

Abstract

We describe a procedure for estimating measures of medical treatment outcomes from published time-to-event summary data, that is, from a finite set of points retrieved from presented Kaplan-Meier curves, corresponding numbers of patients remaining at risk, and, if available, total event counts over the course of the underlying study. These quantities—rather than individuals' time to events and censoring status, which are usually unavailable—are considered the data used for our analyses. To approximate the population time-to-event distribution, we use a Bayesian approach and combine a piecewise constant-hazard survival model with a binomial model of event counts in a sequence of intervals partitioning the period of observation. Uncertainty about observed but unpublished numbers of events and censorings in these intervals is modeled using discrete uniform distributions restricted by the specific constellation of bordering Kaplan-Meier estimates, numbers at risk and total numbers of failures. We use our model to sample from posterior distributions of chances to survive selected time points and from posterior distributions of selected time-to-event quantiles under two competing treatments. Modeling uncertainty about these quantities leads naturally to posterior distributions of their between-treatment differences and ratios. We also estimate numbers of patients that need to be treated to prevent one event over an interval of interest. Moreover, our method opens a way to estimate expected survival times, as well as their between-treatment differences, based on an extrapolation of the hazard beyond the end of the period of observation. In case of mature data, this approach will provide a practically useful approximation, although in a general setting, the validity of the involved assumptions can easily be evaluated using sensitivity analyses. Our method is illustrated using the example of a published study report summarizing a clinical trial conducted in the area of oncology.