

# Estimating statistics from multi-state models using simulation with `multistateutils`

Stuart Lacy<sup>1</sup>

<sup>1</sup> Department of Health Sciences, University of York, UK

DOI: [10.21105/joss.00893](https://doi.org/10.21105/joss.00893)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 25 July 2018

Published: 08 October 2018

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

`multistateutils` provides functionality for analysing multi-state models using Discrete Event Simulation (DES) in R.

DES is a method of exploring the behaviour of dynamic systems by discretising trajectories into a series of state transitions that occur at specific times. There exists general-purpose DES software packages in several programming languages, such as `simmer` (R) (Ucar, Smeets, & Azcorra, 2018) or `simPy` (Python) (Matloff, 2008) that can be applied to many systems regardless of the application area. Multi-state modelling is an area of biostatistics that is concerned with the extension of survival analysis to the case where there are multiple time-to-event transitions of interest, such as a patient experiencing various treatments on their disease pathway (Meira-Machado, Uña-Álvarez, Cadarso-Suárez, & Andersen, 2009). There are several measures of interest that can be generated from a multi-state model, such as the probability of being in a given state at a certain time, or the estimated time spent in a given state, which can be obtained analytically provided certain assumptions - most notably the Markov property - are met; however, simulation provides a more flexible alternative and can be used to extract any desired summary statistic.

`multistateutils` is designed for the specific use case of multi-state modelling in biostatistics and provides a means for estimating statistics of interest from a multi-state model by simulation without exposing users to DES specific terminology. The functions in this package operate on parametric statistical models of transition times from the widely-used `flexsurv` (Jackson, 2016) package, allowing the user to fit transition models in a familiar environment before inputting them directly into the simulation, rather than having to learn the parameterisation of a general purpose DES toolbox. The primary functionality provided by `multistateutils` is an interface to running discrete event simulation over a cohort of patients, thereby providing the simulated outcomes from which users can estimate any required statistics. In addition, for two common statistics - transition probabilities and length of stay estimates - there are functions that will set up and run the simulation and return the estimates directly.

The simulation engine is written in C++ to provide an efficient implementation of a DES. The source code has been archived in Zenodo with the DOI (Lacy, 2018).

## References

- Jackson, C. (2016). `flexsurv`: A platform for parametric survival modeling in R. *Journal of Statistical Software*, 70(8), 1–33. doi:[10.18637/jss.v070.i08](https://doi.org/10.18637/jss.v070.i08)
- Lacy, S. (2018, July). `Stulacy/multistateutils`: 1.2.0. doi:[10.5281/zenodo.1320356](https://doi.org/10.5281/zenodo.1320356)

- Matloff, N. (2008). Introduction to discrete-event simulation and the simpy language.
- Meira-Machado, L., Uña-Álvarez, J. de, Cadarso-Suárez, C., & Andersen, P. K. (2009). Multi-state models for the analysis of time-to-event data. *Statistical methods in medical research*, 18(2), 195–222. doi:[10.1177/0962280208092301](https://doi.org/10.1177/0962280208092301)
- Ucar, I., Smeets, B., & Azcorra, A. (2018). simmer: Discrete-event simulation for R. *Journal of Statistical Software*, *accepted for publication*.