Modeling marginal features in studies of recurrent events in the presence of a terminal event

Per Kragh Andersen

Jules Angst

Henrik Ravn

Research Report 18/01

Department of Biostatistics
University of Copenhagen

Modeling marginal features in studies of recurrent events in the presence of a terminal event

Per Kragh Andersen,

Jules Angst[1],

Henrik Ravn[2]

Abstract: We study models for recurrent events with special emphasis on the situation where a terminal event acts as a competing risk for the recurrent events process and where there may be gaps between periods during which subjects are at risk for the recurrent event. We focus on marginal analysis of the expected number of events and show that an Aalen-Johansen type estimator proposed by Cook and Lawless is applicable in this situation. A motivating example deals with psychiatric hospital admissions where we supplement with analyses of the marginal distribution of time to the competing event and the marginal distribution of the time spent in hospital. Pseudo-observations are used for the latter purpose.

Key words: competing risks; expected number of events; intensity-based models; marginal models; pseudo-observations; recurrent events.

[1]: Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zürich, Switzerland,

[2]: Novo-Nordisk, Bagsværd, Denmark.

1 Introduction

Many studies in epidemiology, clinical medicine, and other fields involve events which may happen multiple times in each subject. Examples include recurrent hospital admissions (e.g., Kessing et al., 2004), repeated tumor occurrences (e.g., Byar 1980) or recurrent pulmonary exacerbations (Fuchs et al., 1994). Methods for analyzing such data are important since restricting attention to the first occurrence of the event will be utilizing data inefficiently. As nicely summarized by Cook and Lawless (2007) there are two major approaches to the analysis of recurrent events: *intensity-based* models and *marginal* models. In the former, the situation is considered as a special case of a multi-state model based on *counting processes* (Andersen et al., 1993) and standard Cox-type or other well-known hazard models from survival analysis apply to the intensities of event occurrence. Examples of such models include the 'PWP' model (Prentice et al., 1981), the 'AG' model (Andersen and Gill, 1982), and models based on *gap times* between events (e.g., Cook and Lawless, 2007, Ch. 4). Dependence among repeated events may in these approaches be modelled by regressing the intensity on information on previous events via time-dependent covariates, see e.g. Andersen and Gill (1982), Aalen et al. (2004) and Fosen et al. (2006). Extended versions of such hazard models may include random effects ('frailties' e.g., Kessing et al., 1999; Hougaard, 2000) to model this intra-individual association. Intensity-based models have the advantage that the probability distribution of the entire stochastic process may be specified which, further, allows for likelihood-based inference and for model-based simulations. Furthermore, quite general (e.g., event-dependent) censoring is easily accommodated. On the other hand, having the ambition of modelling the distribution of the entire recurrent events process, one also runs the risk of model mis-specification and, for such reasons, the second major, 'marginal', modelling approach has been put forward. Here, attention is restricted to certain marginal parameters, most prominently the marginal mean $\mu(t) = E(N(t))$ where $N(t)$ is the recurrent events counting process (e.g., Lawless and Nadeau, 1995; Lin et al., 2000) but also models for the marginal distributions of times to first, second, etc. event have been studied (Wei et al., 1989). To estimate such marginal parameters censoring should, in the simplest situation, be completely independent of the recurrent events process, though inverse probability of censoring weighted versions of the various estimators also exist, see e.g., Cook and Lawless (2007).

In biological applications of methods for recurrent events there will almost inevitably be *competing risks* in the form of terminal events, the occurrence of which prevents further events from happening. Thus, the death of a patient will, obviously, have the consequence that no further hospital admissions, tumors etc. can occur and while intensity-based models readily accommodate this situation by extending the state space for the multi-state model (Figure 1), more care must be exercised for marginal analysis.

The situation is similar to survival analysis in the presence of competing risks where hazards ('rates', 'intensities') immediately generalize to cause-specific

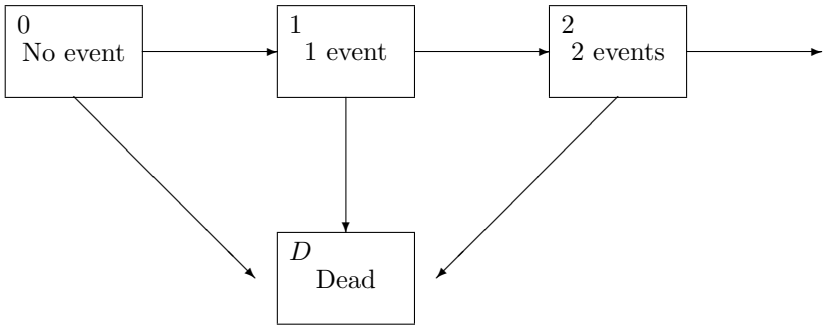

Figure 1: A recurrent events process with a terminal event and no 'gaps' between at-risk periods.

hazards and where methods for estimation in hazard models for survival data can also be applied with few changes to inference for cause-specific hazards (e.g., Prentice et al. 1978; Andersen et al., 1993). Similarly, standard hazard-based models may be applied to the intensities in the model depicted in Figure 1, and a correct modelling of all intensities will specify the entire probability distribution of the process and functionals like $\mu(t)$ may be estimated by plug-in (or by micro-simulation, e.g. Mitton et al., 2000; Iacobelli and Carstensen, 2013).

However, also in the presence of a terminal event, focus for the inference may be on a parameter like $\mu(t)$, and methods for estimating this marginal mean non-parametrically have been derived (Cook and Lawless, 1997; Ghosh and Lin, 2000), as well as regression models (Ghosh and Lin, 2002; Cook et al., 2009). Such models, again similarly to survival analysis with competing risks, only partially specify the process in Figure 1 and, as detailed, e.g. by Ghosh and Lin (2002), attention must also be paid to the survival part of the process. This may be done using a standard model for $S(t) = P(T > t)$ with T being the time to the terminal event. However, it is not entirely obvious neither how to formally combine the two models, nor whether such a formal combination is desirable. The need to study also $S(t)$ even though the main focus may be on $\mu(t)$ is much in line with the recommendations given by Latouche et al. (2013) for competing risks, namely that all cause-specific hazards and all cumulative incidences should be studied even if the primary interest is on a single cause.

A further complication that may arise in the analysis of recurrent events (both with or without a terminal event) is that there may be time periods dur-

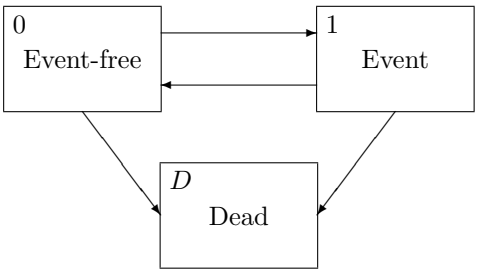

Figure 2: A recurrent events process with a terminal event and 'gaps' between at-risk periods.

ing which a subject is not at risk for a new event. An obvious example is in a study of hospital admissions where, once the patient is in hospital, no new admissions can occur. The situation is depicted in Figures 2 and 3. In the latter figure, the numbers of previous events are shown explicitly but the two ways of showing the model are equivalent. In such a model, one may still target the parameters $\mu(t)$ and $S(t)$ but, at least the former, will depend on the distribution of times not at risk (e.g., the expected number of hospitalizations in a given period will depend on the distribution of length of stay in hospital).

In this paper we will study recurrent events with competing risks in the form of a terminal event, paying special attention to the situation where there may be gaps between periods during which subjects are at risk for new events. We will focus on marginal analysis since, as mentioned above (and discussed by, e.g. Hu et al., 2011) intensity-based models are, in principle, straightforward. We will demonstrate, in a simulation study, how the non-parametric estimator for $\mu(t)$ introduced by Cook and Lawless (1997) still applies when there are gaps between at-risk periods. We will further, in a case study concerning psychiatric hospital admissions, discuss how inference, at least informally, may be based on separate models for $\mu(t)$, $S(t)$, and $Q_1(t)$, the probability of being not-at-risk, i.e. being in the state '1' of Figure 2. Regression analysis via *pseudo-observations* (Andersen and Pohar Perme, 2010) will play a prominent role. The data come from a clinical study of patients with affective disorders conducted in Switzerland (Angst et al., 2003; Kessing et al., 2004).

The structure of the paper is, as follows. In Section 2 we will establish the notation and review previously used methods. We will also introduce pseudo-observations. Section 3 presents a small simulation study to support non-parametric estimation of $\mu(t)$, and Section 4 presents the example concern-

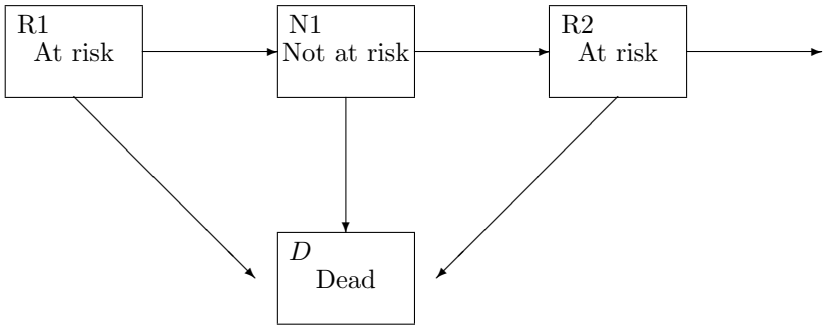

Figure 3: A recurrent events process with a terminal event and 'gaps' between at-risk periods.

ing psychiatric admissions. The final Section 5 contains a brief discussion of our findings.

2 Notation and methods

We consider independent subjects labelled $i = 1, \ldots, n$ and, for subject i, we let $T_{i1} < T_{i2} < \ldots$ be the times of occurrences for the recurrent event, that is, T_{ij} is the time of entry into state j in Figure 1 or the time of entry into state Nj in Figure 3 (or $T_{i1}, T_{i2}, \ldots$ are the times of entry or re-entry into state 1 in Figure 2). The associated counting process is

$$N_i(t) = \sum_j I(T_{ij} \leq t)$$

with mean

$$\mu(t) = E(N(t))$$

and (marginal) rate function

$$\rho(t) = E(dN(t))/dt$$

such that $\mu(t) = \int_0^t \rho(u)du$. Let, further, T_i be the survival time, i.e. the time of entry into state D in either of the Figures 1-3, and let the survival function be $S(t) = P(T > t)$, i.e., $F(t) = 1 - S(t) = Q_D(t)$, say, is the state occupation probability for the absorbing state D. Similarly, let $Q_j(t)$ denote the state occupation probability for any state, j in Figures 1-3. Both $N_i(t)$ and T_i may

be incompletely observed because of right-censoring at C_i, i.e. the observed counting process is

$$N_i^*(t) = \int_0^t I(C_i \geq u) dN_i(u),$$

and the right-censored survival time is $T_i^* = T_i \wedge C_i$. For subject i, a vector Z_i of time-fixed covariates may also be available.

The parameters $\mu(t), \rho(t), Q_j(t), S(t)$ and also the survival functions $S_j(t) = P(T_{ij} > t)$ are all examples of *marginal* parameters where, intuitively, one places oneself at $t = 0$ and asks about probabilities or expected values of future events. Alternatively, one can focus on *intensities* for the multi-state processes, say $X(t)$ depicted in Figures 1-3, i.e.

$$\lambda_{hk}(t) = P(X(t + dt) = k \mid \mathcal{F}_{t-}, X(t-) = h)/dt,$$

where $\mathcal{F}_t$ is the *history* $(X(u), u \leq t; Z)$ of the process at time t and h, k are states. Here, one places oneself at time t and asks about probabilities of events in the next little time interval from t to $t + dt$ *conditionally* on information available just before time t.

From the intensities, marginal features may, in principle, be obtained via plug-in. Thus, assuming that the model in Figure 2 is *Markovian* we have the Doob-Meyer decomposition (e.g., Andersen et al., 1993)

$$N(t) = \int_0^t \alpha_{01}(u) I(X(u-) = 0) du + M(t),$$

where $\alpha_{01}(t)$ is the transition *hazard*, that is, the intensity is $\lambda_{01}(t) = \alpha_{01}(t) I(X(t-) = 0)$, and $M(t)$ is a martingale. Thereby, the marginal mean becomes

$$\mu(t) = E(N(t)) = \int_0^t \alpha_{01}(u) Q_0(u) du, \tag{1}$$

which may be estimated by combining estimators for $A_{01}(t) = \int_0^t \alpha_{01}(u) du$ and for the state occupation probability $Q_0(t)$.

In the following we will, however, rather study estimators for the marginal mean $\mu(t)$ that are not based on a specification of all the transition intensities of the multi-state model. Such estimators were introduced for the model in Figure 1 by Cook and Lawless (1997) with detailed asymptotic results by Ghosh and Lin (2000). This non-parametric estimator builds on writing

$$\mu(t) = E(N(t)) = \int_0^t S(u) dR(u)$$

with $R(t) = E(N(t) \mid T > t)$ yielding the estimator

$$\widehat{\mu}(t) = \int_0^t \widehat{S}(u-) \frac{\sum_i dN_i^*(u)}{\sum_i I(T_i^* \geq u)}, \tag{2}$$

where $\widehat{S}$ is the Kaplan-Meier estimator based on the censored survival times T_i^* and associated death indicators $I(T_i \leq C_i)$. Note that the parameter $R(t)$ itself is not easily interpretable (because it conditions on the future) but the resulting estimator (2) is, indeed, estimating the marginal mean $\mu(t)$. This is because, for uncensored data, the estimating equations

$$\sum_i \int_0^\infty \big(dN_i(u) - dR(u)I(T_i > u)\big)$$

are unbiased for $R(\cdot)$ and so are

$$\sum_i \int_0^\infty \big(dN_i^*(u) - dR(u)I(T_i^* > u)\big) \tag{3}$$

under a censoring scheme where the censoring times C_i are independent of $(N_i(t), T_i)$. For more general (e.g., event-dependent) censoring schemes, inverse probability of censoring weighted versions of (3) exist (Cook and Lawless, 2007).

Even though (2) was derived for the situation without periods where subjects are still alive but currently not at risk for another recurrent event, we will use it also for the model in Figures 2-3. The arguments outlined above for the properties of $\widehat{\mu}(t)$ from the situation without times not-at-risk carry over to this situation, and the estimator was in fact used by Cook and Lawless (2007, Section 5.5.1) in a such a case. To support this view we present a small simulation study in the next section. In studies of psychiatric admissions, of which we present an example in Section 4, ignoring times when patients are hospitalized is some times referred to as studying disease 'cycles' rather than studying times from discharge to admission (which is known as studying disease 'episodes', e.g., Angst et al., 2003).

One should notice the discrepancy between the estimator $\widehat{\mu}(t)$ given in (2), the Nelson-Aalen type estimator $\widehat{\mu}_{LN}(t) = \int_0^t \sum_i dN_i^*(u) / \sum_i I(T_i^* \geq u)$ (which is here identical to $\widehat{R}(t)$), and the Nelson-Aalen estimator $\widehat{A}_{01}(t) = \int_0^t \sum_i dN_i^*(u) / \sum_i I(X_i^*(u) = 0)$ (where $X_i^*(\cdot)$ is the observed, censored, multi-state process for subject i). The former, as mentioned, estimates $\mu(t)$ in the presence of a terminating event; the second estimates the mean of $N(t)$ is the situation *without* a terminating event (Lawless and Nadeau, 1995) (and will overestimate $\mu(t)$ in the presence of a terminating event because it then treats deaths as censorings), and the latter estimates the integrated transition hazard, $A_{01}(t)$, under a Markov assumption for the multi-state process (e.g., Andersen et al., 1993).

Similarly to (2), the regression model for $\mu(t)$ discussed by Ghosh and Lin (2002) is also applicable in situations with periods not-at-risk. However, it is important to keep in mind (as also discussed by Ghosh and Lin, 2002) that a model for $\mu(t)$ only gives a *partial specification* of the multi-state model, and if, e.g. a covariate is associated with a reduction of $\mu(t)$ then this could be a consequence of either reducing the event intensity for the recurrent event or

increasing the mortality rate (or both). For this reason it is important not to let an analysis of $\mu(t)$ stand alone but to supplement it with studies of both $S(t)$ and some aspect of the distribution of times not-at-risk, such as the probability $Q_1(t)$ or the average length of stay up to time t in state 1 of Figure 2, i.e. $\int_0^t Q_1(u)du$. We will illustrate this in the example of Section 4 where we will use *pseudo-observations* (e.g., Andersen and Pohar Perme, 2010) for this purpose.

Pseudo-observations are defined, as follows. Let θ be a marginal mean parameter (such as $\mu(t) = E(N(t))$ or $Q_1(t) = E(I(X(t) = 1))$) and let $\widehat{\theta}$ be an estimator of θ based on independent subjects $i = 1, \ldots n$, possibly incompletely observed because of right-censoring. The ith *pseudo-observation* for the incompletely observed random variable (e.g., $N(t)$ or $I(X(t) = 1)$) is

$$\theta_i = n\widehat{\theta} - (n-1)\widehat{\theta}^{-i}, \tag{4}$$

where $\widehat{\theta}^{-i}$ is the estimator applied to the sample of size $n-1$ obtained by eliminating subject i.

The main application of pseudo-observations has been to study regression models for parameters θ for which no other simple techniques are available, such as the cause-specific number of years lost (Andersen, 2013) or the average length of stay in the diseased state of an illness-death model (Grand and Putter, 2016). Asymptotic properties of the pseudo-observations when censoring is independent of covariates have been derived in special cases, including those based on the Kaplan-Meier estimator for $S(t)$ and the Aalen-Johansen estimator for the cumulative incidence function with competing risks (Graw et al., 2009; Martinussen and Jacobsen, 2016; Overgaard et al., 2017). For covariate-dependent censoring, Binder et al. (2014) suggested to base the computation of pseudo-observations in (4) on alternative, inverse probability of censoring weighted estimators. In addition to $S(t)$ we will, in our example in Section 4, use pseudo-observations for the parameters $\mu(t)$ and $Q_1(t)$ even though asymptotic properties have not been formally proven for these situations. The technique based on von Mises expansions used by Jacobsen and Martinussen (2016) and Overgaard et al. (2017) will likely carry through to our situation but it is beyond the scope of the present paper to pursue this.

3 A small simulation study

In this section we present a small simulation study of how the estimator (2) works in situations where there are periods during which subjects are not at risk for the recurrent event. We will use the fact that when all transition hazards in Figure 2 are constant, the state occupation probabilities $Q_0(t), Q_1(t), Q_D(t)$ are explicit functions of the transition hazards α_{hk} (given by matrix exponentials, e.g. Chiang, 1980). Let

$$\alpha_0 = -(\alpha_{01} + \alpha_{0D}), \quad \alpha_1 = -(\alpha_{10} + \alpha_{1D}),$$

and

$$\begin{aligned}\rho_0 &= \big(\alpha_0+\alpha_1+\sqrt{(\alpha_0-\alpha_1)^2+4\alpha_{01}\alpha_{10}}\big)/2,\\ \rho_1 &= \big(\alpha_0+\alpha_1-\sqrt{(\alpha_0-\alpha_1)^2+4\alpha_{01}\alpha_{10}}\big)/2.\end{aligned}$$

Then

$$Q_0(t)=\frac{\rho_0-\alpha_1}{\rho_0-\rho_1}\exp(\rho_0 t)+\frac{\rho_1-\alpha_1}{\rho_1-\rho_0}\exp(\rho_1 t)$$

and, thereby by (1)

$$E(N(t))=\alpha_{01}\left((\frac{k_0}{\rho_0}(\exp(\rho_0 t)-1)+\frac{k_1}{\rho_1}(\exp(\rho_1 t)-1)\right), \tag{5}$$

where $k_0=\frac{(\rho_0-\alpha_1)\alpha_{0D}+\alpha_{01}\alpha_{0D}}{\rho_0(\rho_0-\rho_1)}$, and $k_1=\frac{(\rho_1-\alpha_1)\alpha_{0D}+\alpha_{01}\alpha_{0D}}{\rho_1(\rho_1-\rho_0)}$.
We considered the following scenarios where, in all cases, $\alpha_{01}=2$:

Scenario	α_{10}	$\alpha_{0D}=\alpha_{1D}$	$Q_0(10)$	$Q_1(10)$	$Q_D(10)$
1	0.4	0.2,	0.022	0.113	0.865
2	1.0	0.2	0.045	0.090	0.865
3	0.4	0.4	0.003	0.015	0.982

So, in all scenarios we have the same hazard for the recurrent event; in scenario 2, the hazard α_{10} of going back 'at risk' is higher, thereby, increasing $Q_0(t)$ and $\mu(t)$, and in scenario 3 the mortality rates are higher than in scenario 1, thereby decreasing $\mu(t)$.

We focus on $\mu(10)$ and repeat the three scenarios for sample sizes $n=1000, 250, 100$ and with either administrative censoring at time 10 or with random exponential censoring with rate 0.3. All combinations were repeated 5000 times and Table 1 summarizes the results. We also show the bias resulting from treating death as 'independent censoring', i.e. by estimating $\mu(t)$ by the Nelson-Aalen type estimator $\widehat{\mu}_{LN}(t)=\int_0^t \sum_i dN_i^*(u)/\sum_i I(T_i^*\geq u)$. We also calculated the Nelson-Aalen estimator for $A_{01}(10)$ but since that resulted in estimates close to the correct value of $10\alpha_{01}=20$ in all cases, results are not included in the table.

It is seen that the estimator $\widehat{\mu}(t)$ defined in (2) is everywhere unbiased with an SD that decreases with n. The simple Nelson-Aalen type estimator $\widehat{\mu}_{LN}(t)$ treating deaths as censoring has a large positive bias. We added a scenario where $\alpha_{0D}=\alpha_{1D}=0$, i.e. there is no mortality, in which case the two estimators coincide and both are unbiased (for $n=1000$ and $\alpha_{10}=0.4$ the mean bias was 0.00047 (SD=0.0478) with administrative censoring and 0.00364 (SD=0.127) with random censoring).

Figure 4 shows the true $\mu(t)$ together with estimates based on (2) for 100 realizations from scenario 1 with administrative censoring.

				Bias (SD) for $\widehat{\mu}(t)$ (2)		Bias for $\widehat{\mu}_{LN}(t)$
α_{10}	α_D	$E(N(10))$	n	Administrative censoring	Random censoring	Administrative censoring
0.4	0.2	2.08213	1000	0.00009 (0.050)	-0.00037 (0.087)	1.95
			250	-0.00154 (0.097)	-0.00442 (0.182)	1.94
			100	-0.00217 (0.150)	-0.02737 (0.280)	1.94
1.0	0.2	3.29888	1000	-0.00068 (0.082)	-0.00155 (0.140)	3.81
			250	-0.00256 (0.161)	-0.00572 (0.288)	3.81
			100	-0.00184 (0.255)	-0.04459 (0.453)	3.82
0.4	0.4	1.41331	1000	-0.00071 (0.037)	-0.00204 (0.056)	2.61
			250	-0.00084 (0.073)	-0.00925 (0.111)	2.62
			100	-0.00109 (0.113)	-0.01989 (0.167)	2.56

Table 1: Simulations, 5000 samples, n= no of individuals, $\alpha_{01} = 2.0$, α_D denotes the common value for $\alpha_{0D} = \alpha_{1D}$.

The computations were performed using SAS 9.4. Here, the estimator $\widehat{\mu}(t)$ (2) is available in `PROC PHREG` by treating the situation as competing risks with several records per subject.

4 Example: psychiatric admissions

We study readmission data for 119 psychiatric patients who had a first diagnosis of affective disorder given at Psychiatric Hospital, University of Zürich, Switzerland between 1959 and 1963 (e.g., Kessing et al., 2004). A first diagnosis of unipolar disorder was given to 98 patients while the remaining 21 had bipolar disorder; 84 patients were females and the average age at first diagnosis was 47.2 years (SD=17.1 years). During follow-up, which averaged 17.1 years, 78 patients died and the patients had on average 5.6 episodes (range 1 to 26).

We first studied models for the admission intensity $\lambda_{01}(t)$, i.e. patients were considered not at-risk while in hospital. Figure 5 shows the estimated cumulative transition hazards for unipolar and bipolar patients. The higher hazard for bipolar patients seen in the figure is also illustrated by fitting a Cox-type ('AG') model for the admission intensity, yielding an estimated ratio of $\exp(0.372) = 1.45$ (with 95% confidence limits from 1.08 to 1.43) between the two groups. Adjusting the model for the number of previous admissions as a time-dependent covariate, however, attenuated the ratio to $\exp(0.095) = 1.10$ (0.83, 1.43). This change in the estimated ratio is as one would expect, since we now condition on an intermediate variable that masks the difference between the two groups.

Such scenarios have been taken as an argument for focussing on marginal models instead, in particular when analyzing data from randomized studies (e.g., Cook and Lawless, 2007). We therefore estimated the expected numbers of re-admissions according to diagnosis (bipolar vs. unipolar disorder) using

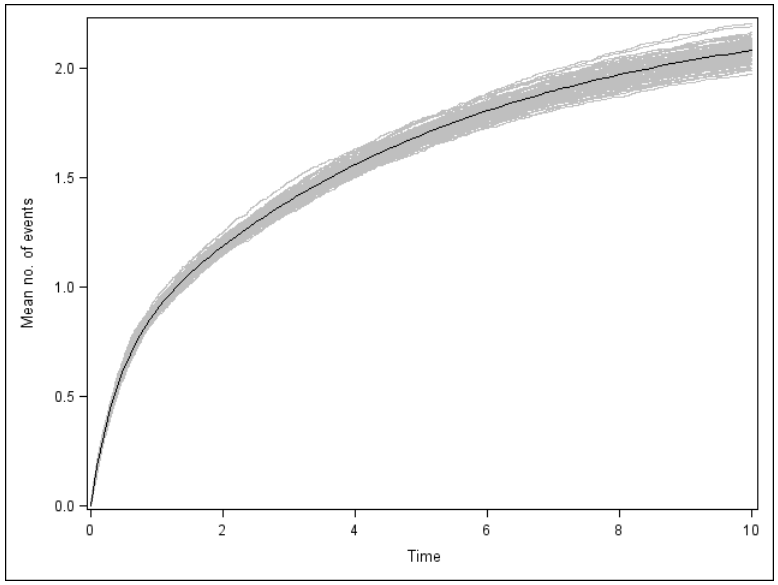

Figure 4: True $\mu(t)$ with 100 simulated curves: scenario 1

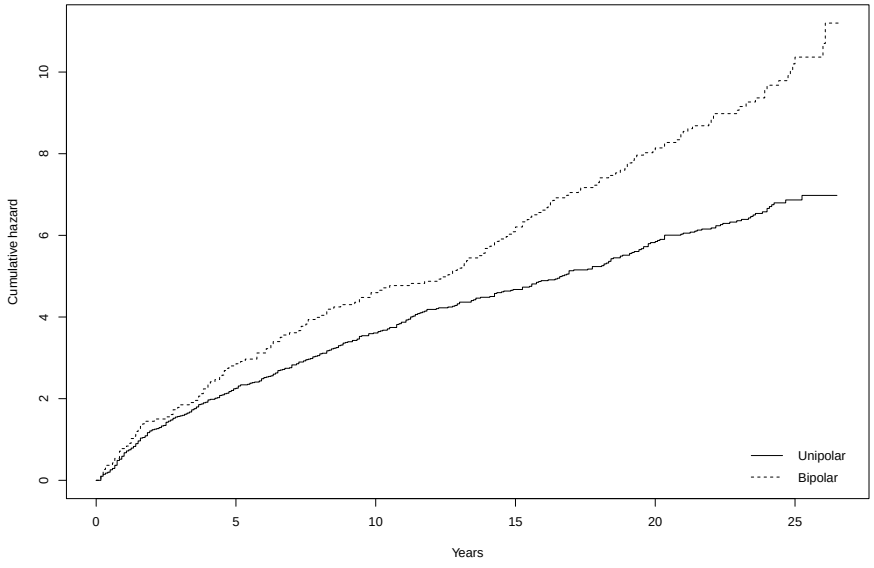

Figure 5: Cumulative hazards of re-admission for unipolar and bipolar patients.

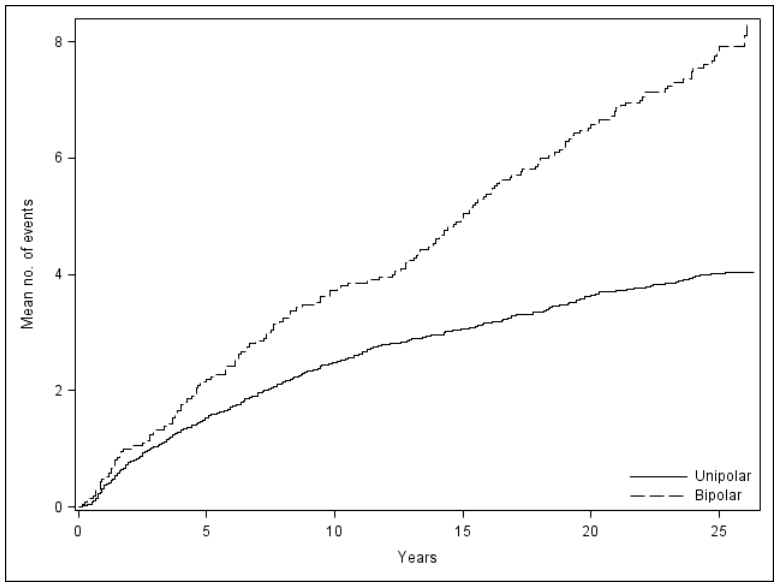

Figure 6: Expected numbers of re-admissions for unipolar and bipolar patients.

$\widehat{\mu}(t)$ defined in (2) - see Figure 6. It is seen that bipolar patients have more re-admissions, a result that is sustained by fitting the multiplicative means model of Ghosh and Lin (2002) which results in an estimated mean ratio of $\exp(0.673) = 1.96$ (1.49, 2.57). Adjustment for potential confounders like sex and age at disease onset changed the estimated mean ratio into $\exp(0.464) = 1.59$ (1.20, 2.11). Thus, bipolar patients, on average, have more admissions than unipolar patients. However, analysis of the marginal mean does not reveal to what extent this can be ascribed to a lower mortality rate and/or shorter time spent in hospital. We therefore supplemented the above analyses of the marginal mean by an analysis of pseudo-observations at time 10 years for: (a) the marginal mean, $\mu(10)$, (b) the survival probability, $S(10)$, and (c) the average time spent in hospital up to 10 years after diagnosis, $\int_0^{10} Q_1(t)dt$. The pseudo-observations for parameter (c) were based on the Aalen-Johansen estimator (e.g., Andersen et al., 1993) for the state occupation probabilities evaluated at time 10 years, see Figure 7. The results from the regression analysis of pseudo-observations are summarized in Table 2. It is seen that the adjusted coefficient for bipolar disease vs. unipolar disease ($0.330 = \log(1.39)$) is somewhat smaller than the corresponding estimate based on the Ghosh-Lin model and with a larger standard error. Analyses of the survival probability and expected length of stay suggest that bipolar patients both live longer (hazard ratio $\exp(-1.123) = 0.325$) and spend shorter time (estimate 0.109 years during 10 years) in hospital than unipolar patients, see Figures 7-8. These tendencies, however, are insignificant when based on pseudo-values at 10 years.

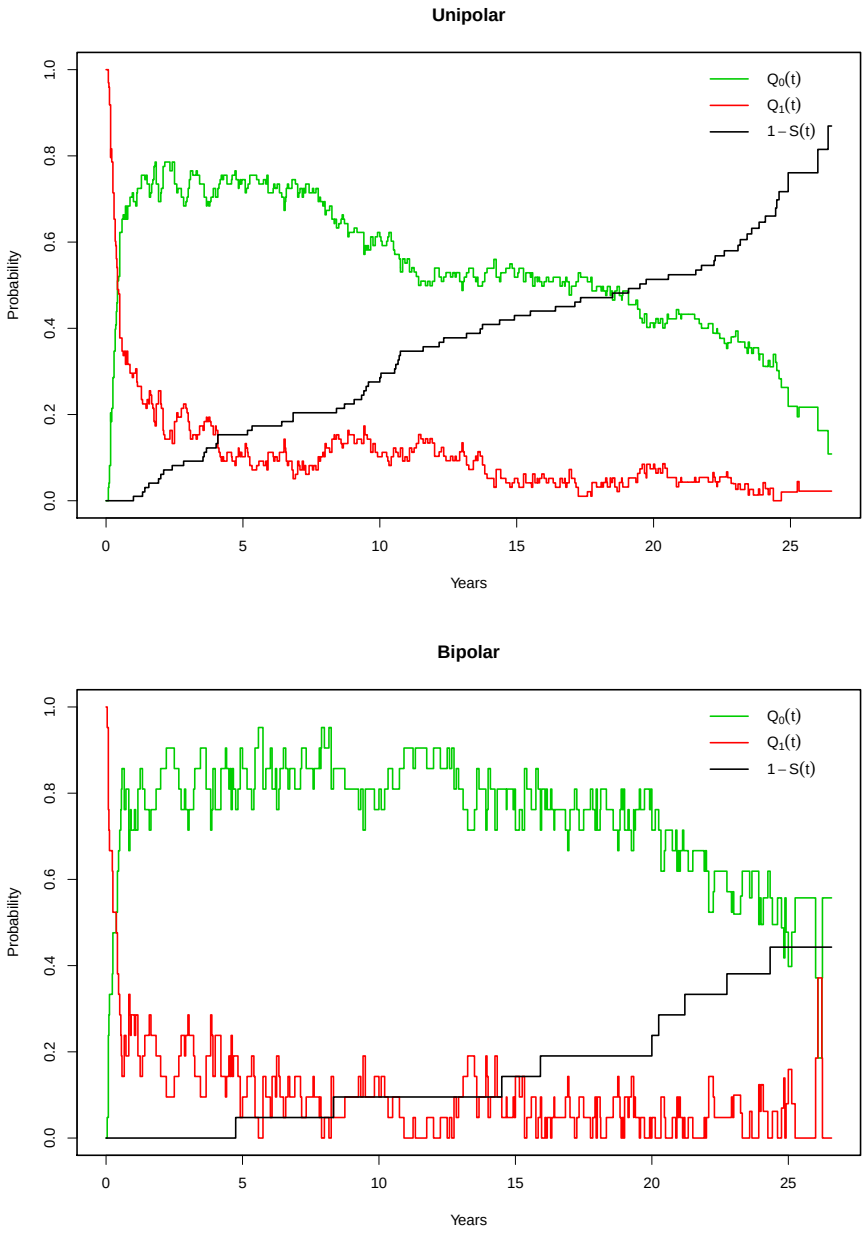

Figure 7: State occupation probabilities for unipolar and bipolar patients.

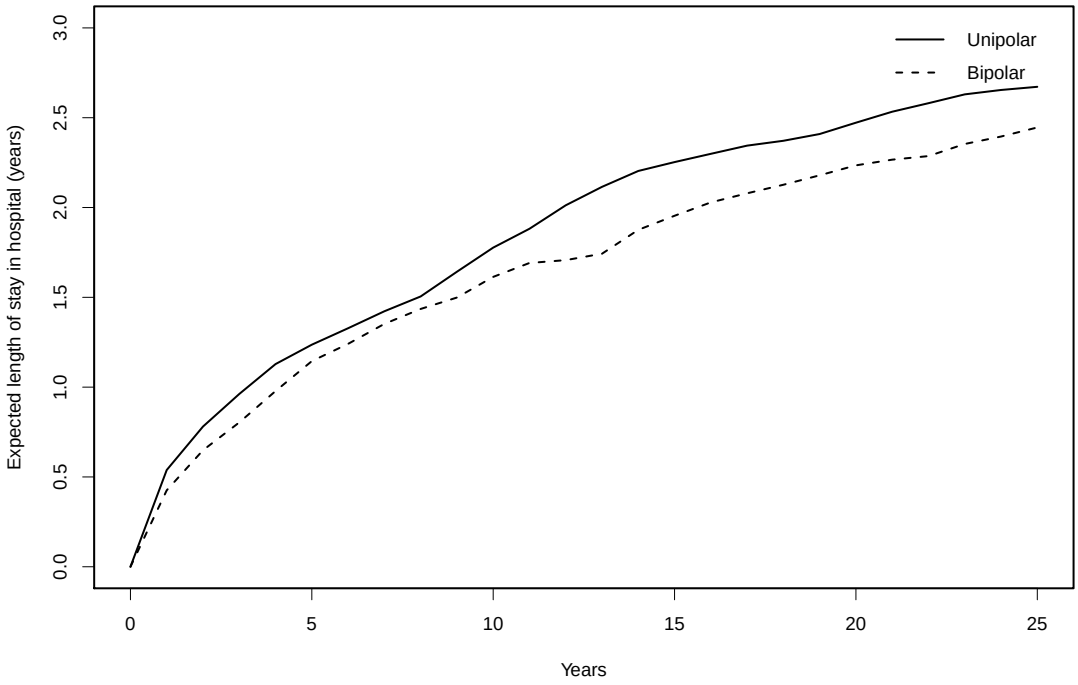

Figure 8: Expected length of stay in hospital for unipolar and bipolar patients.

Parameter	Link	Bip. vs. unip.		Adjusted*	
		Coeff.	(SD)	Coeff.	(SD)
$\mu(10)$	log	0.404	(0.200)	0.330	(0.212)
$S(10)$	cloglog	-1.213	(0.732)	-1.123	(0.913)
$\int_0^{10} Q_1(u)du$	identity	-0.163	(0.312)	-0.109	(0.327)

Table 2: Analysis of pseudo-observations at 10 years. (*: for age at diagnosis and sex)

5 Discussion

We have studied models for recurrent events focussing on marginal models for the mean number of events. We demonstrated that the estimator suggested by Cook and Lawless (1997) for the situation with competing risks works well even when there are times during which subjects are not at risk for new events. We emphasized that when there are competing risks and/or periods where subjects are not at risk then a model for $\mu(t)$ only partially describes the process and analyses of $\mu(t)$ should be supplemented by studies of the mortality rate and/or time not-at-risk. In our example concerning psychiatric admissions such supplementary analyses were conducted using pseudo-observations.

Marginal models have the nice feature that dependence between successive events need not be specified. This is in contrast to intensity-based models where such dependencies are typically modelled using time-dependent covariates or by using random effects models. In fact, such 'frailty models', e.g. negative binomial models with a gamma distributed frailty, seem to be quite frequently used in medical applications (e.g., Bulsara et al., 2004; Rogers et al., 2014). However when there are competing risks such frailty models may be less appealing since also a specification of how frailty affects the mortality rates is needed. Such models have been studied, e.g. by Huang and Wang (2004) and Rondeau et al. (2007) and some versions are implemented in the `R` package `frailtypack` (Rondeau et al., 2012).

In survival analysis, *left-truncation* is typically easily dealt with in hazard-based models. However, for intensity-based models for recurrent events, left-truncation may impose some difficulties - both when conditioning on past events as time-dependent covariates and for frailty models where the conditional frailty distribution given the past is needed. In both cases, the necessary information on the past may be unavailable at the time of delayed entry. For marginal models, however, the estimating equations (3) stay unbiased under left-truncation as long as the times of delayed entry are independent of the recurrent events process and of survival times.

Marginal models are most simple when independent censoring can be assumed. Otherwise, models for censoring are needed to create weights to be used for making the estimating equations unbiased.

It should be mentioned that an alternative approach to the analysis of recurrent events with a terminal event, again in analogy with methods used in practice for competing risks, would be to define a *composite end-point* by combining information for $N(t)$ and T (e.g. Rogers et al., 2014), possibly by attaching different weights to terminal and non-terminal events (e.g., Mao and Lin, 2016). Such an approach, however, entails some new challenges (tackled by Mao and Lin, 2016) owing to the fact that distinction needs to be made between the two types of events when evaluating the risk set.

References

[Aalen et al., 2004] Aalen, O. O., Fosen, J., Weedon-Fekjaer, H., and Borgan, Ø. (2004). Dynamic analysis of multivariate failure time data. *Biometrics*, 60:764–773.

[Andersen, 2013] Andersen, P. K. (2013). Decomposition of number of years lost according to causes of death. *Statistics in Medicine*, 32:5278–5285.

[Andersen et al., 1993] Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.

[Andersen and Gill, 1982] Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.*, 10:1100–1120.

[Andersen and Perme, 2010] Andersen, P. K. and Perme, M. P. (2010). Pseudo-observations in survival analysis. *Statist. Meth. Med. Res.*, 19:71–99.

[Angst et al., 2003] Angst, J., Gamma, A., Selaro, R., Lavori, P. W., and Zhang, H. (2003). Recurrence of bipolar disorders and major depression. *Eur. Arch. Psychiatry Clin. Neurosci.*, 253:236–240.

[Binder et al., 2014] Binder, N., Gerds, T. A., and Andersen, P. K. (2014). Pseudo-observations for competing risks with covariate dependent censoring. *Lifetime Data Analysis*, 20:303–315.

[Bulsara et al., 2004] Bulsara, M. K., Holman, C. D. J., Davis, E. A., and Jones, T. W. (2004). Evaluating risk factors associated with severe hypoglycaemia in epidemiology studies - what method should we use? *Diabetic Medicine*, 21:914–919.

[Byar, 1980] Byar, D. P. (1980). The Veterans Administrations study of chemoprophylaxis for recurrent stage I bladder tumors: comparisons of placebo, pyridoxine, and topical thiotepa. In Pavone-Macaluso, M., Smith, P. H., and Edsmyr, P., editors, *Bladder Tumors and Other Topics in Urological Oncology*, pages 363–370. Plenum, NewYork.

[Chiang, 1980] Chiang, C.-L. (1980). *An Introduction to Stochastic Processes and their Applications*. Krieger, New York.

[Cook and Lawless, 1997] Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statist. in Med.*, 16:911–924.

[Cook and Lawless, 2007] Cook, R. J. and Lawless, J. F. (2007). *The Statistical Analysis of Recurrent Events*. Springer, New York.

[Cook et al., 2009] Cook, R. J., Lawless, J. F., Lakhal-Chaieb, L., and Lee, K.-A. (2009). Robust estimation of mean functions and treatment effects for recurrent events under event-dependent censoring and termination: Application to skeletal complications in cancer metastatic to bone. *J. Amer. Statist. Assoc.*, 104:60–75.

[Fosen et al., 2006] Fosen, J., Borgan, Ø., Weedon-Fekjaer, H., and Aalen, O. O. (2006). Dynamic analysis of recurrent event data using the additive hazard model. *Biometrical J.*, 48:381–398.

[Fuchs et al., 1994] Fuchs, H. J., Borowitz, D. S., Christiansen, D. H., Morris, E. M., Nash, M. L., Ramsey, B. W., Rosenstein, B. J., Smith, A. L., and Wohl, M. E. (1994). Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *New Engl. J. Med.*, 331:637–642.

[Ghosh and Lin, 2000] Ghosh, D. and Lin, D. Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics*, 56:554–562.

[Ghosh and Lin, 2002] Ghosh, D. and Lin, D. Y. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica*, 12:663–688.

[Grand and Putter, 2016] Grand, M. K. and Putter, H. (2016). Regression models for expected length of stay. *Statist. in Med.*, 35:1178–1192.

[Graw et al., 2009] Graw, F., Gerds, T. A., and Schumacher, M. (2009). On pseudo-values for regression analysis in competing risks models. *Lifetime Data Analysis*, 15:241–255.

[Hougaard, 2000] Hougaard, P. (2000). *Analysis of Multivariate Survival Data.* Springer, New York.

[Hu et al., 2011] Hu, X. J., Lorenzi, M., Spinelli, J. J., Ying, S. C., and McBride, M. L. (2011). Analysis of recurrent events with non-negligible event duration, with application to assessing hospital utilization. *Lifetime Data Analysis*, 17:215–233.

[Huang and Wang, 2004] Huang, C. and Wang, M. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *J. Amer. Statist. Assoc.*, 99:1153–1165.

[Iacobelli and Carstensen, 2013] Iacobelli, S. and Carstensen, B. (2013). Multiple time scales in multi-state models. *Statist. in Med.*, 30:5315–5327.

[Jacobsen and Martinussen, 2016] Jacobsen, M. and Martinussen, T. (2016). A note on the large sample properties of estimators based on generalized linear models for correlated pseudo-observations. *Scand. J. Statist.*, 43:845–862.

[Kessing et al., 2004] Kessing, L. V., Hansen, M. G., Andersen, P. K., and Angst, J. (2004). The predictive effect of episodes on the risk of recurrence in depressive and bipolar disorder - a life-long perspective. *Acta Psych. Scand.*, 109:339–344.

[Kessing et al., 1999] Kessing, L. V., Olsen, E. W., and Andersen, P. K. (1999). Recurrence in affective disorder: Analyses with frailty models. *Amer. J. Epidemiol.*, 149:404–411.

[Latouche et al., 2013] Latouche, A., Allignol, A., Beyersmann, J., Labopin, M., and Fine, J. P. (2013). A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J. Clin. Epidemiol.*, 66:648–653.

[Lawless and Nadeau, 1995] Lawless, J. F. and Nadeau, J. C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37:158–168.

[Lin et al., 2000] Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. Roy. Statist. Soc. ser. B*, 62:711–730.

[Mao and Lin, 2016] Mao, L. and Lin, D. Y. (2016). Semiparametric regression for the weighted composite endpoint of recurrent and terminal events. *Biostatistics*, 17:390–403.

[Mitton et al., 2000] Mitton, L., Sutherland, H., Week, M., and (eds.) (2000). *Microsimulation Modelling for Policy Analysis. Challenges and Innovations.* Cambridge University Press, Cambridge.

[Overgaard et al., ress] Overgaard, M., Parner, E. T., and Pedersen, J. (2017, in press). Asymptotic theory of generalized estimating equations based on jack-knife pseudo-observations. *Ann. Statist.*

[Prentice et al., 1978] Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 34:541–554.

[Prentice et al., 1981] Prentice, R. L., Williams, B. J., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika*, 68:373–379.

[Rogers et al., 2014] Rogers, J. K., Pocock, S. J., McMurray, J. J. V., Granger, C. B., Michelson, E. J., Ostergren, J., Pfeffer, M. A., Solomon, S. D., Swedberg, K., and Yusuf, S. (2014). Analysing recurrent hospitalizations in heart failure: a review of ststistical methodology, with application to CHARM-preserved. *Eur. J. Heart Failure*, 16:33–40.

[Rondeau et al., 2007] Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V., and Soubeyran, P. (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: Application on cancer events. *Biostatistics*, 8:708–721.

[Rondeau et al., 2012] Rondeau, V., Mazroui, Y., and Gonzalez, J. R. (2012). frailtypack: An R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *J. Statist. Software*, 47:Issue 4.

[Wei et al., 1989] Wei, L. J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.*, 84:1065–1073.

Research Reports available from Department of Biostatistics

http://www.pubhealth.ku.dk/bs/publikationer

Department of Biostatistics
University of Copenhagen
Øster Farimagsgade 5
P.O. Box 2099
1014 Copenhagen K
Denmark

15/01 Mansourvar, Z., Martinussen, T. & Scheike, T. Semiparametric regression for restricted mean residual life under right censoring.

15/02 Mansourvar, Z., Martinussen, T. & Scheike, T. An additive-multiplicative restricted mean residual life model.

15/03 Mansourvar, Z. & Martinussen, T. Estimation of average causal effect using the restricted mean residual lifetime as effect measure.

15/04 Ekstrøm, C.E., Gerds, T.A., Jensen, A.K. & Brink-Jensen, K. Sequential rank agreement methods for comparison of ranked lists.

15/05 Christensen, K.B., Makransky, G. & Horton, M. Critical Values for Yen's *Q3*: Identification of Local Dependence in the Rasch model using Residual Correlations.

15/06 Müller, M. & Kreiner, S. Item Fit Statistics in Common Software for Rasch Analysis.

16/01 Andersen, P.K. Life years lost among patients with a given disease.

16/02 Strohmaier, S., Haase, N., Wetterslev, J. & Lange, T. A simple to implement algorithm for natural direct and indirect effects in survival studies with a repeatedly measured mediator.

16/03 Andersen, P.K., Syriopoulou, E. & Parner, E. Causal inference in survival analysis using pseudo-observations.

16/04 Paul Blanche, Michael W. Kattan & Thomas A. Gerds. The c-index is not proper for the evaluation of *t*-year predicted risks.

17/01 Liis Starkopf, Mikkel Porsborg Andersen, Thomas Alexander Gerds, Christian Torp-Pedersen & Theis Lange. Comparison of five software solutions to mediation analysis.

17/02 Klemen Pavlič, Torben Martinussen & Per Kragh Andersen. Goodness of fit tests for estimating equations based on pseudo-observations.

17/03 Julie K. Furberg, Christian B. Pipper & Thomas Scheike. Flexible tests for treatment equivalence in a survival setting.

17/04 Luise Cederkvist, Klaus K. Holst, Klaus K. Andersen & Thomas H. Scheike. Modelling the cumulative incidence function of multivariate competing risks data allowing for within-cluster dependence of risk and timing.

17/05 Cristina Boschini, Klaus K. Andersen & Thomas H. Scheike. Excess risk estimation for matched cohort survival data.

17/06 Paul Blanche. Confidence intervals for the cumulative incidence function via constrained NPMLE.

18/01 Per Kragh Andersen, Jules Angst & Henrik Ravn. Modeling marginal features in studies of recurrent events in the presence of a terminal event.