# Estimation and prediction in multistate models

Bendix Carstensen
Senior Statistician, Steno Diabetes Center Copenhagen, Gentofte, Denmark
b@bxc.dk http://BendixCarstensen.com

This talk will be a largely non-technical description of concepts from statistical models for persons' changing status over time, the so-called multistate models (often mis-termed Markov-models). In clinical studies, the states typically represent different kinds of treatment and complication levels.

Some studies will exclusively focus on comparison of pre-defined sets of transition rates, in which case traditional modeling of rates and comparison via hazard ratios will be the focus: For example we could compare the cancer occurrence rate between persons with and without diabetes, ignoring both the occurrence rates of diabetes as well as the mortality rates.

Other studies will also focus on the absolute measures such as probability of being in a given state or the expected time spent in a state.

As illustration I shall use an example from a clinical trial, the Steno-2 study, initiated at Steno Diabetes Center in the 1990s. In the long-term follow-up we wanted to assess the treatment effect not only on incidence rates of cardiovascular disease (CVD) and mortality (hazard ratios), but also the effect on cumulative measures such as survival, lifetime lost, and time spent with cardiovascular disease (CVD). We wanted to study how much time persons in the two arms spent being "alive", respectively "alive without CVD".

I will describe the concepts used to derive hazard(ratio)s for mortality and CVD occurrence as well as baseline hazrads and outline how these can be used to construct the cumulative measures of interest, and what assumptions are needed in order to do this in a meaningful way.

I shall argue that models with smooth effects of timescales are both conceptually and computationally preferable over the much used non-parametric approaches, such as the Kaplan-Meier estimator and much used Cox-analysis. I will describe the `Lexis` machinery from the `Epi` package for `R` facilitating this and also briefly describe the ideas behind the simulation tool `simLexis` that allows predicions from multistate models.

The emphasis of the talk will be on describing the concepts and the logic behind reported results and not on the technical details.

# References

[1] S. Iacobelli and B. Carstensen. Multiple time scales in multi-state models. *Stat Med*, 32(30):5315–5327, Dec 2013.

[2] Bendix Carstensen. Simulation of multistate models with multiple timescales: `simLexis` in the `Epi` package. http://BendixCarstensen.com/Epi/simLexis.pdf or `Epi` vignette, 2015.