

# UNDERSTANDING AND AVOIDING SURVIVAL BIAS: AN APPLICATION OF MULTISTATE MODELS

Arthur Allignol<sup>1,2,\*</sup>

<sup>1</sup>Institute of Medical Biometry and Medical Informatics, University Medical Center Freiburg, Freiburg, Germany

<sup>2</sup>Freiburg Center of Data Analysis and Modelling, University Medical Center Freiburg, Freiburg, Germany

\*`arthur.allignol@fdm.uni-freiburg.de`

Survival bias is remarkably common and often overlooked in clinical and epidemiological articles [1]. We consider two types of such bias known as length bias and time-dependent bias, respectively. Length bias occurs, if the time interval between study entry and the final outcome is implicitly analysed as being longer than the actual time interval under observation. This especially happens in epidemiological cohort studies if age is modelled as the time scale without left-truncation. Time-dependent bias occurs, if the timing of a time-dependent exposure is not taken into account, i.e., treated as known at baseline. Hospital-acquired infections are examples for time-dependent exposures in clinical epidemiology [2]. Both types of bias distort the statistical analysis through misclassified at-risk time, which is best understood from a multistate modelling point of view. Exemplarily, the methodology is demonstrated in a recent study on the potential survival advantage of Academy Award-winning actors and actresses [3]. Another example is discussed, in which we consider the estimation of the rate of spontaneous abortion in drug-exposed pregnancy studies [4]. The multistate modelling approach clearly reveals how the biased analyses led to distorted risk sets and, hence, biased estimation. A better understanding of these issues may prevent these biases due to improperly analysed studies in epidemiology.

**Key words:** survival bias, cumulative hazard, event history, left truncation, hazard ratio

[1] van Walraven, C., Davis, D., Forster, A., and Wells, G. (2004), Time-dependent bias was common in survival analyses published in leading clinical journals, *J Clin Epidemiol*, 57, 672-682.

[2] Beyersmann, J., Wolkewitz, M., and Schumacher, M. (2008), The impact of time-dependent bias in proportional hazards modelling, *Stat Med*, 27, 6439-6454.

[3] Redelmeier, D. and Singh, S. (2001), Survival in Academy Award-winning actors and actresses, *Ann Intern Med*, 134, 955-962.

[4] Meister, R. and Schaefer, C. (2008). Statistical methods for estimating the probability of spontaneous abortion in observational studies—Analyzing pregnancies exposed to coumarin derivatives. *Reproductive Toxicology*, 26, 31–35.