Re: Visualizing Length of Survival in Time-to-Event Studies: A Complement to Kaplan–Meier Plots

Royston et al. (1) recently suggested a method of presenting data that was complementary to Kaplan–Meier curves, with the aim to supplement ordinary survival probability plots with a direct and intuitive indication of the variability in survival times among individual patients. They showed how survival time histograms could be obtained by imputing censored survival times with values predicted by a log-normal model.

We agree with Royston et al. (1) that Kaplan–Meier plots tend to conceal survival time variability, thus obscuring the impact of the treatment. We also agree the actual benefits of the treatment depend not only on the hazard ratio but also on the shape of the underlying probability distribution, which is disease related (2). Hence, common measures of effects on the vertical axis of the survival function and time plane (eg, the hazard ratio) ought to be supplemented with measures of the corresponding effects on the horizontal axis (eg, the median or any appropriate quantile time).

We think the histograms proposed by Royston et al. (1) might be useful as diagnostic tools for modeling purposes (eg, for mixture or subgroup recognition). We believe, however, that the imputation method they suggested is not necessary for graphic display purposes and that some clinicians might not feel comfortable with plots that use random-generated data rather than the actual data. We also believe that simple distribution percentiles highlight differences on the magnitude of the effect between populations better than histogram plots by focusing attention on a few striking points. A simple box-and-whiskers plot can be drawn by use of empirical quantiles from Kaplan–Meier survival estimates and should suffice to provide a clear picture of the location and scale of the time-to-event distribution. Box-plot whiskers could be drawn from the 5th to the 95th percentiles or to the largest percentile estimable from the survival curve.

We present a simple simulation study to show box plots for a hypothetical trial in which experimental and control treatment events were generated from two exponential distributions with rates $\lambda_{\text{EXP}} = 0.75\lambda_{\text{CTRL}}$ and $\lambda_{\text{CTRL}} = 0.1$, respectively. Total sample size was set to 300 and censoring was uniformly distributed on $(0,c)$, where $c$ was chosen to provide 10% and 50% censoring rates. Theoretical median values of $\log(2)/\lambda$ are equal to 9.24 and 6.93 for the experimental and control arms, respectively. Sample medians computed from the simulated (not censored) dataset were 9.18 and 7.63, respectively. The corresponding median values estimated from survival curves were 9.18 and 7.6, respectively, for a 10% censoring rate and 9.23 and 7.67 for a 50% censoring rate.

Box plots for the two censoring rates of 10% (Figure 1, A) and 50% (Figure 1, B) are presented. The upper limit in Figure 1, B is the 87th percentile, the largest estimate from both the survival curves, although the high 50% censoring rate. This graph appears to be a convenient way to graphically depict and compare the two distributions. Low confidence should, however, be attached to the whiskers’ upper limits because of the sparse data in the right tails of survival curves, typically for high censoring rates. Reliability could be improved by choosing a rule for the upper quantile to be estimated (eg, a minimum of 20 subjects at risk).

NICOLA LAMA
CIRO GALLO

References

Funding
Ciro Gallo was Partially supported by non-profit charity AIRC (Italian Association for Research on Cancer) (CG).

Notes
Affiliations of authors: Department of Medicine and Public Health, Medical Statistics Unit, Second University of Napoli (NL, CG).
Correspondence to: Nicola Lama, PhD, Department of Medicine and Public Health, Second University of Napoli, Via Luciano Armanni 6, 80138 Napoli, Italy (e-mail: nicola.lama@unina2.it).
DOI: 10.1093/jnci/djn251
© The Author 2008. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.

Response
Lama and Gallo propose an alternative to our approach in which individual survival times are displayed with box-and-whiskers plots (box plots) of percentiles of the survival distribution. Their suggestion is not new [eg, see figures 7 and 8 in Gentleman and Crowley (2)]. As stated by Gentleman...
and Crowley, “Thus, these methods [box plots] and the others suggested [including scatter plots] are not competitors, but, rather are supplementary [to Kaplan–Meier curves]” (our square brackets). This is also the primary message of our paper—to expand the presentation of survival data by the use of informative methods in addition to Kaplan–Meier curves.

Box plots are simple in the sense that their values can be read easily from Kaplan–Meier curves. However, they lose a lot of information because many individual values are summarized with only a few percentile points. The loss is considerable in the upper end of the survival distribution (ie, for those with the longest survival). Further, we do not believe that box plots have the visual impact of our graphical method. In Figure 1, we compare figure 4 from our article with a box-plot depiction of the same data. The graph was produced from observed and imputed data, but our software was not able to produce box plots from percentile estimates that were derived from censored survival data. The box plot, however, does provide information on percentiles of the distribution, which is valuable.

We note that there are difficulties in producing box plots when the amount of censoring is high. For example, Lama and Gallo compared estimates of survival percentiles in a simulated two-arm trial with a hazard ratio of 0.75 and either 10% or 50% censored observations. Because of the censoring, high percentiles, such as the 95th percentile, are not estimable directly. Because not all percentile estimates are available, it is difficult to standardize the appearance of the box plot.

Because box plots in essence present summary statistics, to be most useful they should also present an indication of precision (eg, 95% confidence intervals). As far as we are aware, 95% confidence intervals are never presented, probably because it is impractical. As Lama and Gallo indicate, because of censoring estimates of upper percentiles of survival time (ie, long times) in box plots are very imprecise, but such information cannot easily be displayed on the plot. Our scatter plots do not have and do not need an indication of precision. They are the closest you can reasonably get to raw observations of survival time.

Our method is more flexible than box plots. A prime example is figure 5 from our article, which is a scatter plot of observed versus predicted survival times and which cannot reasonably be emulated by box plots. We believe that the ability to explore outcomes at the individual level (eg, by using such scatter plots) gives a much better feel for the data than only that provided by summary statistics and is indeed a useful complement to Kaplan–Meier plots.

**Figures**

![Figure 1](https://jnci.oxfordjournals.org/content/100/16/1188/F1.large.jpg)

**Figure 1.** Two depictions of the survival–time distribution (observed and imputed data) from the Medical Research Council RE01 trial in metastatic renal carcinoma. Data are as in figure 4 from Royston et al. (1). Three prognostic groups (good, moderate, and poor) are shown. Left: Data displayed as dot plots. Right: Identical data displayed as box plots, as produced by Stata (3).