Editorial Commentary

Appropriate statistical methods are available to handle biases encountered in blinded, independent, central review (BICR) determined progression-free survival

Jessica A. Lavery, Katherine S. Panageas

Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Correspondence to: Katherine S. Panageas, DrPH. Attending Biostatistician, Memorial Sloan Kettering Cancer Center, 485 Lexington Ave. 2nd Fl. New York, NY 10017, USA. Email: panageak@mskcc.org.

Comment on: Stone A, Gebski V, Davidson R, *et al.* Exaggeration of PFS by blinded, independent, central review (BICR). Ann Oncol 2019;30:332-8.

Received: 26 March 2019; Accepted: 08 April 2019; Published: 09 April 2019.

doi: 10.21037/jhmhp.2019.04.01

View this article at: http://dx.doi.org/10.21037/jhmhp.2019.04.01

While overall survival (OS) has traditionally been the standard evaluation for a new treatment in oncology since it is easily obtained and unambiguous, the endpoint of progression-free survival (PFS) is appealing due to the shortened observation time required to determine treatment efficacy, smaller sample size requirements and no confounding due to subsequent treatments. For many cancers, PFS has been demonstrated to be a valid measure of surrogacy for OS and an acceptable trial endpoint from regulatory agencies (1). FDA approval for many therapies including sorafenib for renal cell carcinoma, gemcitabine for ovarian cancer, and rituximab for non-Hodgkin's lymphoma were based on PFS (1-4). However, PFS can be associated with measurement error and bias. Radiologic scans are the primary mode of assessment for solid tumors, and the Response Evaluation Criteria in Solid Tumors (RECIST) criteria were developed to define changes to a scan that constitute a progression (5). Because of the potential for discrepancy in interpretation of RECIST criteria across radiologists, a process of blinded, independent, central review (BICR) was established in order to attempt to provide reliable and unbiased PFS assessments (6). Furthermore, BICRs are generally recommended for clinical trials submitted to the FDA for regulatory consideration(7).

Intended to mitigate assessment variability among local site evaluations, BICR-based analyses potentially introduce other biases, resulting from varying evaluation times or differential attrition rates between study arms, interval censoring, and informative censoring (8-11). BICR is typically performed retrospectively for the purposes of quality control across radiologic assessments rather than for individual treatment decisions (12). Typically, if a patient is deemed to have progressed by local evaluation, this triggers a sequence of events: the patient is off treatment, off protocol, and will not undergo additional scans. If the BICR cannot confirm the locally determined progression, the FDA has recommended that these cases be censored at the time of local progression for the BICR analysis of PFS (7). This then violates the independent censoring assumption required for standard survival analyses since patients in the BICR analysis are considered lost to follow-up for reasons related to the study and are then not representative of all censored observations (13,14). In the recent article entitled "Exaggeration of PFS by blinded, independent, central review (BICR)", Stone *et al.* addressed the impact of this informative censoring on Kaplan-Meier (KM) estimates of median progression-free survival (15).

Stone *et al.* present a simulation study considering various scenarios of true progression times and correlations in timing between local and BICR identification of progression under the scenario of a 12-week radiographic imaging assessment schedule. The authors demonstrate scenarios under which KM estimates of median PFS are both under- and over-estimated due to informative censoring and conclude that KM estimates of median survival are biased even under the scenario in which local and BICR PFS times are identical (i.e., local review is sufficiently standardized). Given the interval censoring inherent in studies of PFS, the bias resulting from informative censoring in BICR

analyses should instead be assessed utilizing appropriate analytic methods that take this into account, such as the nonparametric extension of the KM estimator (11,16-18). Since the traditional KM estimate of median PFS ignores interval censoring, the observed bias cannot be attributed entirely to informative censoring.

In general, methods that appropriately account for interval censoring should be the standard analytic approach for analyses of PFS (19). In studies of PFS that utilize BICR, the local estimate of PFS may be biased due to lack of standardization across radiologic reviewers and the BICR estimate of PFS may be biased due to informative censoring. We recommend that estimates from both analyses be presented and sensitivity analyses should be conducted in the BICR analyses to assess the potential impact of interval censoring. As indicated by Stone *et al.*, potential analyses include inverse probability weighting and multiple imputation (20,21). An additional option is to consider sensitivity analyses at the extremes, where patients censored due to local progression can be considered to progress at the time of local progression or after all other patients in the sample. Though extreme, this provides estimates of the range of potential impact (22). In the past, implementing these more complex statistical methods may have been challenging due to lack of available software, but in recent years numerous resources have been developed to implement these analyses (23).

While Stone *et al.* demonstrated bias in KM estimates of the median survival, phase III trials are generally intended to assess treatment efficacy, often via a hazard ratio (HR) comparing PFS in the treatment arm to the control arm. Prior meta-analyses as well as the case study within the article by Stone *et al.* have demonstrated that the estimate of HR and it's 95% confidence interval are consistent for analyses based on BICR and local review (24-26). Though the estimate of treatment efficacy is unbiased, the biased KM median PFS is often additionally reported. Thus, the above recommendations are pertinent only when median survival is reported.

Beyond modifications to the statistical method employed, a potential remedy that has been proposed is to move from retrospective BICR to real-time BICR. However, this presents a costly and logistically challenging solution that is potentially unnecessary after appropriately accounting for the interval and informative censoring (8).

In an era of increased awareness of the importance of reproducibility, we recommend that studies of PFS that utilize BICR implement rigorous analytic approaches and present sensitivity analyses when informative censoring mechanisms are potentially violated.

Acknowledgments

Funding: Authors acknowledge Memorial Sloan Kettering Cancer Center Support Grant/Core Grants (P30 CA008748).

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Section Editor Jianrong Zhang (MPH Candidate, George Warren Brown School; Graduate Policy Scholar, Clark-Fox Policy Institute, Washington University in St. Louis, St. Louis, USA).

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jhmhp.2019.04.01). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Gill S, Berry S, Biagi J, et al. Progression-free survival as a primary endpoint in clinical trials of metastatic colorectal cancer. Curr Oncol 2011;18 Suppl 2:S5-10.
2. Escudier B, Eisen T, Stadler WM, et al. Sorafenib in advanced clear-cell renal-cell carcinoma. N Engl J Med 2007;356:125-34.
3. Pfisterer J, Plante M, Vergote I, et al. Gemcitabine plus carboplatin compared with carboplatin in patients with platinum-sensitive recurrent ovarian cancer: an intergroup

trial of the AGO-OVAR, the NCIC CTG, and the EORTC GCG. J Clin Oncol 2006;24:4699-707.
4. Hochster H, Weller E, Gascoyne RD, et al. Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: results of the randomized phase III ECOG1496 Study. J Clin Oncol 2009;27:1607-14.
5. Duffaud F, Therasse P. New guidelines to evaluate the response to treatment in solid tumors. Bull Cancer 2000;87:881-6.
6. Dancey JE, Dodd LE, Ford R, et al. Recommendations for the assessment of progression in randomised cancer treatment trials. Eur J Cancer 2009;45:281-9.
7. Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. United States Food and Drug Administration. 2007;72:27575-6.
8. Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? J Clin Oncol 2008;26:3791-6.
9. Amit O, Bushnell W, Dodd L, et al. Blinded independent central review of the progression-free survival endpoint. Oncologist 2010;15:492-5.
10. Zeng L, Cook RJ, Wen L, et al. Bias in progression-free survival analysis due to intermittent assessment of progression. Stat Med 2015;34:3181-93.
11. Panageas KS, Ben-Porat L, Dickler MN, et al. When you look matters: the effect of assessment schedule on progression-free survival. J Natl Cancer Inst 2007;99:428-32.
12. Ford R, Schwartz L, Dancey J, et al. Lessons learned from independent central review. Eur J Cancer 2009;45:268-74.
13. Ranganathan P, Pramesh CS. Censoring in survival analysis: Potential for bias. Perspect Clin Res 2012;3:40.
14. DuBois R, Berry D, Doroshow J, et al. Conference on Clinical Cancer Research, September 2008. PANEL 2: Improved Insights into Effects of Cancer Therapies. 2008:11-8.
15. Stone A, Gebski V, Davidson R, et al. Exaggeration of PFS by blinded, independent, central review (BICR). Ann Oncol 2019;30:332-8.
16. Odell PM, Anderson KM, D'Agostino RB. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. Biometrics 1992;48:951-9.
17. Lindsey JC, Ryan LM. Tutorial in biostatistics methods for interval-censored data. Stat Med 1998;17:219-38.
18. Peto R. Experimental Survival Curves for Interval-Censored Data. J Royal Stat Soc 1973;22:86-91.
19. Qi Y, Allen Ziegler KL, Hillman SL, et al. Impact of disease progression date determination on progression-free survival estimates in advanced lung cancer. Cancer 2012;118:5358-65.
20. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 2000;56:779-88.
21. Hsu CH, Taylor JM. Nonparametric comparison of two survival functions with dependent censoring via nonparametric multiple imputation. Stat Med 2009;28:462-75.
22. Allison PD. Survival Analysis Using SAS: A Practical Guide, Second Edition. SAS Institute; 2010.
23. Bogaerts K, Komarek A, Lesaffre E. Survival Analysis with Interval-Censored Data: A Practical Approach with Examples in R, SAS, and BUGS. CRC Press; 2017.
24. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. Eur J Cancer 2011;47:1772-8.
25. Zhang J, Zhang Y, Tang S, et al. Systematic bias between blinded independent central review and local assessment: literature review and analyses of 76 phase III randomised controlled trials in 45 688 patients with advanced solid tumour. BMJ Open 2018;8:e017240.
26. Zhang JJ, Chen H, He K, et al. Evaluation of Blinded Independent Central Review of Tumor Progression in Oncology Clinical Trials: A Meta-analysis. Ther Innov Regul Sci 2013;47:167-74.

doi: 10.21037/jhmhp.2019.04.01
Cite this article as: Lavery JA, Panageas KS. Appropriate statistical methods are available to handle biases encountered in blinded, independent, central review (BICR) determined progression-free survival. J Hosp Manag Health Policy 2019;3:8.

