

Transition Probability Estimates for Non-Markov Multi-State Models

Andrew C. Titman

July 6, 2015

Abstract

Non-parametric estimation of the transition probabilities in multi-state models is considered for non-Markov processes. Firstly, a generalization of the estimator of Pepe et al, 1991 (Statistics in Medicine) is given for a class of progressive multi-state models based on the difference between Kaplan-Meier estimators. Secondly, a general estimator for progressive or non-progressive models is proposed based upon constructed univariate survival or competing risks processes which retain the Markov property. The properties of the estimators and their associated standard errors are investigated through simulation. The estimators are demonstrated on datasets relating to survival and recurrence in patients with colon cancer and prothrombin levels in liver cirrhosis patients.

1 Introduction

Multi-state models provide a way of describing event history data that are in the form of transition times between a finite set of states. In biostatistical contexts the states may represent the presence, absence or severity of disease or diseases (Andersen and Keiding (2002)). Often death is included as an absorbing state, or states, within the model and the approach can be considered as a generalization of both standard and competing risk survival analysis. In many cases the transition probabilities, representing the probability of occupying a particular state at a given time given occupancy of the same or a different state at an earlier time, are of particular interest.

Under a Markov assumption, the Aalen-Johansen (AJ) estimator (Aalen and Johansen (1978)) provides non-parametric estimates of the transition probabilities, $P_{ij}(s, t) = P(X(t) = j | X(s) = i)$ of a multi-state process, $\{X(t), t > 0\}$, $X(t) \in \{0, 1, \dots, R\}$ observed up to right-censoring. If we are only interested in obtaining the marginal state occupation probabilities, i.e. estimates of quantities of the form $P_{ij}(0, t)$ or $p_j(t) = \sum_i \pi_i P_{ij}(0, t)$ where π_i are the initial state occupancies, then the AJ estimator has been shown to remain consistent for non-Markov models (Datta and Satten (2001, 2002); Glidden (2002)). If instead we are interested in $P_{ij}(s, t)$ for some $s > 0$, the AJ estimator may be biased for a general non-Markov model. An intuitive explanation for this bias is that the estimate of the transition probability depends on Nelson-Aalen estimates of the transition intensities which use the data from all subjects in a particular state at a particular time regardless of those subject's state occupancy at time s .

Pepe et al. (1991) and Pepe (1991) developed a robust estimator of occupancy in the illness state of a three-state illness-death model based on the difference between Kaplan-Meier estimators of survival in the healthy state and overall survival. Meira-Machado et al. (2006) considered non-parametric estimation of the transition probabilities for a non-Markov three-state illness-death model based on multi-time-dimensional Kaplan-Meier estimates. Allignol et al. (2014) constructed a simplified representation of the estimator in

This is the peer reviewed version of the following article:

Titman, A.C. (2015) Transition Probability Estimates for Non-Markov Multi-State Models. *Biometrics*. DOI:10.1111/biom.12349,

which has been published in final form at

<http://onlinelibrary.wiley.com/doi/10.1111/biom.12349/abstract>.

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

terms of the limiting probability of a particular competing risks process. Both estimators have the drawback of requiring that the support of the time to absorption (e.g. death) in the illness-death model is contained within the right-censoring distribution.

In this paper, alternative estimators are proposed for the transition probabilities of non-Markov models. Firstly, a generalisation of the robust estimator of Pepe et al. (1991), applicable to a class of progressive models is proposed. Secondly, a general estimator, similar to that of Allignol et al. (2014) in being based on a constructed competing risks process, is proposed which is applicable to any multi-state model with a defined initiation population, including models where recovery is possible or there are an arbitrary number of absorbing states.

The remainder of the paper is organised as follows. Section 2 presents the conditional Pepe estimator and also illustrates the principle by which more efficient estimates than the estimator of Allignol et al. (2014) can be obtained for the progressive illness-death model. Section 3 details the general estimation procedure for an arbitrary multi-state process. Section 4 uses simulation to compare the performance of the proposed estimator for the illness-death model with those of existing estimators and also to investigate the finite sample properties of the proposed estimator and asymptotic standard errors for a range of models. Section 5 presents data examples to illustrate the proposed estimator, firstly with progressive illness-death model on an example dataset involving patients with colon cancer and subsequently with a reversible illness-death model on data relating to liver cirrhosis. The paper concludes with a discussion.

2 Estimators for the progressive illness-death model

2.1 Conditional Pepe estimator

Suppose we are interested in estimation of the transition probabilities in a progressive three-state illness-death model, $\{X(t), t > 0\}$ where $X(t) \in \{0, 1, 2\}$. Here, state 0 represents a healthy state, state 1 represents illness from which recovery to state 0 is not possible and state 2 represents death which is assumed to be directly reachable from either state 0 or state 1.

A simple robust estimator of the marginal transition probabilities in a progressive three-state illness-death model can be constructed by following the approach of Pepe et al. (1991) of using the difference between two Kaplan-Meier estimates. Let T_0 be the time of exit from state 0 and T the time of entry into state 2. Also suppose that subjects are independently right-censored at a time C , if $C < T$.

If we are interested in estimating $P_{01}(s, t)$, we can consider the subset of patients for which $T_0 > s$ and $C > s$, i.e. those under observation at time s and still in state 0. We can identify survival processes; time to first exit from state 0 and time to entry into state 2. Define ${}_sN_0(t)$ to be the counting process of exits from state 0 among subjects in state 0 at time s and let ${}_s\bar{Y}_0(t)$ be the number at risk of exit from state 0 at time t among subjects in state 0 at time s . Similarly let ${}_sN_1(t)$ be the counting process of entries into state 2 among subjects in state 0 at time s and let ${}_s\bar{Y}_1(t)$ be the number under observation and not in state 2 among subjects in state 0 at time s . A conditional Pepe-type estimator of the transition probability is then given by

$$\begin{aligned} \hat{P}_{01}(s, t) &= \hat{S}_1(t) - \hat{S}_0(t) \\ &= \prod_{v \in [s, t)} \left(1 - \frac{d_s N_1(v)}{{}_s\bar{Y}_1(v)} \right) - \prod_{v \in [s, t)} \left(1 - \frac{d_s N_0(v)}{{}_s\bar{Y}_0(v)} \right). \end{aligned}$$

Note that this estimator has the same form as that of the estimator of the marginal state occupancy probabilities in Pepe et al. (1991) except that only patients still healthy at time s are used in the estimation. It is also the estimator recently independently proposed by Uña-Álvarez and Meira-Machado (2015).

The estimator depends on two correlated Markov processes. As a consequence, standard martingale based methods for obtaining variance estimates are not applicable. Uña-Álvarez and Meira-Machado (2015) used a simple bootstrap to obtain variance estimates. However, asymptotic estimates are also possible using moment-type variance estimators as developed by Pepe (1991). The form of these variance estimates for \hat{P}_{01} in the three-state illness-death model is given in Appendix A1.

Estimators of the transition probabilities from more general progressive multi-state models can also be constructed from Pepe-type estimates. Uña-Álvarez and Meira-Machado (2015) showed that to estimate the transition probability from state i to state j it is necessary to define the set of states \mathcal{X}_j which are reachable from i and from which j is reachable. An estimator for $P_{ij}(s, t)$ can then be found by considering the difference between Kaplan-Meier estimates of the sojourn distribution in \mathcal{X}_j and the sojourn distribution in $\mathcal{X}_j \cup \{j\}$.

2.2 Allignol *et al* estimator

Allignol *et al.* (2014) also considered the estimation of transition probabilities in a progressive three-state illness-death model. Their approach involves constructing a competing risks process defined as

$$\kappa_s(u) = \begin{cases} 0 & \text{if } X(u) \in \{0, 1\} \\ 1 & \text{if } X(u) = 2 \text{ and } I(s < T_0 \leq t, t < T) = 1 \\ 2 & \text{if } X(u) = 2 \text{ and } I(s < T_0 \leq t, t < T) = 0, \end{cases} \quad (1)$$

note that the times s and t are fixed and a separate process needs to be constructed for each distinct pair of times (s, t) .

Since this competing risks process involves at most one event for each patient, it is trivially Markov and hence the AJ estimator can be used to estimate the transition probabilities of this process. Moreover, we have that

$$P_{01}(s, t) = \lim_{u \rightarrow \infty} P(\kappa_s(u) = 1)$$

and therefore $\hat{P}_{01}(s, t)$ may be taken to be the estimate of $\lim_{u \rightarrow \infty} P(\kappa_s(u) = 1)$.

A disadvantage of this construction is that, in order to identify this limiting probability, it is necessary for the support of T , the times of absorption, to be contained within the support of right-censoring times. This requirement is rather restrictive. In particular, it would preclude any study of a fixed length shorter than the maximum possible lifetime. It is also a much stronger requirement than is necessary for the AJ estimator, which only requires the time of interest, t , to be contained in the support of right-censoring times.

A further drawback of the construction is that no distinction can be made between patients who are censored before reaching death, regardless of whether the condition $I(s < T_0 \leq t < T)$ is met or not, since in either case the subject will be censored as $\kappa_s = 0$. The extent of information lost due to the need for subjects to reach the absorbing state will depend on the censoring rate and the particular t of interest.

Regardless of assumptions about the nature (Markov or non-Markov) of $X(t)$, the transition probability $P_{01}(s, t)$ has no dependence on the underlying transition intensities,

$$\lambda_{ij}(u; \mathcal{F}_u) = \lim_{\delta u \downarrow 0} \frac{P(X(u + \delta u) = j | X(u) = i, \mathcal{F}_u)}{\delta u},$$

where \mathcal{F}_u is the history or *filtration* of $X(t)$ up to but not including time u , for times $u > t$. As a consequence we can construct a process $X^*(u)$ for which the transition intensities, $\lambda_{ij}^*(u; \mathcal{F}_u)$ have the property $\lambda_{i2}^*(u; \mathcal{F}_u) = \lambda_{i2}(u; \mathcal{F}_u)$ for $u \leq t$ and $\lambda_{i2}^*(u; \mathcal{F}_u) > \lambda_{i2}(u; \mathcal{F}_u)$ for $u > t$.

Note that, by construction, $P_{01}^*(s, t) = P(X^*(t) = 1 | X^*(s) = 0) = P_{01}(s, t)$. However, for identical censoring distributions, the estimate of $P_{01}^*(s, t)$ via $\kappa_s^*(u)$ based on $X^*(u)$ will be more efficient than the estimate of $P_{01}(s, t)$ for $X(u)$ because a greater proportion of subjects will reach the absorbing state. Moreover, the most efficient estimate will occur for a process, $X^*(t)$, that takes arbitrarily large transition intensities, λ_{i2} , at all times $u > t$, which is equivalent to assuming all patients reach the absorbing state at time $t+$. The estimator developed in the remainder of this paper works on the basis that estimates of the transition probabilities can be obtained through the behaviour of the process up to $t-$ and information on the observed state occupancy of subjects at risk at time t .

3 Proposed estimator

In this section an estimator for the transition probability between states, or sets of states, in a general multi-state model is developed.

Let $\{X(t), t > 0\}$ be a multi-state process where $X(t) \in \mathcal{S} = \{1, 2, \dots, R\}$. Suppose that the set of inadmissible direct transitions, i.e. the (i, j) such that $\lambda_{ij}(t; \mathcal{F}_t) = 0$ for all t within the range of times of interest $(0, \tau)$ and all possible histories, \mathcal{F}_t , is taken to be known. For additional generality, suppose we are interested in estimation of $P_{\mathcal{I}\mathcal{J}}(s, t) = P(X(t) \in \mathcal{J} | X(s) \in \mathcal{I})$, where \mathcal{I} and \mathcal{J} are non-empty, but not necessarily disjoint, subsets of \mathcal{S} .

We can identify two important sets of states:

- $\mathcal{R}_{\mathcal{J}}$ - the set of states reachable from \mathcal{I} from which \mathcal{J} cannot be reached.
- $\mathcal{A}_{\mathcal{J}}$ - the set of states from which entry at time t implies $X(t') \in \mathcal{J}$ for all $t' \geq t$.

$\mathcal{A}_{\mathcal{J}}$ will equal \mathcal{J} if \mathcal{J} consists of a single, or multiple, absorbing states, or if \mathcal{J} consists of states from which it is only possible to enter other states in \mathcal{J} . In contrast if, for instance, \mathcal{J} consists of an illness state from which recovery is possible and a death state, then $\mathcal{A}_{\mathcal{J}}$ will include only the death state.

Having defined these sets we can then define a competing risks process, $\{Z_s(u), u \geq s\}$, as follows:

$$Z_s(u) = \begin{cases} 0 & \text{if } X(u) \notin \{\mathcal{R}_{\mathcal{J}} \cup \mathcal{A}_{\mathcal{J}}\} \\ 1 & \text{if } X(u) \in \mathcal{A}_{\mathcal{J}} \\ 2 & \text{if } X(u) \in \mathcal{R}_{\mathcal{J}}. \end{cases}$$

This process can be related back to the probability of interest through the identity

$$P(X(t) \in \mathcal{J} | X(s) \in \mathcal{I}) = P(Z_s(t) = 1 | X(s) \in \mathcal{I}) + P(Z_s(t) = 0 | X(s) \in \mathcal{I})P(X(t) \in \mathcal{J} | Z_s(t) = 0, X(s) \in \mathcal{I}).$$

Suppose we have a sample of subjects, continuously observed up to independent left-truncation and right-censoring. For notational simplicity the subject subscript is omitted. Let ${}_sY(u)$ be the at risk indicator for the process $Z_s(u)$ for a given subject and let $\mathcal{N}_{\mathcal{I}}(s)$ denote the set of subjects satisfying $X(s) \in \mathcal{I}$, ${}_sY(s) = 1$. Moreover, denote ${}_s\bar{Y}(u) = \sum {}_sY(u)$ as the number of subjects at risk for the process $Z_s(u)$ at time u , where the summation here and in subsequent summations is over subjects in $\mathcal{N}_{\mathcal{I}}(s)$.

Under independent left-truncation and right-censoring, but without any other assumptions regarding the process, $P(Z_s(t) = 0 | X(s) \in \mathcal{I})$ and $P(Z_s(t) = 1 | X(s) \in \mathcal{I})$ can be estimated by considering the AJ estimator of the transition probabilities for $Z_s(t)$ amongst the patients in $\mathcal{N}_{\mathcal{I}}(s)$. A natural estimator for $P(X(t) \in \mathcal{J} | Z_s(t) = 0, X(s) \in \mathcal{I})$ is simply the observed proportion of subjects with $X(t) \in \mathcal{J}$ amongst all patients with $Z_s(t) = 0$, ${}_sY(t) = 1$ and $X(s) \in \mathcal{I}$, which we can denote $\hat{p}_{\mathcal{J}|\mathcal{I}}(t)$.

Specifically we denote ${}_sN_1(t)$ to be the counting process of observed events of type 1 in the competing risks model and ${}_sN(t)$ to be the counting process of observed events of all types among the patients in $\mathcal{N}_{\mathcal{I}}(s)$. The relevant quantities of the competing risks process can be estimated by

$$\hat{F}_0(t) = \hat{P}(Z_s(t) = 0 | X(s) \in \mathcal{I}) = \prod_{v \in [s, t]} \left(1 - \frac{d_s N(v)}{{}_s\bar{Y}(v)} \right)$$

and

$$\hat{F}_1(t) = \hat{P}(Z_s(t) = 1 | X(s) \in \mathcal{I}) = \int_s^t \prod_{v \in [s, t]} \left(1 - \frac{d_s N(v)}{{}_s\bar{Y}(v)} \right) \frac{d_s N_1(u)}{{}_s\bar{Y}(u)},$$

while

$$\hat{p}_{\mathcal{J}|\mathcal{I}}(t) = \frac{\sum I(X(s) \in \mathcal{I}, X(t) \in \mathcal{J}, {}_sY(t) = 1)}{\sum I(X(s) \in \mathcal{I}, {}_sY(t) = 1)}.$$

The overall estimator is then defined as

$$\hat{P}_{\mathcal{I}\mathcal{J}}(s, t) = \hat{F}_1(t) + \hat{F}_0(t)\hat{p}_{\mathcal{J}|\mathcal{I}}(t).$$

Note that this estimator is only defined for s such that ${}_s\bar{Y}(s) > 0$, meaning there is at least one patient still under observation satisfying $X(s) \in \mathcal{I}$.

A delta-method based approach for obtaining variance estimates of the proposed estimator is given in the Appendix A2.

3.1 Estimates of transitions between pairs of states

Usually we will only be interested in estimating the transition probabilities between a pair of distinct states i and j . In which case the general formulation simplifies somewhat.

If j is an absorbing state then it follows that $P(X(t) = j|Z_s(t) = 0) = 0$ and hence the estimate of the transition probability is just $\hat{P}_{ij}(s, t) = \hat{F}_1(t)$.

Alternatively, when j is a transient state, the competing risks process reduces to a simple survival process

$$Z_s(t) = \begin{cases} 0 & \text{if } X(t) \notin \{\mathcal{R}_j \cup \mathcal{A}_j\} \\ 1 & \text{if } X(t) \in \mathcal{R}_j \end{cases}.$$

and then

$$P(X(t) = j|X(s) = i) = P(Z_s(t) = 0|X(s) = i)P(X(t) = j|Z_s(t) = 0, X(s) = i),$$

meaning $\hat{P}_{ij}(s, t) = \hat{F}_0(t)\hat{p}_{j|i}(t)$.

When all states within the process are transient such that $\mathcal{R}_{\mathcal{J}} = \emptyset$, no additional information can be inferred from subjects who are censored before the time of interest t . The estimator in this case reduces to the observed proportion of subjects in the set of states \mathcal{J} amongst all patients still under observation at time t with $X(s) = i$.

3.2 Three-state illness-death model

Returning to the progressive three-state illness-death model considered in Section 2, under the proposed method, estimation of $P_{01}(s, t)$ is based on a constructed survival process

$$Z_s(u) = \begin{cases} 0 & \text{if } X(u) \in \{0, 1\} \\ 1 & \text{if } X(u) = 2 \end{cases} \quad (2)$$

and the estimated quantity

$$\hat{P}(X(t) = 1|X(t) \in \{0, 1\}) = \frac{\sum I(X(t) = 1, X(s) = 0, {}_sY(s) = {}_sY(t) = 1)}{\sum I(X(s) = 0, {}_sY(s) = {}_sY(t) = 1)} \quad (3)$$

where we take $\hat{P}_{01}(s, t) = \hat{P}(Z_s(t) = 0|X(s) = 0)\hat{P}(X(t) = 1|Z_s(t) = 0, X(s) = 0)$.

The difference between this estimator and that of Allignol et al. (2014) will be most pronounced for estimates of $P_{01}(s, t)$ when t is small in relation to the typical times-to-death of the process or when there is a maximum follow-up time in the study which is before the time at which all subjects will have reached state 2.

The proposed general estimator differs from the Pepe-type estimator of Section 2.1 only in the way in which the conditional probability of occupancy in state 1 given occupancy in state 0 or 1 is estimated. In the Pepe estimator this is estimated via $1 - \hat{S}_0(t)/\hat{S}_1(t)$ whereas the general estimator uses the moment estimate in (3). Since the Pepe-type estimator uses information from subjects censored before time t to estimate the quantity we would expect some improvement in efficiency. Whether this improved efficiency outweighs the more complicated procedures required for variance estimation will be considered in the simulation study in Section 4.

4 Simulation study

4.1 Illness-death model without recovery

To assess the performances of the proposed general estimator and the Pepe-type estimator and to compare them to the AJ estimator and the Allignol et al. (2014) estimator for realistically sized datasets, data are first simulated from a three-state illness-death model without recovery. Three processes are considered: a Markov process such that the AJ estimator will be consistent; a non-Markov process induced via a shared Gamma frailty and a pathological non-Markov process where future transition intensities depend on the state occupied at a fixed point in time. For the Markov process a time-homogeneous process with intensities

$\alpha_{01} = 0.12, \alpha_{02} = 0.03, \alpha_{12} = 0.1$ is used. The frailty model has the same baseline intensities but the individual intensities for the $0 \rightarrow 1$ and $1 \rightarrow 2$ transitions are of the form $\alpha_{ij}V$ where $V \sim \Gamma(1/2, 1/2)$. The non-Markov process also has the same intensities except that

$$\alpha_{12}(t) = \begin{cases} 0.05 & \text{if } X(4) = 0 \\ 0.1 & \text{if } X(4) \neq 0. \end{cases}$$

For each process, data are generated under two censoring scenarios: independent exponentially distributed right-censoring times, where $C \sim \text{Exp}(0.04)$, and uniformly distributed censoring times where $C \sim U[5, 40]$. Each scenario is also generated with sample sizes of $N = 200$ and $N = 500$ where N is the number of subjects in state 0 at time 0.

Table 1 gives the bias and empirical standard deviation across 1000 simulations for the three estimators for each of the scenarios when assessing the transition probabilities between $\tau_{.15}$ and $\tau_{.45}$, corresponding to the 15th and 45th percentiles of the time-to-absorption distribution T , respectively. Both the Pepe-type and the proposed general estimator give close to unbiased estimates in all the scenarios considered. The bias and standard deviation of the proposed estimator is generally lower than the estimator of Allignol et al. (2014). Note that the estimator of Allignol et al. (2014) exhibits quite considerable bias under the shared Gamma frailty model, even with an exponential censoring distribution. This is presumably due to the heaviness of the tail of the marginal survival distribution of the frailty illness-death process and the higher censoring rate for the scenario considered.

As would be expected, the AJ estimator performs better than the proposed estimator and the Pepe-type estimator when the true process is Markov. In particular, it is more efficient. For instance, in the scenario considered the empirical standard deviation is around 80% that of the proposed estimator. When the process is not Markov, the AJ estimator is biased. The bias is much more marked for the pathological non-Markov process than for the Gamma frailty.

Broadly similar results were found for other transition probability estimates. Web Tables 2 and 3 give results for analogous transition probabilities between times corresponding to the 15th and 30th and 15th and 60th percentiles of T . The Pepe-type estimator performs slightly better than the proposed estimator tending to have comparable empirical bias but slightly lower variance. The apparent improvement in efficiency was more marked for estimates of $P_{01}(\tau_{.15}, \tau_{.60})$, which has a greater proportion of subjects who are censored before the upper time point.

Table 1 also gives the estimated standard error and the resulting empirical coverage of nominal 95% confidence intervals for each of the scenarios and estimators. The Greenwood-type estimates of variance were close to being unbiased and confidence intervals based on them had reasonably close to nominal coverage. The asymptotic variance estimator for the Pepe-type estimator is known to underestimate the true variance to greater extent than standard Greenwood-type estimates (Pepe, 1991). However, in the scenarios considered there was little evidence of this resulting in coverage further from the nominal 95% level than the Greenwood-type estimates.

4.2 Illness-death model with recovery

To assess the performance of the proposed estimator in situations where backward transitions are possible, scenarios based around a three-state illness-death model with recovery are simulated. Once again, three processes are considered: a Markov process; a non-Markov process induced via a shared Gamma frailty and a pathological non-Markov process. For this model, the Pepe-type estimator and Allignol *et al* estimators are not applicable so we only consider a comparison between the proposed general estimator and the AJ estimator.

Table 2 shows the bias, empirical standard deviation, standard error and coverage of 95% confidence intervals for each of the estimators for $P_{01}(\tau_{.15}, \tau_{.45})$. The proposed estimator is shown to be close to unbiased in each scenario considered. In contrast, though not unexpectedly, the Aalen-Johansen estimator exhibits considerable bias. The constructed asymptotic 95% confidence intervals for the proposed estimator were found to have reasonably close to nominal coverage in most scenarios, the exception being in the case of the frailty model for $N = 200$. Similarly results were also found for other transition probabilities. Full results are given in Web Tables 5 and 6.

Table 1: Comparison of estimates of $P_{01}(\tau_{15}, \tau_{45})$ for simulated illness-death without recovery model datasets. Bias, empirical standard deviation (SD) and average standard error (SE) $\times 10^2$. Cov refers to the empirical coverage of nominal 95% confidence intervals.

N	Scenario		Proposed				Pepe type				Allignol <i>et al</i>				Aalen-Johansen			
	Model	Cens	Bias	SD	SE	Cov	Bias	SD	SE	Cov	Bias	SD	SE	Cov	Bias	SD	SE	Cov
200	Markov	Unif	-0.061	4.832	4.819	0.946	-0.256	4.686	4.646	0.943	-2.401	6.718	5.834	0.868	-0.198	4.112	4.137	0.950
200	Markov	Exp	0.129	5.587	5.446	0.941	-0.105	5.377	5.207	0.939	-0.742	6.912	6.475	0.913	-0.045	4.772	4.656	0.935
500	Markov	Unif	0.120	3.086	3.057	0.944	0.057	2.954	2.969	0.953	-2.212	4.137	3.850	0.869	0.029	2.637	2.624	0.949
500	Markov	Exp	0.225	3.354	3.456	0.957	0.082	3.243	3.334	0.958	-0.009	4.408	4.234	0.935	0.143	2.903	2.950	0.953
200	Non-Markov	Unif	0.141	5.317	5.346	0.951	-0.107	5.194	5.135	0.950	-3.239	6.817	6.387	0.871	4.821	4.537	4.582	0.832
200	Non-Markov	Exp	0.260	5.967	6.085	0.945	0.126	5.691	5.788	0.944	-0.628	7.657	7.107	0.911	4.827	5.069	5.185	0.869
500	Non-Markov	Unif	0.239	3.449	3.396	0.943	0.152	3.419	3.290	0.946	-2.488	4.666	4.305	0.866	4.981	2.836	2.904	0.610
500	Non-Markov	Exp	0.182	3.928	3.856	0.938	0.140	3.785	3.713	0.944	-0.361	4.785	4.638	0.928	4.900	3.319	3.281	0.684
200	Frailty	Unif	0.117	3.497	3.525	0.951	0.092	3.452	3.366	0.934	-4.152	4.647	4.085	0.706	-0.132	3.082	3.050	0.942
200	Frailty	Exp	0.021	4.020	3.977	0.943	-0.054	3.847	3.761	0.939	-1.447	6.225	5.078	0.831	-0.198	3.507	3.411	0.937
500	Frailty	Unif	0.186	2.287	2.244	0.942	0.124	2.184	2.150	0.939	-3.768	3.438	2.830	0.610	-0.069	1.981	1.935	0.933
500	Frailty	Exp	-0.007	2.554	2.529	0.934	-0.050	2.461	2.402	0.933	-1.081	4.384	3.592	0.822	-0.245	2.215	2.16	0.923

Table 2: Comparison of estimates of $P_{01}(\tau_{15}, \tau_{45})$ for simulated illness-death with recovery model datasets. Bias and empirical standard deviation $\times 10^2$

N	Scenario		Proposed				Aalen-Johansen			
	Model	Cens	Bias	SD	SE	Cov	Bias	SD	SE	Cov
200	Markov	Unif	0.077	4.633	4.754	0.947	0.065	3.358	3.469	0.953
200	Markov	Exp	0.016	5.574	5.265	0.928	-0.034	3.854	3.844	0.935
500	Markov	Unif	0.199	2.975	3.020	0.951	0.091	2.164	2.199	0.956
500	Markov	Exp	0.185	3.394	3.355	0.944	0.092	2.502	2.443	0.942
200	Non-Markov	Unif	-0.036	5.006	4.827	0.936	-2.169	3.644	3.603	0.885
200	Non-Markov	Exp	0.030	5.508	5.364	0.948	-2.020	4.100	4.009	0.905
500	Non-Markov	Unif	-0.073	3.217	3.067	0.940	-2.154	2.377	2.287	0.816
500	Non-Markov	Exp	0.050	3.380	3.406	0.950	-2.044	2.483	2.540	0.858
200	Frailty	Unif	0.125	5.371	5.270	0.943	8.388	3.647	3.789	0.393
200	Frailty	Exp	-0.182	5.907	5.826	0.927	8.226	4.225	4.199	0.518
500	Frailty	Unif	0.071	3.210	3.350	0.956	8.389	2.302	2.400	0.043
500	Frailty	Exp	0.088	3.783	3.733	0.939	8.468	2.573	2.672	0.096

5 Data examples

5.1 Example 1: Adjuvant chemotherapy for colon cancer

As a first example, the proposed estimator is applied to data relating to patients with colon cancer. The data originate from a trial assessing the efficacy of different forms of adjuvant chemotherapy. Here the subset of 315 patients who were randomised to observation is considered. Note that, considering the full dataset without adjusting for treatment may result in an apparent lack of Markovianity.

Patient follow-up was between 5 and 9 years. The date of recurrence of colon cancer was known if it occurred within the follow-up period. A three-state progressive illness-death model can be formulated by considering pre-recurrence, post-recurrence and death states. The data have a relatively high censoring rate. 168 patients died within the follow-up period, of whom 13 died before recurrence. 125 were censored after recurrence, with the remaining 22 being censored before recurrence.

Figure 1 shows estimates of $P_{01}(365, t)$ for the Aalen-Johansen, Allignol, conditional Pepe and general non-Markov estimates, where t is measured in days. There is close agreement between the proposed general estimator and the Pepe-type estimator, with divergence only at late follow-up times where the censoring rate is high. There is considerable disagreement between these estimates and the Allignol estimator. Note that unlike the example considered in Allignol et al. (2014), the colon cancer dataset has a relatively high censoring rate (49% reach the absorbing state) and it is not reasonable to assume the support of the censoring distribution contains the support of the absorption time distribution. There is some disagreement between the AJ estimator and the non-Markov estimates with the AJ estimate having a lower peak. However, there is relatively little evidence against the Markov assumption. This can also be verified by considering the Kendall's τ based tests of Markovianity proposed by Rodríguez-Girondo and Uña-Álvarez (2012) in which estimated values of τ different from zero indicate departures from a Markov process. In this case, neither the local test at $t = 365$ ($|\hat{\tau}| = 0.12, p = 0.17$) nor the global supremum based test ($|\hat{\tau}| = 0.21, p = 0.28$) suggest a significant departure.

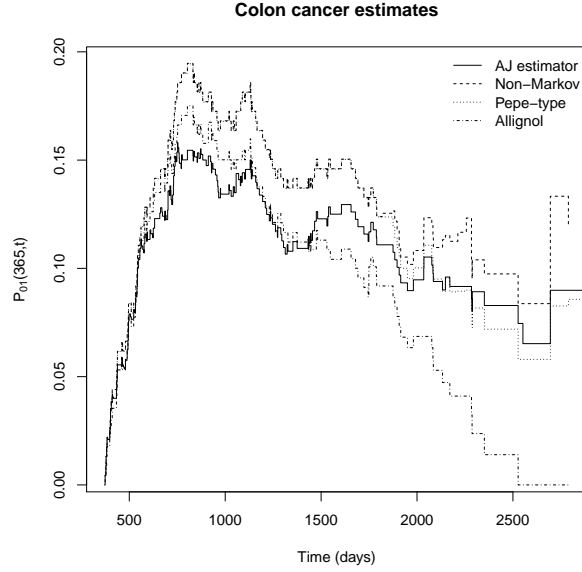
Web Figure 1 shows the estimated standard errors. The standard errors of the estimates are generally lower for the AJ estimator, as might be expected. For the conditional Pepe and proposed non-Markov estimators the standard errors are very similar until around 5 years, with the conditional Pepe estimator giving slightly lower standard errors after this time. The Allignol *et al* estimator's standard errors are also broadly comparable up to 5 years but thereafter decrease to zero because $\hat{P}_{01}(365, t)$ goes to zero.

5.2 Example 2: Liver cirrhosis data

In the second example, we consider the liver cirrhosis data set considered by Andersen et al. (1993) and subsequently by many other authors. The data are available in the R package *mstate* de Wreede et al. (2011). The data consist of 488 patients from a randomized clinical trial and consider abnormal prothrombin levels as a reversible illness state in an illness-death model with recovery.

Patients were randomized to either to a treatment of prednisone (251 patients) or a placebo (237 patients).

Figure 1: Comparison of estimated transition probabilities from the no recurrence to recurrence states in the colon cancer dataset



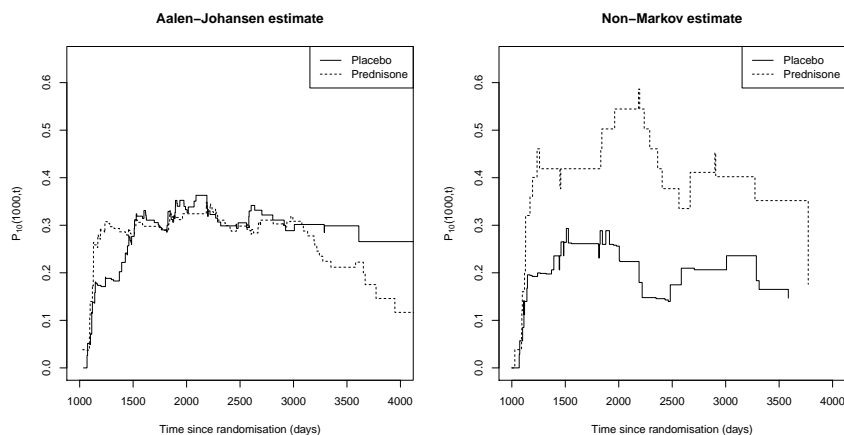
The natural time scale in the model is time since randomization. Patients can enter the study with either normal or elevated prothrombin levels. 51% of patients died before the end of follow-up, with the remainder right-censored. We wish to compare estimates of $P_{10}(s, t)$, representing the transition probabilities from elevated to normal prothrombin levels, using the AJ estimator which makes a Markov assumption, and the proposed estimator. Note that, since the model allows recovery from the illness state, the constructed Markov process has $\mathcal{A}_{\mathcal{J}} = \emptyset$ and $\mathcal{R}_{\mathcal{J}} = \{2\}$.

Estimates of the transition probabilities via the AJ estimator are performed using the R package *etm* (Allignol et al. (2013)). Estimates of $P_{10}(1000, t)$ are produced for each trial arm via an Aalen-Johansen estimator and also by the proposed non-Markov estimator and given in Figure 2. A starting point of 1000 days is chosen, in the absence of any specific clinical milestone, because it roughly coincides with the median follow-up time in the dataset. From this particular starting point, the Aalen-Johansen estimates suggest little difference between placebo and prednisone. In contrast, the estimates based on the non-Markov estimate suggest a reasonably marked difference between the two treatments, with a higher proportion of patients on prednisone returning to normal levels of prothrombin. The conclusions about the efficacy of continued treatment beyond 1000 days may therefore be different if the non-Markov estimate is considered compared to the AJ estimator. The discrepancy between the estimates may be due to unobserved heterogeneity in the data. Note that, even if the process is conditionally Markov given some covariates, the marginal process will be non-Markov if those covariates are unobserved. Some patients make substantially more transitions between the normal and elevated prothrombin states than expected.

6 Discussion

In the case of a progressive three-state illness-death model, both the conditional Pepe estimator and the proposed general estimator are more efficient than the Allignol *et al* estimator and remain unbiased even if the support of the censoring distribution does not encompass the support of the time-to-absorption distribution. The Pepe-type estimator is slightly more efficient than the general estimator, but the simpler form of variance estimates for the general estimator may make it more attractive in practice. The general estimator also has the advantage of offering a single form of estimator that is applicable to a wide range of multi-state models

Figure 2: Estimated transition probabilities from elevated to normal prothrombin levels for placebo and prednisone groups using the Aalen-Johansen estimator (left-panel) and proposed non-Markov estimator (right-panel).



including models permitting recovery.

An obvious extension to the proposed methods would be the incorporation of regression models for the effect of explanatory variables. It is not possible to use a traditional approach of regression of the individual transition intensities. However, methods based upon pseudo-observations of quantities of interest (Andersen et al. (2003)) should be applicable.

Only an informal indication of the sensitivity to the Markov assumption can be obtained by comparing the AJ and non-Markov transition probability estimates. In particular, the discrepancy may be dependent on an often arbitrary choice of starting time, s . It would therefore be useful to devise formal tests of global Markovianity, similar to those of Rodríguez-Girondo and Uña-Álvarez (2012), for the case of processes with recovery.

The proposed estimator assumes that the left-truncation and right-censoring times are independent of the multi-state process $X(t)$. In situations where censoring is dependent on the current state or the history of the multi-state process, the estimates will be biased. A further investigation of the impact of state dependent censoring in the case of the progressive illness-death model is given in the Web Appendix which indicates that all the non-Markov estimators are to some extent biased, with the Pepe-type estimates appearing more robust in the cases considered. It may be possible to obtain unbiased estimates in this situation by adopting an inverse-probability-of-censoring weights approach, similar to that used by Datta and Satten (2002) for the state occupancy probabilities.

Supplementary materials

Web Appendices, Tables, and Figures referenced in Sections 4.1, 4.2, 5.1 and 6 are available with this paper at the Biometrics website on Wiley Online Library.

Appendix

A1: Variance estimation for the conditional Pepe-estimator

Following Pepe (1991), a moment-type estimator can be obtained for the variance of the conditional Pepe estimator of $P_{01}(s, t)$ for a progressive three-state illness-death model. Let n_s be the number of patients

under observation and in state 0 at time s . For the conditional Pepe-estimator

$$\hat{P}_{01}(s, t) = \hat{S}_1(t) - \hat{S}_0(t)$$

defined in Section 2.1, the asymptotic variance of

$$\sqrt{n_s}(\hat{P}_{01}(s, t) - P_{01}(s, t))$$

is given by

$$\frac{1}{n_s} \sum_{j=1}^{n_s} \left(S_1(t) \left\{ \int_0^t \frac{1}{y_1(u)} dN_1^j(u) - \int_0^t \frac{Y_1^j(u)}{y_1(u)} \lambda_1(u) du \right\} - S_0(t) \left\{ \int_0^t \frac{1}{y_0(u)} dN_0^j(u) - \int_0^t \frac{Y_0^j(u)}{y_0(u)} \lambda_0(u) du \right\} \right)^2$$

where $N_k^j(t)$ is the counting process for event k for the j th subject and $Y_k^j(t)$ is the corresponding at risk indicator,

$$y_k(t) = P(\text{observed at risk for an event of type } k \text{ just before } t)$$

and $\lambda_k(t) = \lim_{\delta \downarrow 0} P(T_k \in (t, t + \delta) | T_k \geq t) / \delta$ for $k = 0, 1$. It is assumed that $y_k(u) > 0, s \leq u \leq t$. A consistent-moment type variance estimator for $\hat{P}_{01}(s, t)$ is then

$$\sum_{j=1}^{n_s} \left(\hat{S}_1(t) \left\{ \int_0^t \frac{1}{\hat{y}_1(u)} dN_1^j(u) - \int_0^t \frac{Y_1^j(u)}{\hat{y}_1(u)^2} dN_1(u) \right\} - \hat{S}_0(t) \left\{ \int_0^t \frac{1}{\hat{y}_0(u)} dN_0^j(u) - \int_0^t \frac{Y_0^j(u)}{\hat{y}_0(u)^2} dN_0(u) \right\} \right)^2$$

where $N_k(u) = \sum_{j=1}^{n_s} N_k^j(u)$ and $\hat{y}_k(u) = Y_k(u) / n_s, k = 0, 1$.

A2: Variance estimation for the general estimator

In its most general form, the proposed estimator can be written in terms of a competing risks counting process and a binary random variable. Let ${}_s Y(t)$ be the subject's at-risk indicator and let $V(t) = I(X(t) \in \mathcal{J}) {}_s Y(t)$. We can express $\hat{p}_{\mathcal{J}|\mathcal{I}}(t) = \sum V(t) / \sum {}_s Y(t)$. Note that $V(t)$ is independent of the filtration of the competing risks counting processes up to time t conditional on ${}_s Y(t)$. As a consequence, $\hat{p}_{\mathcal{J}|\mathcal{I}}$ is independent of the cumulative incidence function estimates $(\hat{F}_1(t), \hat{F}_0(t))$. Standard theory for the AJ estimator of cumulative incidence functions (Andersen et al. (1993)) can be used to obtain an estimate of the variance-covariance matrix of $(\hat{F}_1(t), \hat{F}_0(t))$,

$$\mathbf{\Sigma}_t = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}.$$

In addition, $\hat{p}_{\mathcal{J}|\mathcal{I}}$ may be considered as an independent binomial proportion, with estimated variance $\hat{\sigma}_{\mathcal{J}} = \frac{1}{{}_s Y(t)} \hat{p}_{\mathcal{J}|\mathcal{I}}(1 - \hat{p}_{\mathcal{J}|\mathcal{I}})$. Using a delta method argument, an asymptotic approximation to the variance of the overall estimator can then be obtained via

$$\text{Var}(\hat{F}_1(t) + \hat{p}_{\mathcal{J}|\mathcal{I}} \hat{F}_0(t)) = \Sigma_{11} + 2\hat{p}_{\mathcal{J}|\mathcal{I}} \Sigma_{21} + \hat{p}_{\mathcal{J}|\mathcal{I}}^2 \Sigma_{22} + \hat{\sigma}_{\mathcal{J}} \hat{F}_0^2.$$

References

- Aalen, O.O, and Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics* **5**, 141-150.
- Allignol, A., Schumacher, M. and Beyersmann, J. (2011). Empirical Transition Matrix of Multi-State Models: The etm Package. *Journal of Statistical Software*, **38(4)**, 1-15.

- Allignol, A., Beyersmann, J., Gerds, T. and Latouche, A. (2014). A competing risks approach for nonparametric estimation of transition probabilities in a non-Markov illness-death model. *Lifetime Data Analysis* **20**, 495-513.
- Andersen, P.K., Borgan, Ø., Gill, R. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer, New York.
- Andersen, P.K., and Keiding, N. (2002). Multi-state models for event history analysis. *Statistical Methods in Medical Research* **11**, 91-115.
- Andersen, P.K., Klein, J.P., Rosthøj, S. (2003). Generalised linear models for correlated pseudoobservations, with applications to multistate models. *Biometrika*. **90**, 15-27.
- Beyersmann, J., Allignol, A., and Schumacher, M. (2012). *Competing risks and multistate models with R*, Springer, New York.
- Datta, S. and Satten, G.A. (2001). Validity of the Aalen-Johansen estimators of stage occupancy probabilities and Nelson-Aalen estimators of integrated transition hazards for non-Markov models. *Statistics and Probability Letters* **55**, 403-411.
- Datta, S. and Satten, G.A. (2002). Estimation of integrated transition hazards and stage occupancy probabilities for non-Markov systems under dependent censoring. *Biometrics* **58**, 797-802.
- Glidden, D. (2002). Robust inference for event probabilities with non-Markov data *Biometrics*. **58**, 361-368.
- Meira-Machado, L., de Uña-Álvarez, J. and Cadarso-Suárez, C. (2006). Nonparametric estimation of transition probabilities in a non-Markov illness-death model. *Lifetime Data Analysis* **12**, 325-344.
- Pepe, M.S., Longton, G. and Thornquist, M. (1991). A qualifier Q for the survival function to describe the prevalence of a transient condition. *Statistics in Medicine* **10**, 413-421.
- Pepe, M.S. (1991). Inference for Events With Dependent Risks in Multiple Endpoint Studies. *Journal of the American Statistical Association* **86**, 770-778.
- Rodríguez-Girondo, M., and Uña-Álvarez, J. (2012). A nonparametric test for Markovianity in the illness-death model. *Statistics in Medicine* **31**, 4416-4427.
- de Wreede, L.C., Fiocco, M., and Putter, H. (2011). mstate: An R package for the analysis of competing risks and multi-state models. *Journal of Statistical Software* **38**, 1-30.
- Uña-Álvarez, J., and Meira-Machado, L. (2015). Nonparametric Estimation of Transition Probabilities in the Non-Markov Illness-Death Model: A Comparative Study. *Biometrics* DOI: 10.1111/biom.12288