Mathematics & Statistics | Lancaster University

Multi-state Models: An Overview

Andrew Titman

Lancaster University

14 April 2016

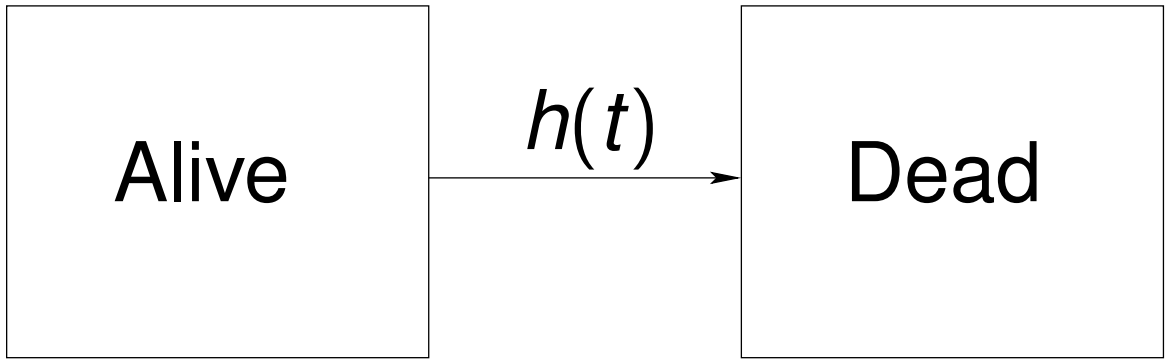
Overview

- Introduction to multi-state modelling
- Examples of applications
- Continuously observed processes
- Intermittently observed processes
- State Misclassification

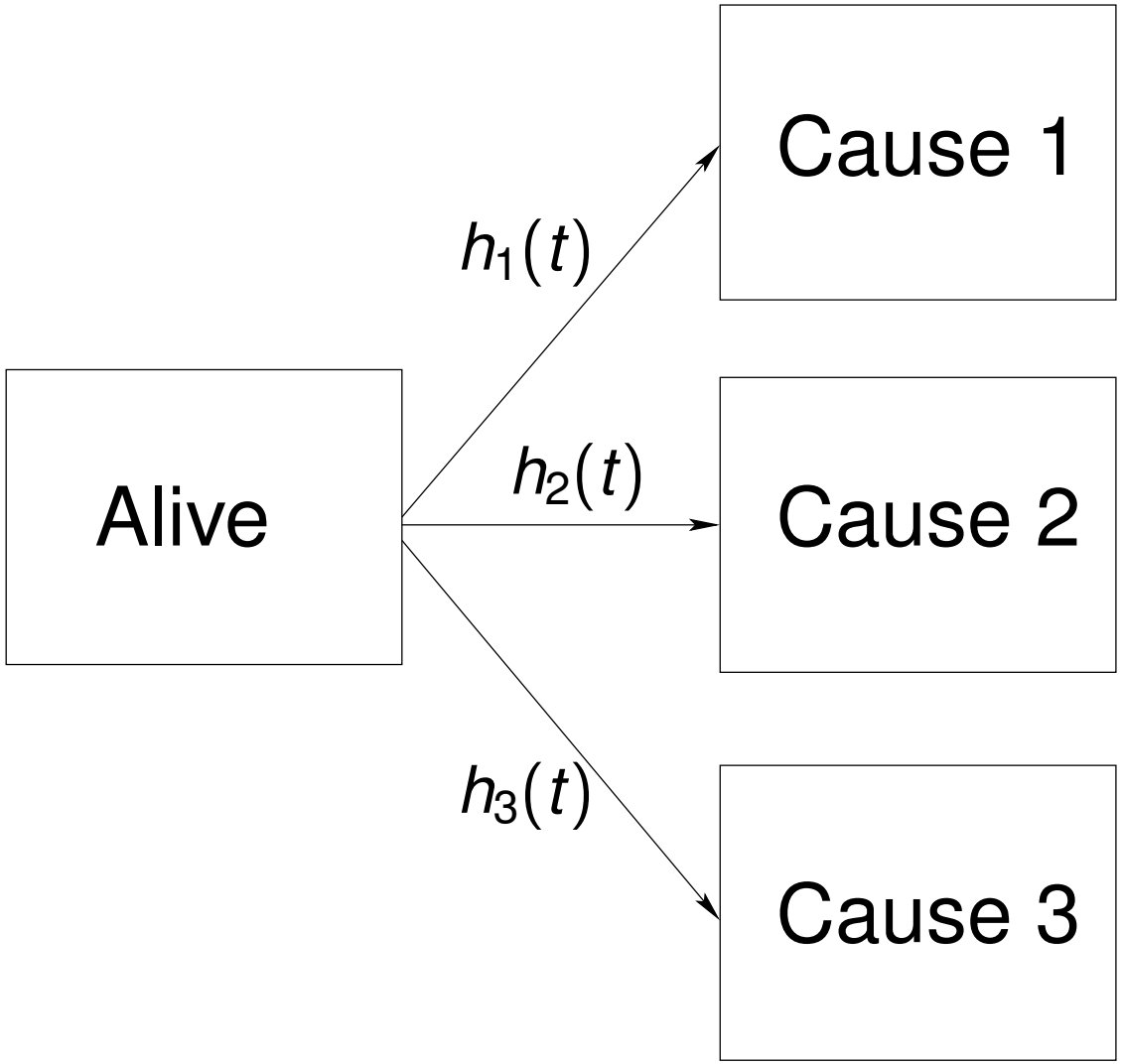
Multi-state models

- Models for *event history analysis*.
  - Times of occurrence of events
  - Types of event that occurred
- Joint modelling of survival and important (categorical) time dependent covariates
- Model the *transition intensities* between states of a process
  - E.g. healthy to diseased
  - E.g. diseased to death
- Modelling onset and progression of chronic diseases
- Process usually modelled in continuous time

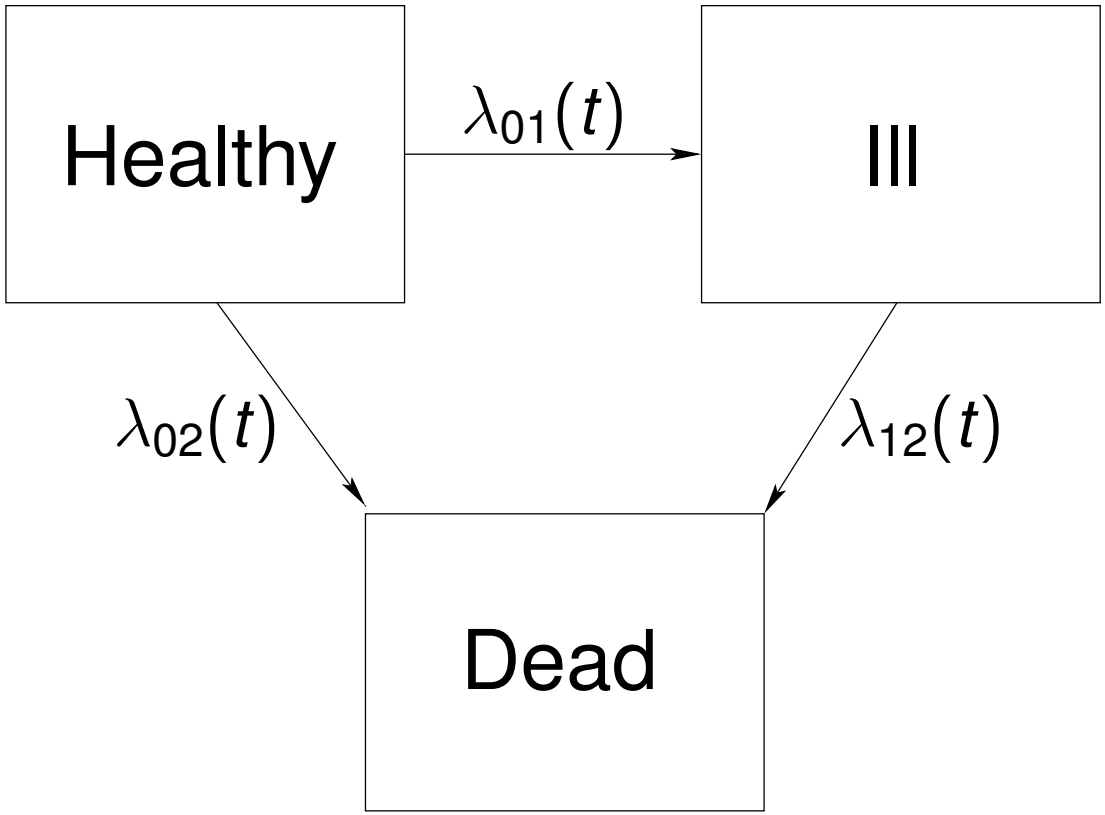
Example: Simple survival model



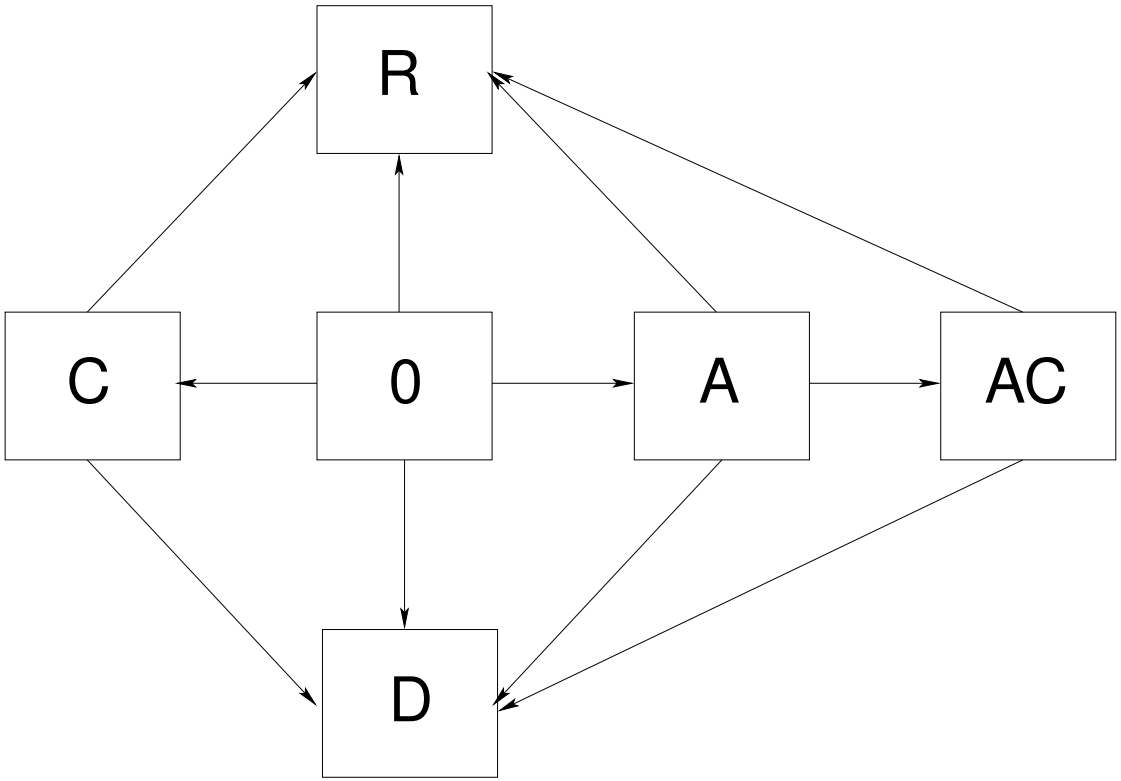
Example: Competing risks model



Example: Progressive Illness-death model



Example: Outcome prediction in bone marrow transplantation



Keiding et al (2001)

O: Initally transplanted
R: Relapse
D: Die in Remission
A: Acute graft-vs-host
C: Chronic graft-vs-host
AC: Acute+Chronic gvhd

Modelling interdependence

- Standard survival data, and also competing risks data, involve patients having at most one event of interest
- Once each subject can experience more than one event, assumptions need to be made about dependencies between events
- Most commonly a *Markov* assumption is adopted, where only the current state and time govern the trajectory of the process.
- Potentially other time scales and summaries of past history may be important

Mathematical framework

- Models defined by transition intensities between states

$$\lambda_{rs}(t, \mathcal{F}_t) = \lim_{\delta t \downarrow 0} \frac{\mathbb{P}(X(t+\delta t) = s | X(t) = r, \mathcal{F}_t)}{\delta t}$$

  where $\mathcal{F}_t$ is the history, or filtration, of the process up to time t.
- Common assumptions
  - Markov:

$$\lambda_{rs}(t, \mathcal{F}_t) = \lambda_{rs}(t)$$

  - (Homogeneous) Semi-Markov:

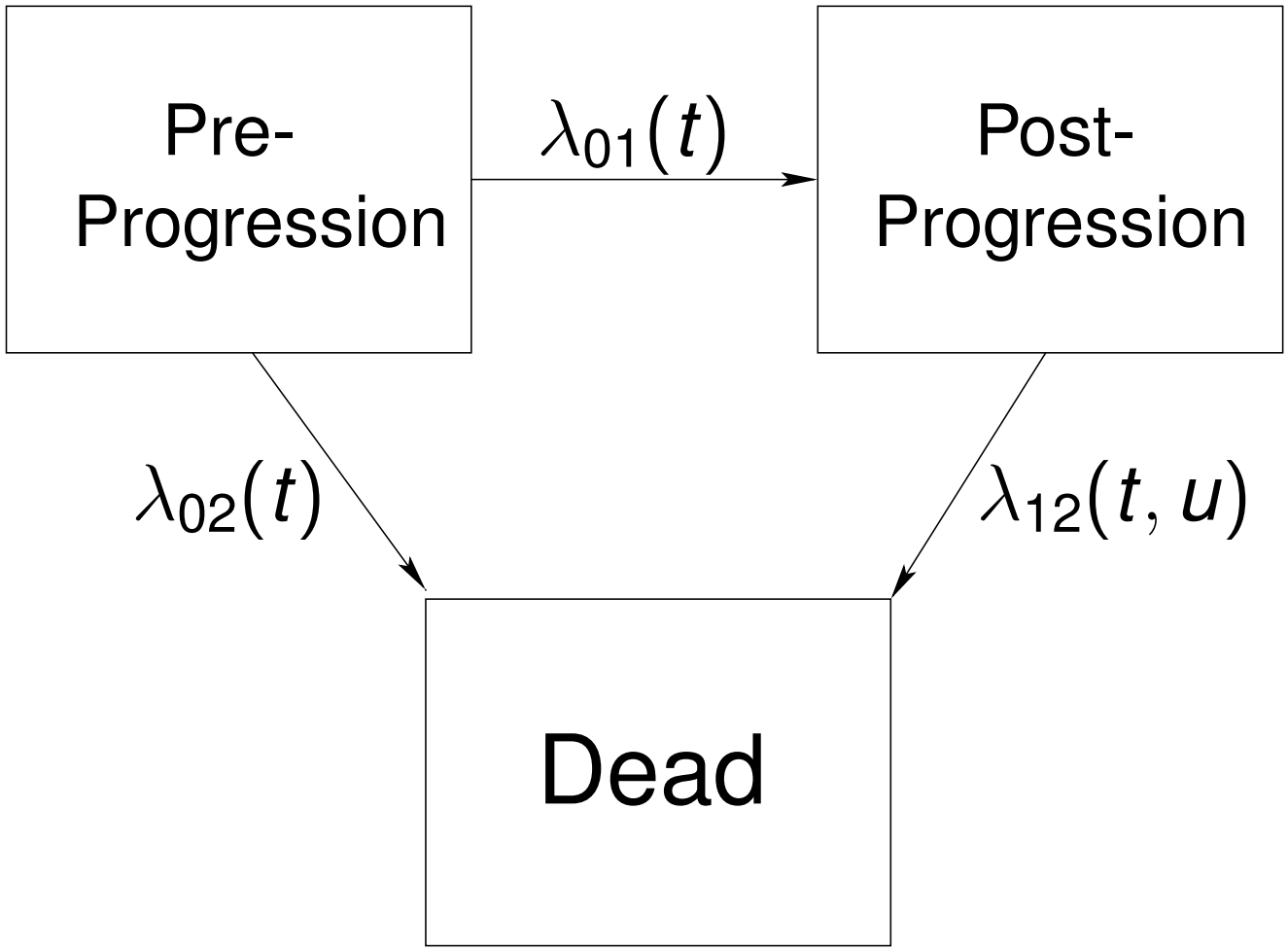
$$\lambda_{rs}(t, \mathcal{F}_t) = \lambda_{rs}(t - t_r)$$

    where t_r is the time of entry into current state r.

Application: Cancer progression and survival

- Overall survival (OS) after diagnosis of cancer is heavily influenced by an intermediate event of cancer progression
- Progression-free survival (PFS) is commonly reported in addition to overall survival in cancer trials and used as the primary endpoint in some cases
    - Defined as the minimum of time of death and time of progression
- Process of progression and survival can be represented through an illness-death multi-state model

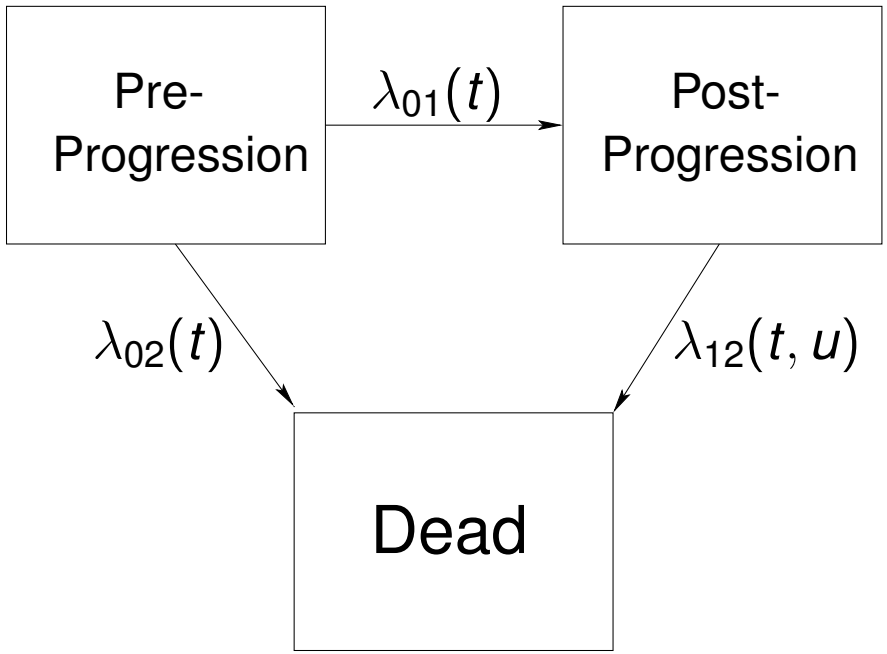
Application: Cancer progression and survival



Application: Cancer progression and survival

- Appropriateness of PFS as a surrogate endpoint depends on association with OS

- Can be characterized through the transition intensities
- Often reasonable to assume $\lambda_{12}(t,u) = \lambda_{12}(u)$ e.g. clock-reset at time of
- For PFS to be a good surrogate need $\lambda_{12}(t) \gg \lambda_{02}(t)$ progression



- Also need primary effect of treatments to be on $\lambda_{01}(t)$

Model fitting

- Estimation for multi-state models with continuous observation (up to right-censoring) is quite straightforward
- Under a Markov or semi-Markov assumption, the likelihood factorizes into periods of time at risk for each transition intensity
- Same formulation as for left-truncated (i.e. delayed entry) survival data
    - A patient becomes at risk of transitions out of state *i* from their time of entry into state *i*
- Models for each transition intensity can be maximized independently, provided there are no shared parameters
- Can be fitted using the *survival* package in **R**
- Determining quantities of interest, such as transition probabilities, from the models is more challenging.

Example: Illness-death model

Patient 1 enters illness state at 968 days. Dies at 1521 days

In counting process format would represent data as follows:

id	entry	exit	from	to	event	transtype
1	0	968	0	1	1	1
1	0	968	0	2	0	2
1	968	1521	1	2	1	3

Model without covariates in **R**:

```
fit0 <- survfit(Surv(entry,exit,event)~strata(transtype))
```

Transition Probabilities

The transition probabilities are defined as

$$P_{rs}(t_0, t_1) = P(X(t_1) = s | X(t_0) = r)$$

- Complicated function of the transition intensities
- In progressive models the transition probabilities can be written as an integral of the transition intensities with respect to the possible times of intermediate events.
- For more general Markov models, they are given by solving Kolmogorov Forward Equations (a system of differential equations)

Aalen-Johansen estimator

Under a Markov assumption, the matrix of transition probabilities of an R state multi-state model can be estimated non-parametrically using the *Aalen-Johansen estimator*.

$$\hat{\mathbf{P}}(t_0, t_1) = \prod_{k: t_0 \leq t_k \leq t_1} (\mathbf{I} + d\hat{\mathbf{\Lambda}}_k)$$

where $d\hat{\mathbf{\Lambda}}_k$ is an $R \times R$ matrix with (i, j) entry

$$d\hat{\Lambda}_{ijk} = \frac{d_{ijk}}{r_{ik}} \quad \text{for} \quad i \neq j,$$

$d\hat{\Lambda}_{iik} = -\sum_{j \neq i} d\hat{\Lambda}_{ijk}$ where d_{ijk}: number of $i \to j$ transitions at t_k, r_{ik}: number of subjects under observation in state i at t_k

Covariate models

- Standard approach to incorporating covariates is to assume proportional intensities (Cox-Markov model)

$$\lambda_{rs}(t; \mathbf{z}) = \lambda_{rs0}(t) \exp(\mathbf{z}' \beta_{rs})$$

- Extension of the Cox model, where the baseline intensities, $\lambda_{rs0}(t)$, are non-parametric
- Potentially allow different β_{rs} for each transition intensity and hence fit separate models to each.

Model building

- In practice, may have insufficient numbers of transitions between particular states to reliably fit independent models
- Constraints possible
  - Allow common regression coefficients across transitions out of a state

$$\beta_{rs} = \beta_r, \text{ for } s = 1, \ldots, R$$

  - Allow proportionality between transition intensities

$$\lambda_{rs0}(t) = \lambda_{r10}(t) \exp(\theta_s), s \neq 1$$

Interpretation issues

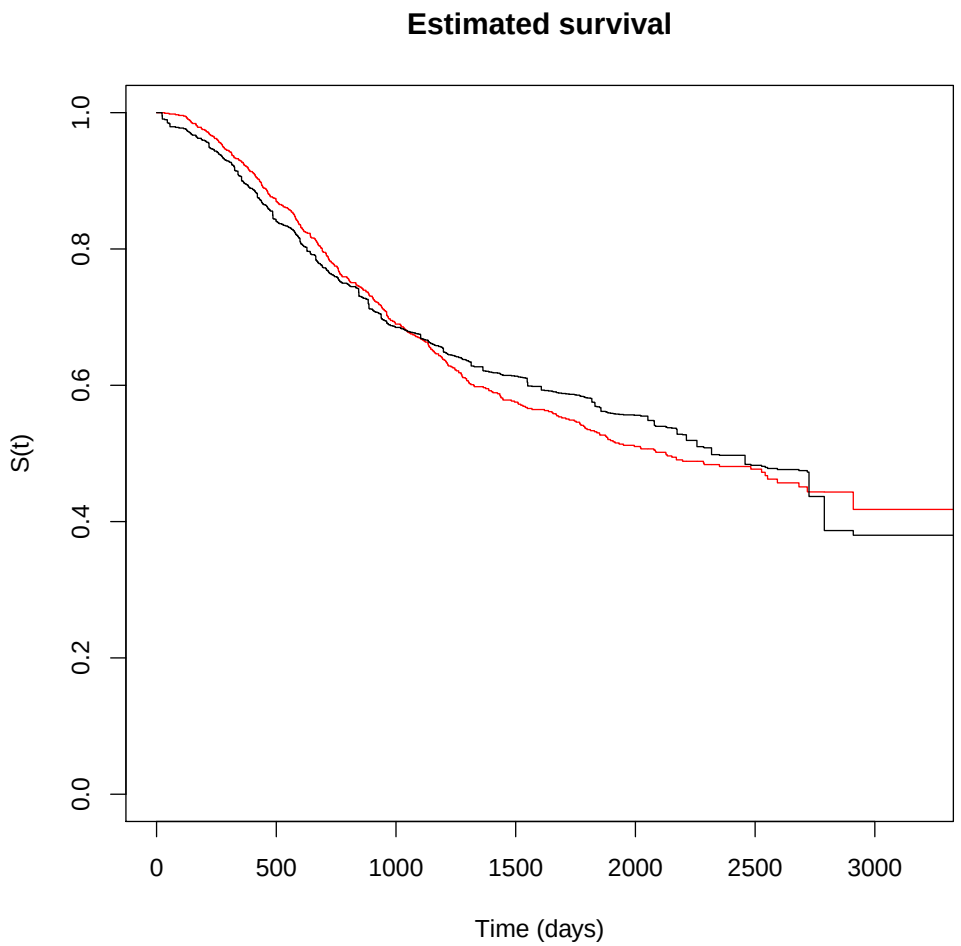
Having fitted a model for the effect of covariates on transition intensities, often want to also determine the transition probabilities for different covariate patterns

- This is reasonably straightforward to implement
  1. Calculate the estimates of the transition intensities for a given patient with covariates **z** (using the Nelson-Aalen-Breslow estimator)

$$d\hat{\Lambda}_{ijk}(\mathbf{z}) = d\hat{\Lambda}_{ijk}(0)\exp(\mathbf{z}'\hat{\beta})$$

  2. Plug these into the Aalen-Johansen formula
- Procedure performed by the *mstate* package in **R**
- But no guarantee of a simple relationship that explains the covariate effect
  - c.f. regression of cause-specific hazards in competing risks analysis

Example: Cancer
Overall survival



- Patients with different covariates may have crossing survival curves
- e.g. if a covariate is associated with faster progression but improved survival post-progression
- Direct regression methods possible via *pseudo-observations*

Application: Health Economics

- In Health Economics, cost-effectiveness usually determined by Incremental Cost-Effectiveness Ratio (ICER):

$$\text{ICER} = \frac{\Delta\text{Cost}}{\Delta\text{QALY}}$$

- Quality Adjusted Life Years (QALY) is an estimate of life expectancy weighted by health state
- A multi-state model with states corresponding to the distinct health states can be built
- QALY depends on state occupancy probabilities, $P_{1r}(0,t)$ and the health utilities, u_r, in each state r.

$$\text{QALY} = \sum_{r=1}^{R} \int_0^{\tau} u_r P_{1r}(0,t)dt$$

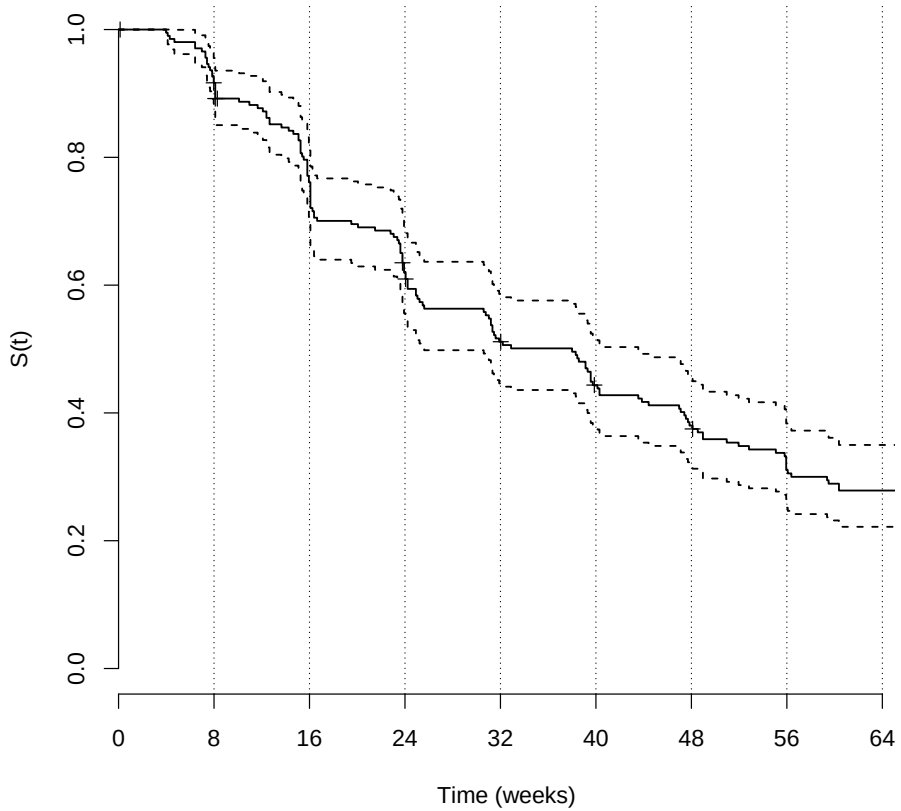
Intermittently observed processes

- It is not always possible to assume that the exact time of transitions between states can be observed (up to right censoring)
- Disease status may only be diagnosable at clinic visits, leading to interval censoring.
- In addition the precise transition(s) that took place may not be known.
- Time of entry into the absorbing state (e.g. death) may still be known up to right-censoring

Example: Cancer progression and survival

Mathematics & Statistics | Lancaster University

- While death will be known to the nearest day, progression is typically determined at scheduled follow-up visits
- Leads to dual censored data
- Almost always ignored in standard analysis, but does lead to bias (Zeng et al; 2015)



Intermittently observed processes

- Interval censoring makes estimation more difficult, particularly non- or semi-parametric approaches
- Usually need to assume that the observation times are fixed or, if random, non-informative.
  - e.g. next clinic visit time determined at current visit and not influenced by disease progression in between.
- In the Markov case the likelihood can be expressed as a product of transition probabilities.
  Subject in states $x_0, x_1, \ldots, x_m$ at times $t_0, t_1, \ldots, t_m$ has likelihood
$$\prod_{i=1}^{m} P_{x_{i-1}x_i}(t_{i-1}, t_i; \theta)$$
- Modifications needed to allow for exact times of death and when state occupied at end of follow-up is unknown.

Estimation: Markov case

- For time homogeneous models the transition probabilities can be expressed as a matrix exponential of the generator matrix of transition intensities

$$\mathbf{P}(t_0, t_1) = \exp(\mathbf{\Lambda}(t_1 - t_0))$$

  where $\mathbf{\Lambda}$ is the $R \times R$ matrix with (r, s) entry λ_{rs} for $r \neq s$ and $\lambda_{rr} = -\sum_j \lambda_{rj}$.
- Piecewise constant transition intensities can also be accommodated straightforwardly.
- Kalbfleisch & Lawless (1985) developed methods for maximum likelihood estimation
- Implemented in the *msm* package in **R** (Jackson, 2011).

State Misclassification

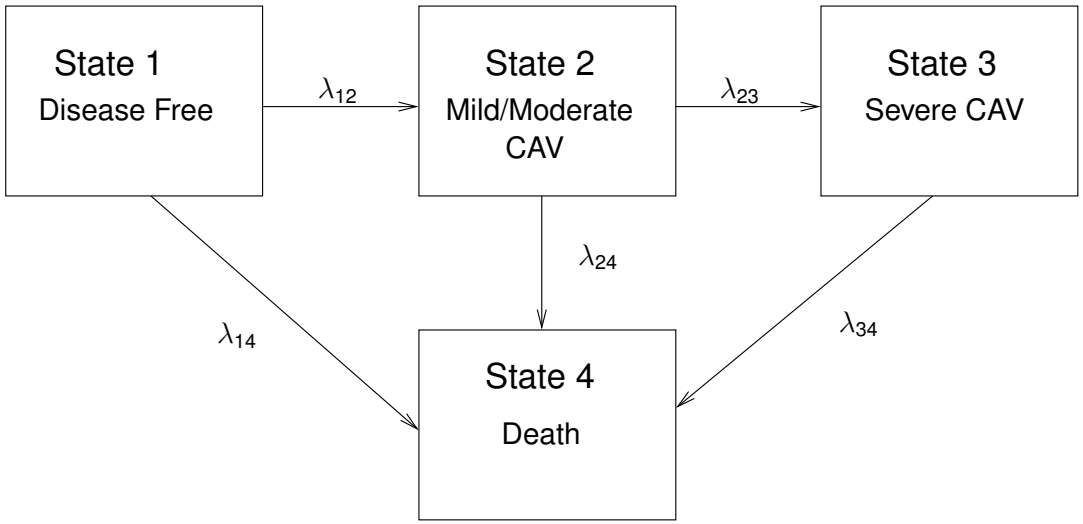
- Often the underlying disease of interest is known (or assumed) to be progressive - but this is contradicted by individual patients' sequences of states
  - Cognitive decline
  - Development of progressive chronic diseases
- Assume that the observed state at a clinic visit is subject to classification error
  - e.g. diagnosis of dementia via memory tests
  - e.g. patient undergoes a diagnostic test with $< 100\%$ sensitivity and specificity

Example: CAV in post heart-transplantation patients

Mathematics & Statistics | Lancaster University

4 state progressive illness-death model for development of cardiac allograft vasculopathy in post-heart transplant patients

- Diagnosis of disease status via angiogram at discrete clinic visits.
- 1972 clinic visits from 596 patients
- Angiogram potentially under diagnoses disease
- Underlying model assumed to be homogeneous Markov



Example: CAV in post heart-transplantation patients

	To				
From	1	2	3	4	Cen
1	1366	204	44	122	276
2	**46**	134	54	48	69
3	**4**	**13**	107	55	26

- Observed transitions include some cases that contradict the assumption of a progressive disease
- Ad hoc approach: Assume true state is maximum observed up to that time
- Instead assume angiogram may misspecify disease to adjacent states

$$e_{rs} = P(O_k = s | X_k = r)$$

State Misclassification

- A multi-state Markov model observed with misclassification can be represented by a hidden Markov model (HMM)
- HMM requires important assumption that conditional on the underlying states $x_1, x_2, \ldots, x_m$, the observed states $o_1, o_2, \ldots, o_m$ are independent.
  - Plausibility depends on method of diagnosis and regularity of tests
- Likelihood can be calculated iteratively through a forward recursion

$$\begin{aligned} P(O_1, \ldots, O_m, X_m) &= P(O_m|X_m) \sum_{X_{m-1}} P(O_1, \ldots, O_{m-1}, X_{m-1}) P(X_m|X_{m-1}) \\ &= e_{X_m O_m} \sum_{X_{m-1}} P(O_1, \ldots, O_{m-1}, X_{m-1}) P_{X_{m-1} X_m}(t_{m-1}, t_m) \end{aligned}$$

Example: CAV in post heart-transplantation patients

Mathematics & Statistics | Lancaster University

Parameter	Naive Markov	Hidden Markov
λ_{12}	0.039 (0.027,0.057)	0.033 (0.021,0.050)
λ_{14}	0.022 (0.017,0.030)	0.021 (0.015,0.029)
λ_{23}	0.199 (0.162,0.246)	0.190 (0.143,0.252)
λ_{24}	0.041 (0.023,0.075)	0.053 (0.029,0.099)
λ_{34}	0.146 (0.116,0.184)	0.155 (0.120,0.201)
$\beta_{12}^{(\text{IHD})}$	0.446 (0.185,0.706)	0.520 (0.234,0.807)
$\beta_{12}^{(\text{dage})}$	0.022 (0.011,0.033)	0.025 (0.013,0.037)
e_{12}		0.025 (0.015,0.042)
e_{21}		0.186 (0.123,0.272)
e_{23}		0.065 (0.038,0.108)
e_{32}		0.102 (0.051,0.194)

State Misclassification

- Can also assume misclassification for processes with backwards transitions
  - e.g. (discretization of) CD4 count in patients with HIV
- Becomes more difficult to distinguish between different parts of the underlying and observation process
  - Markov assumption (vs. say semi-Markov dependence)
  - Assumption of conditional independence of observations given true states
  - Assumption of non-informative observation times

Further issues

- In some settings may have clustered multi-state processes, e.g. multiple processes from the same person
  - Either account for via shared random effects or via robust marginal analyses
- Data may have complicated sampling schemes
  - e.g. Truncation because patients only including in a study if in a given state - requires modified likelihoods to estimate population-level quantities
  - Patient initiated visit times - observation times no longer non-informative: limited methods to deal with this currently.
- Goodness-of-fit/Model diagnostics
  - Models, particularly those for intermittently observed data, often make strong assumptions. Important to assess model fit.

Conclusions

- Multi-state models are particularly useful if we want to build a model for an overall process of survival
- Potential to obtain dynamic predictions, not just marginal quantities
- Able to accommodate observational data with complex sampling, e.g. intermittent patient-specific, unequally spaced observation times.

References

Mathematics & Statistics | Lancaster University

- Kalbfleisch, J. D., & Lawless, J. F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, **80**, 863-871.
- Keiding, N., Klein, J. P., & Horowitz, M. M. (2001). Multistate models and outcome prediction in bone marrow transplantation. *Statistics in Medicine*, **20**, 1871-1885.
- Jackson, C. H. (2011). Multi-state models for panel data: the msm package for R. *Journal of Statistical Software*, **38**(8), 1-29.
- Zeng, L., Cook, R. J., Wen, L., & Boruvka, A. (2015). Bias in progressionfree survival analysis due to intermittent assessment of progression. *Statistics in Medicine*, **34**, 3181-3193.