical trials involving competing risks endpoints. Trials 12:127. doi:10.1186/1745-6215-12-127
Therneau T (2011) Survival: survival analysis, including penalised likelihood. http://CRAN.R-project.org/package=survival, R package version 2.36-5
Therneau TM, Grambsch PM (2000) Modeling survival data: extending the Cox model (statistics for biology and health). Springer, New York
Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci USA 72(1):20–22
Urbanek S (2011) multicore: parallel processing of R code on machines with multiple cores or CPUs. http://CRAN.R-project.org/package=multicore, R package version 0.1-5