Published in final edited form as:

JAMA Oncol. 2018 August 01; 4(8): 1137–1138. doi:10.1001/jamaoncol.2018.1087.

Estimating and Interpreting the Overall Survival Benefit of Checkpoint Inhibitors via Meta-analysis

Hajime Uno, PhD,

Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts

Lu Tian, ScD,

Department of Health Research and Policy, Stanford University School of Medicine, Palo Alto, California

Lee-Jen Wei, PhD

Department of Biostatistics, Harvard University, Boston, Massachusetts

To the Editor Lee et al¹ conducted an interesting meta-analysis to estimate the relative efficacy of checkpoint inhibitor vs docetaxel for treatment of advanced non-small cell lung carcinoma. The meta-analysis consists of 5 comparative clinical trials (CheckMate-017, CheckMate-057, Keynote-010, OAK, POPLAR) with the overall survival (OS) end point. For each study, the hazard ratio (HR) was used to quantify the treatment effect. A weighted average of 5 HRs was constructed as the pooled treatment effect from checkpoint inhibitors using the fixed-effects inverse-variance-weighted method. This resulted in a combined HR (checkpoint inhibitor vs docetaxel) of 0.69 (95% CI, 0.63-0.75).

There are a couple of issues regarding this meta-analysis. First, except for CheckMate-017, checkpoint inhibitors had delayed clinical OS benefit. That is, Kaplan-Meier curves for 2 treatment groups in each trial overlapped considerably for the early part of the study. Thus, the HR was not a constant over the entire study follow-up time. For this situation, it would be difficult to interpret individual HRs clinically and the HR estimate would not be an appropriate measure to quantify the OS benefit from checkpoint inhibitor use.^{2,3} Second, even when the HR was constant over time for each study, one would not be able to identify a meaningful patient population for which the aforementioned pooled estimate of 0.69 could be interpreted as its HR unless those 5 underlying HRs are identical (an unlikely scenario).⁴

For a single study, a robust alternative summary for the between-treatment difference in OS could be the difference of 2 survival rates or restricted mean survival times (RMST) at a specific time point.^{2,3} For example, for CheckMate-057, RMSTs for OS with 24-month follow-up were 13.0 and 11.3 months for nivolumab and docetaxel, respectively.³ That is, patients treated with nivolumab would survive an average of 13.0 months with 24-month follow-up. The difference of the 2 RMSTs is 1.7 months. Unlike HR, no model assumption is required for this clinically interpretable summary to quantify the between-group

Corresponding Author: Lee-Jen Wei, PhD, Department of Biostatistics, Harvard University, 655 Huntington Ave, Boston, MA 02115 (wei@hsph.harvard.edu).

Conflict of Interest Disclosures: None reported.

difference. One can then use a weighted average of such differences from 5 studies with weights, for example, proportional to study size, to quantify the treatment effect. This pooled measure would be the RMST difference estimate for a welldefined population that is a mixture of 5 study populations with weights proportional to study size.^{4,5(pp25-27)} One may also choose other weights to create a mixture population from 5 study populations to match a prespecified target population and obtain a corresponding pooled treatment effect estimate.

4

References

1. Lee CK, Man J, Lord S, et al. Clinical and molecular characteristics associated with survival among patients treated with checkpoint inhibitors for advanced non-small cell lung carcinoma: a systematic review and meta-analysis. *JAMA Oncol.* 2018;4(2):210–216. doi:10.1001/jamaoncol.2017.4427 [PubMed: 29270615]
2. 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(22):2380–2385. doi:10.1200/JCO.2014.55.2208 [PubMed: 24982461]
3. 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(12):1692–1696. doi: 10.1001/jamaoncol.2017.2797 [PubMed: 28975263]
4. Hasegawa T, Claggett B, Tian L, et al. The myth of making inferences for an overall treatment efficacy with data from multiple comparative studies via meta-analysis. *Stat biosci.* 2017;9(1):284–297. doi:10.1007/s12561-016-9179-3
5. Everitt BS, Hand DJ. *Finite Mixture Distributions*. London, England: Chapman and Hall; 1981. doi: 10.1007/978-94-009-5897-5