Published in final edited form as:

Ann Surg. 2019 August ; 270(2): e48–e50. doi:10.1097/SLA.0000000000003115.

Comment on “Interpreting Clinical Benefits of Neoadjuvant Chemoradiation With Gemcitabine Versus Upfront Surgery in Patients With Borderline Resectable Pancreatic Cancer (BRPC)”

Zachary R. McCaw, AM, Lu Tian, ScD, Dae Hyun Kim, ScD, MD, Lee-Jen Wei, PhD

Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

Department of Biomedical Data Science, Stanford University School of Medicine, Palo Alto, CA

Institute For Aging Research, Hebrew SeniorLife, Boston, MA

Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

To the Editor:

Jang et al¹ conducted an important study among Korean patients with BRPC to investigate whether gemcitabine-based neoadjuvant chemoradiation treatment, followed by surgery, was superior to upfront surgery, followed by chemoradiation treatment, in terms of overall survival. The primary endpoint was survival at 2-years after randomization. There were 27 patients randomly allocated to neoadjuvant treatment, and 23 to upfront surgery. In the intention-to-treat analysis, the 2-year survival rates (2-YSRs) were 40.7% and 26.1% for neoadjuvant treatment and upfront surgery, respectively. The corresponding median survival times were 21 months and 12 months. Unfortunately, for neither survival summary measure did the authors report statistical analyses comparing the 2 arms. It is unclear if statistical significance was observed for either comparison. However, in the paper the authors concluded that neoadjuvant treatment was significantly better than upfront surgery with respect to the above 2 summary measures. The study was terminated at the interim analysis based on a superiority claim for neoadjuvant treatment. On the other hand, a statistical analysis was reported for the hazard ratio (HR). The observed HR of 1.97 (95% confidence interval [CI] 1.07–3.62, $P = 0.028$) significantly favored neoadjuvant treatment. The HR is a quite different summary of the treatment benefit from either the primary endpoint, 2-YSR, or the median survival time. A statistically significant HR does not imply a statistically significant difference in either the 2-YSR or the median survival time. Although the 2-YSR and median survival time have clinically meaningful interpretations, the HR may not.

The validity of HR analysis depends on the strong modeling assumption of proportional hazards (PH), which requires that the ratio of the hazard functions for the 2 arms is constant over time. When the PH assumption is not met, the estimated HR is difficult to interpret clinically.^{2–6} Based on visual inspection of the Kaplan-Meier (KM) curves displayed in Figure 2,¹ it appears that the PH assumption may not hold for the present study. This is

wei@hsph.harvard.edu.

The authors report no conflicts of interest.

especially likely in Figure 2B and C, where the KM curves of the 2 treatment arms cross. Even if the PH assumption is plausible, it is not clear how to clinically interpret a statistically significant HR. For example, an HR of 1.97 does not imply that upfront surgery increased mortality by 97%, as compared with neoadjuvant treatment, across the course of the study. Unlike the risk or mortality, the hazard is not a simple probability measure. Rather, it is the “force of mortality,” whose value does not have a straightforward clinical interpretation. Moreover, it is inappropriate to base treatment selection decisions on a single contrast, such as the HR, without a benchmark hazard value over the entire study duration. These issues and concerns have been discussed extensively.^{2–6} Clinically and statistically, the median survival time and the survival rate at a specific time point, such as the 2-YSR, are more interpretable than the HR, but only summarize local, not global, aspects of the survival profile.

The survival curves in Figure 2 provide all survival probabilities for each treatment arm across time. These are the most informative summaries of survivorship throughout the study. The higher a curve, the better the survival experience of patients in that arm. Intuitively, the area under the survival curve is a good overall summary of survivorship. In fact, this area is the expected survival time. Due to limited study follow-up, however, we generally cannot observe the entire survival curve. In Figure 2, the survival curves were observed up to 36 months. The area under the curve up to 36-months is the so-called restricted mean survival time (RMST).^{2–6} This statistic provides a global alternative to the event rate or median survival time for summarizing the survivorship in each treatment arm. Using reconstructed⁷ patient-level survival data from Figure 2A,¹ the RMSTs through 36 months of follow-up were 22.1 months for neoadjuvant therapy versus 15.5 months for upfront surgery. That is, on average, patients treated with neoadjuvant therapy and scheduled for 36 months of follow-up are expected to survive for 22.1 months. The difference in the RMSTs is 6.6 months, which significantly favors neoadjuvant therapy (95% CI: 1.2–12 mo, $P = 0.02$). In contrast to the HR, the validity of RMST point estimates and CIs does not require strong modeling assumptions, such as PH. Moreover, the 6.6-month increase in expected survival, across the first 36 months after baseline, is easier for clinicians and patients to interpret than is the reduction in hazard rate. It is interesting to note that, based on the reconstructed survival data, the difference in 2-YSR, comparing neoadjuvant therapy with upfront surgery, was 14.7% (95% CI: –11.5 to 40.8%, $P = 0.27$), which is not statistically significant. Neither was the difference in median survival times statistically significant. These local comparisons are likely underpowered. In contrast, the difference in RMSTs, which uses all data across the follow-up period, demonstrates a significant benefit from neoadjuvant treatment.

To examine whether the difference in expected survival time was influenced by restriction to 36 months of follow-up, we fit parametric Weibull models to the survival curves in Figure 2A.¹ Based on the Weibull fits, the expected survival times for the neoadjuvant and upfront surgery arms were 23.5 and 15.8 months, respectively. The difference in the (unrestricted) mean survival times was 7.6 months, which again significantly favors neoadjuvant therapy (95% CI: 1.2–14 months, $P = 0.02$). Results from this parametric analysis should be interpreted cautiously due to the extrapolations beyond the range of follow-up. On the other hand, this analysis suggests that, had we continued to observe the treatment arms beyond 36 months, the benefit attributable to neoadjuvant therapy would have continued to increase.

It is important to make quantitative summaries of the treatment difference more clinically interpretable so that clinicians and patients can use them for shared decision making when selecting from among alternative treatments.⁸ While the HR is not a translatable measure of the treatment effect, the difference or ratio of RMSTs is.⁶ The RMST is a useful, additional tool when analyzing time to clinical event data. Inference procedures for RMST are easily implemented using the R packages *surv2sampleComp* and *survRM2*, available on CRAN (<https://cran.r-project.org/>).

REFERENCES

1. Jang JY, Han Y, Lee H, et al. Oncological benefits of neoadjuvant chemoradiation with gemcitabine versus upfront surgery in patients with borderline resectable pancreatic cancer: a prospective, randomized, open-label, multicenter phase 2/3 trial. *Ann Surg* 2018; 268:215–222. [PubMed: 29462005]
2. Royston P, Parmar MK. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Stat Med* 2011; 30:2409–2421. [PubMed: 21611958]
3. Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *J Clin Oncol* 2014; 32:2380–2385. [PubMed: 24982461]
4. A'Hern RP. Restricted mean survival time: an obligatory end point for time-to-event analysis in cancer trials? *J Clin Oncol* 2016; 34:3474–3476. [PubMed: 27507871]
5. Pak K, Uno H, Kim DH, et al. Interpretability of cancer clinical trial results using restricted mean survival time as an alternative to the hazard ratio. *JAMA Oncol* 2017; 3:1692–1696. [PubMed: 28975263]
6. McCaw ZR, Orkaby AR, Wei LJ, et al. Applying evidence-based medicine to shared decision making: value of restricted mean survival time. *Am J Med* 2018; [Epub ahead of print].
7. Guyot P, Ades AE, Ouwens MJ, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012; 12:9. [PubMed: 22297116]
8. Merchant FM, Dickert NW, Howard DH. Mandatory shared decision making by the Centers for Medicare & Medicaid Services for cardiovascular procedures and other tests. *JAMA* 2018; [Epub ahead of print].