**Basic concepts**

1. **The Markov assumption (one clock):** the future (event/status) only depends on the present (event/status); when conditioning on the present (event/status), the future (event/status) is independent of the past (event/status) process.
2. **Different possibilities for violating the Markov assumption (two relevant possibilities):**
3. The future state does not only depend on the current state but also depends on the time spent in the current state.
4. The future state does not only depend on the current state but also depends on the past event. This point can be argued from different perspectives, depending on the data source.
   1. Example: people who reverted from CIND to normal cognition have a higher risk of developing CIND again or dementia compared with people who never had CIND before. This can be addressed by treating states with different state histories as distinct states, but this can be infeasible due to sparse data.
5. **Time-homogeneous Markov model (one clock):** constant transition intensity over the timescale.
6. **Semi-Markov model (two clocks, clock reset):** transition intensity depends on the time spent in the current state. This type of model can not address point 2.2.a by default.

**Comparisons between statistical packages for multistate modelling**

|  |  |  |  |
| --- | --- | --- | --- |
| **Packages** | **Default assumptions** | **Methods to relax the assumptions** | **Suitable data structure** |
| msm (R) | 1. Time-homogeneous transition intensity  2. Markov assumption | 1.1 piecewise-constant transition intensity  1.2 timescale time variable as a time-varying covariate  2.1 phase-type semi-Markov model  2.2 past event as a time-varying covariate; treating states with different state history as distinct states | Panel data (e.g., cognitive state measured at scheduled times), when the exact transition time is unknown, unknown transitions between time intervals, and censored states ([Multi-state modelling with msm: a practical course (chjackson.github.io)](https://chjackson.github.io/msm/msmcourse/index.html)) |
| flexsurv (R)  multistate & merlin (stata)  Flexible transition intensity distribution | 1. Exact transition time (flexsurv allows interval-censored event times)  2. No censored states  3. No misclassification of states  4. Markov/semi-Markov assumption | 1. treating event time as interval-censored (require time interval information)  2. MICE does not suffice  3. None; sensitivity analysis needed  4. past event as a time-varying covariate; treating states with different state histories as distinct states | Time-to-event data, panel data when the exact transition time/even time interval is known (e.g., hospitalization)  (<https://www.jstatsoft.org/article/view/v070i08>  <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.7448>) |
| mstate (R)  Does not model baseline transition intensity, can not make predictions (e.g., transition probability) when there’re reversions from ill to healthy state?? | 1. Exact transition time  2. No censored states  3. No misclassification of states  4. Markov/semi-Markov assumption | 1. treating event time as interval censored (require time interval information)  2. MICE does not suffice  3. None; sensitivity analysis needed  4. past event as a time-varying covariate; treating states with different state histories as distinct states | Time-to-event data, panel data when the exact transition time is known (e.g., hospitalization)  (<https://cran.r-project.org/web/packages/mstate/vignettes/Tutorial.pdf>) |
| Others | SmoothHazard (R, [SmoothHazard: An R Package for Fitting Regression Models to Interval-Censored Observations of Illness-Death Models | Journal of Statistical Software (jstatsoft.org)](https://www.jstatsoft.org/article/view/v079i07)), nhm (R [nhm-manual.pdf (microsoft.com)](https://cran.microsoft.com/snapshot/2022-10-22/web/packages/nhm/vignettes/nhm-manual.pdf)), etc | | |

**Important aspect:** check the goodness of fit for models by (1) plotting observed and expected numbers/prevalence of each state at a set of time points, (2) using fit indices such as AIC, BIC, or (3) other suitable tests.

**Other aspects:**

1. **Sparse data bias-**likelihood-based models give biased point estimates & confidence intervals when the number of events in some covariate combinations is small.
2. **Missing data in covariates or outcomes-**commonly used methods to impute missing data, such as MICE, don’t suffice.
3. Prediction with time-varying covariates is not straightforward.