The Impact of Censoring Assumptions in the Generation of Individual Patient Data from Kaplan-Meier Estimates

Edmondson-Jones M[1]*, Sullivan W[1]

[1]Delta Hat, Nottingham, UK.

MSR65

Background

It is common for publications to report hazard ratios (HRs) or median survival times when making comparisons regarding time-to-event (TTE) endpoints, such as overall or progression-free survival. However, these measures have serious limitations, for example, where extrapolations are of interest, or where proportional hazard (PH) assumptions may be inappropriate. For this reason, alternative approaches have been of interest in applications such as comparative (cost-)effectiveness analyses for health technology appraisal (HTA).

It is standard for publications to include Kaplan-Meier (KM) plots to summarise observed TTE data. Where individual participant data (IPD) are not available, methods to reconstruct (pseudo) IPD from KM figures are useful; an algorithm developed by Guyot et al (2012) for this purpose is popular. This algorithm's use has been facilitated by its availability through statistical software such as R's survHE package.

Typically, the estimation of event times and summary information on censoring has been of primary interest in generating pseudo IPD, as these are sufficient to reconstruct the published KM data. However, we note that the precise placement of censoring times typically also has influence on survival analyses, such as the fitting of parametric models, through the distribution of numbers 'at risk' over time. Rogula et al (2022) proposed extracting censoring data from KM plots by digitising censoring marks. In principle, this can allow for TTE IPD to be recreated, albeit without covariates. However, censoring markings may be obscured, hard to discern, or absent, and in such cases this approach is not available.

Objectives

This study aims to assess the importance of censoring assumptions on pseudo IPD generated from KM estimates, given their importance for HTA. The study focuses on the placement of censoring times, rather than the estimation of event timepoints which it accepts as obtained by the Guyot et al algorithm, using the R survHE package. Rather, we consider alternative strategies for the placement of censor times. The impact of these approaches is assessed using a parametric extrapolated restricted-mean survival time (RMST) as an HTA-relevant measure, in addition to HR.

Methods

As a case study, pseudo-IPD were generated from published overall survival (OS) KM data from a randomised, controlled trial (RCT) which exhibited significant censoring (Baselga et al, 2012). An example with extensive censoring was selected to demonstrate the potential impact of censoring placement decisions. The data used were from an interim analysis of this RCT where censoring was anticipated to be high. In the dataset, 333 out of 402 (83%) pertuzumab and 310 out of 406 (76%) control patients were censored for OS.

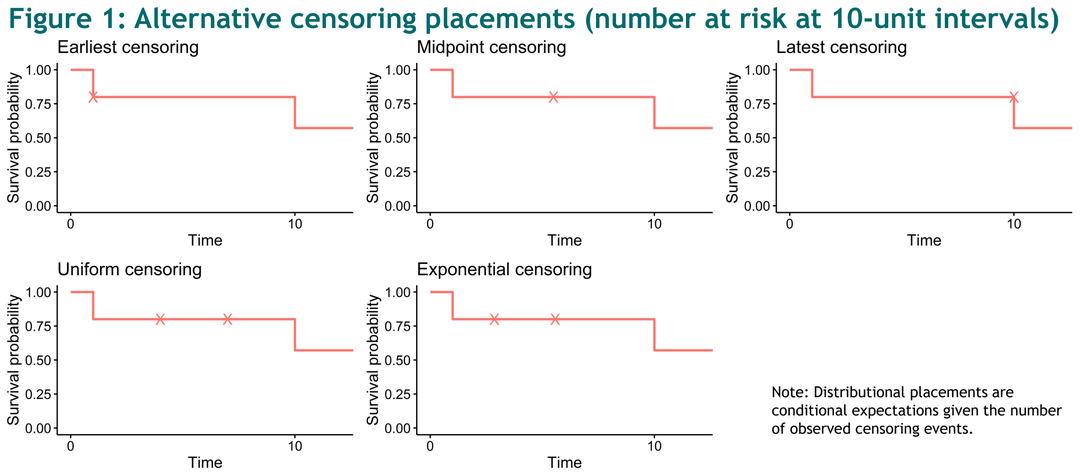
The KM plot was digitised using WebPlotDigitizer (https://apps.automeris.io/wpd/) and pseudo IPD were generated using the digitise function in the R survHE package. For each event, estimated times were used. For each censoring point, the 'window' between consecutive events was noted.

Two types of adjustment were made to censoring time:

1. Where relevant, censoring windows were split either side of a 'number at risk' breakpoint to preserve the numbers at risk at reported time points;
2. Within each possible censoring window locate censorings a) as early as possible; b) as late as possible; c) at the midpoint; d) assuming a uniform distribution; and e) assuming an exponential distribution (constant rate) initially, followed by uniform (administrative) censoring from 10 months

Figure 1 illustrates alternative scenarios for the placement of censoring times.

Figure 1: Alternative censoring placements (number at risk at 10-unit intervals)



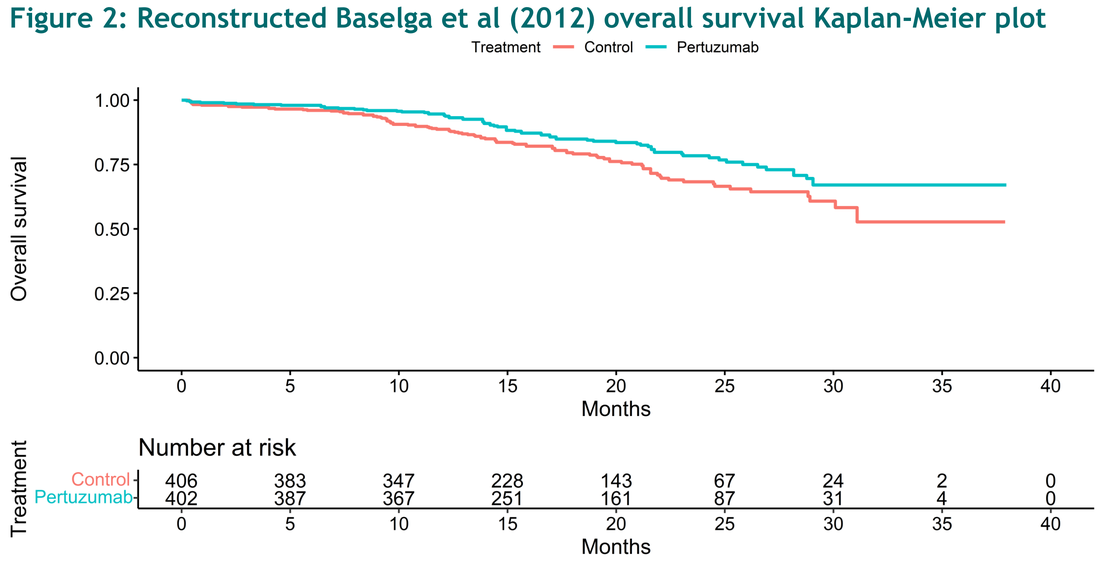
Note: Distributional placements are conditional expectations given the number of observed censoring events.

The importance of different censoring assumptions for HTA-relevant outcomes was assessed by comparison of estimated HRs and by comparison of 10-year RMST based on a Gompertz model extrapolation. (Gompertz was found to be the best fitting standard parametric model according to AIC based on the survHE estimates and was fitted independently to each arm.)

Results

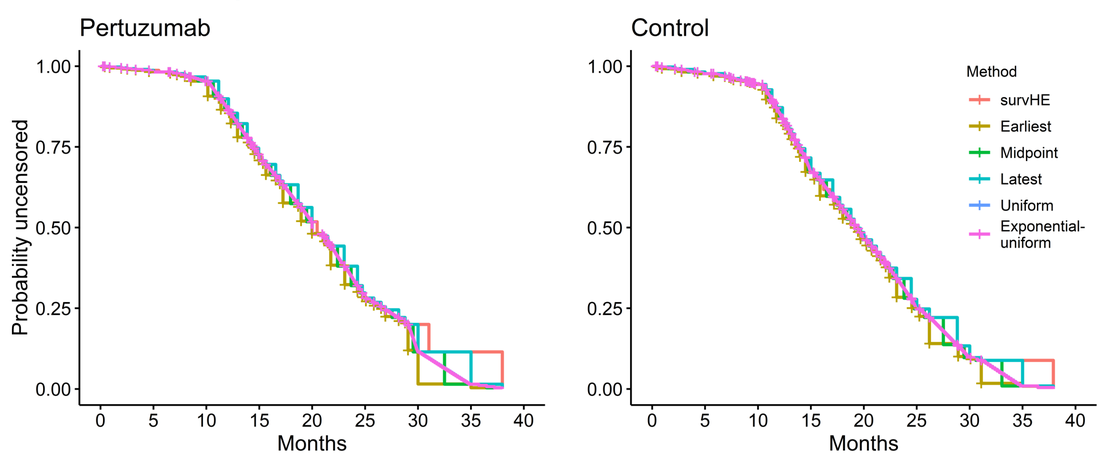
The pseudo IPD generated from the KM plots in Basalga et al (2012) are reproduced, without censoring markings, in Figure 2. The censoring positions are illustrated in the source publication but the KM estimates (recreated from the pseudo IPD) are provided here without any censoring marks.

Figure 2: Reconstructed Baselga et al (2012) overall survival Kaplan-Meier plot



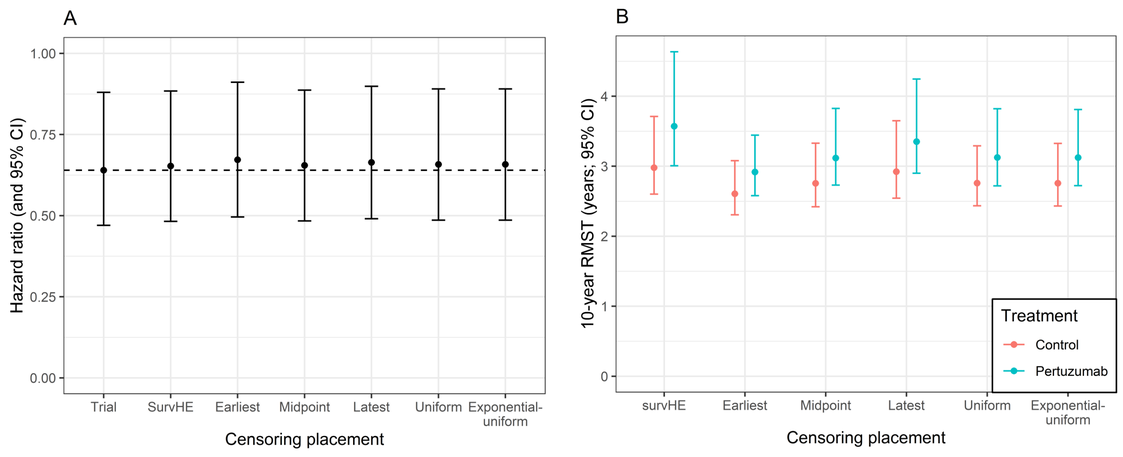
The probability of remaining uncensored under each censoring method is illustrated in Figure 3. It is apparent that the censoring mechanism seems independent of treatment arm, offering opportunities to 'borrow' information on the censoring distribution.

Figure 3: Kaplan-Meier plots of the probability of remaining uncensored under each alternative placement method



The HR and 10-year RMST estimates obtained using each censoring placement approach are summarised in Figure 4.

Figure 4: Panel A = HR estimates (with 95% confidence intervals); Panel B = 10-year RMST estimates (with 95% confidence intervals)



HR estimates spanned 0.653-0.672 (survHE: 0.653) and Gompertz 10-year RMST spanned 2.92-3.57 (Pertuzumab; survHE: 3.57) years and 2.61-2.98 (Control; survHE: 2.98) years. In this example, consistent changes to the placement of censoring points across arms had little impact on the estimated relative treatment effect, with the survHE derived estimate being slightly closer to the published RCT estimate but also close to the midpoint, uniform and exponential-uniform estimates. However, the survHE estimates of 10-year RMST appeared relatively optimistic compared to other approaches. Expected life years gained (pertuzumab vs control) over a 10-year window were 0.591, 0.311, 0.362, 0.428, 0.366 and 0.365 years with survHE, earliest, midpoint, latest, uniform and exponential-uniform censoring placement respectively.

Discussion

The results of this study highlight the importance of reasoned specification of censoring assumptions in the generation of individual patient data from KM estimates if censoring marks are obscured, hard to discern, or absent.

Where participant follow-up is relatively complete, and there are minimal censored observations, decisions on the placement of censoring times, as expected, seem to have minimal impact. However, we have shown that where censoring is more extensive, assumptions regarding censoring times can have a meaningful impact on expected survival times. As mean survival can be a major factor in cost-effectiveness evaluations informing HTA decisions, the distribution of censoring times should be carefully considered and justified.

References

Guyot, Patricia, AE Ades, Mario JNM Ouwens, and Nicky J. Welton. 'Enhanced Secondary Analysis of Survival Data: Reconstructing the Data from Published Kaplan-Meier Survival Curves'. BMC Medical Research Methodology 12, no. 1 (1 February 2012): 9. https://doi.org/10.1186/1471-2288-12-9.

Baselga, José, Javier Cortés, Sung-Bae Kim, Seock-Ah Im, Roberto Hegg, Young-Hyuck Im, Laslo Roman, et al. 'Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer'. The New England Journal of Medicine 366, no. 2 (12 January 2012): 109-19. https://doi.org/10.1056/NEJMoa1113216.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Baio G (2020). "survHE: Survival Analysis for Health Economic Evaluation and Cost-Effectiveness Modeling." _Journal of Statistical Software_, *95*(14), 1-47. doi: 10.18637/jss.v095.i14 (URL: https://doi.org/10.18637/jss.v095.i14).

Rogula, Basia, Greta Lozano-Ortega, and Karissa M. Johnston. 'A Method for Reconstructing Individual Patient Data From Kaplan-Meier Survival Curves That Incorporate Marked Censoring Times'. MDM Policy & Practice 7, no. 1 (31 January 2022): 23814683221077643. https://doi.org/10.1177/23814683221077643.

Presented at ISPOR Europe 2023, 12-15 November 2022, Copenhagen, Denmark.

Contact: medmondson-jones@deltahat.com