Editorial

The search for surrogate endpoints for immunotherapy trials

Marc Buyse[1,2], Tomasz Burzykowski[2,3], Everardo D. Saad[3]

[1]International Drug Development Institute, San Francisco, CA, USA; [2]Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat), Hasselt University, Hasselt, Belgium; [3]International Drug Development Institute, Louvain-la-Neuve, Belgium

Correspondence to: Marc Buyse, ScD. International Drug Development Institute, 775 Francisco St, San Francisco, CA, USA. Email: marc.buyse@iddi.com.

Provenance: This is a Guest Editorial commissioned by Section Editor Jianrong Zhang, MD (Candidate of Master of Public Health, George Warren Brown School of Social Work; Graduate Policy Scholar-in-training, Clark-Fox Policy Institute, Washington University in St. Louis, St. Louis, USA).

Comment on: Mushti SL, Mulkey F, Sridhara R. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. Clin Cancer Res 2018;24:2268-75.

Submitted Apr 23, 2018. Accepted for publication May 04, 2018.

doi: 10.21037/atm.2018.05.16

View this article at: http://dx.doi.org/10.21037/atm.2018.05.16

Mushti and colleagues from the Food and Drug Administration (FDA) have used patient-level data from 13 trials comparing immunotherapeutic agents with standard treatment in various cancer types in an attempt to validate objective response (OR) and progression-free survival (PFS) as surrogate endpoints for overall survival (OS) (1). This is an important issue, given the current and future role of immunotherapy in several cancer types. To this effect they have used a "two-level" approach to investigate whether (I) each of these two potential surrogates was associated with OS in individual patients, adjusting for the randomized treatment; and (II) the effect of immunotherapy on each of these two potential surrogates could be used to reliably predict the effect of immunotherapy on OS. Both questions are of interest: the former for patient management (with the surrogate being a marker of prognosis), and the latter for drug development (with use of the surrogate potentially gaining months or years of development time). In the latter respect, immunotherapy may be a special case in the treatment of metastatic cancer, because at least in phase 3 trials of checkpoint inhibitors (CPIs) the prevailing view is that OS should continue to be the primary endpoint, notwithstanding its limitations in clinical trials for other treatment modalities. We therefore discuss the results of Mushti *et al.* in their own right and in contrast to other treatment modalities.

The results of Mushti *et al.*'s analyses are interesting, and broadly comparable to those obtained in most metastatic settings with drugs other than immunotherapies (2). The patient-level analyses revealed associations between the two surrogates and OS: specifically, OR had a major prognostic impact on OS, with a hazard ratio (HR) =0.14 favoring responders (95% CI: 0.12–0.16, Figure 1A), while the patient-level correlation between PFS and OS was ρ=0.61 (95% CI not provided). In contrast, the trial-level analyses revealed no association between the effects of immunotherapy on either surrogate and on OS: specifically, the trial-level coefficient of determination was R^2=0.13 (95% CI not provided) for both PFS (Figure 2A) and OR (Figure 2B). It is worth noting that 95% CIs were not provided for correlation and determination coefficients, which limits the interpretation of these results.

Had OR or PFS been acceptable surrogates, one would have expected a positive correlation between the effects of treatment on PFS and OS (such that treatments that decrease the PFS hazard rate also decrease the OS hazard rate), and a negative correlation between the effects of treatment on OR and OS (such that treatments that increase the OR probability decrease the OS hazard rate). Should we take these results as a confirmation that OS should continue to be the primary endpoint at least with CPIs, or should we seek alternative explanation for these findings? For example, are there potential confounders that blunt the associations at the trial level, such as cross-over to CPIs in some of the trials? Since we cannot answer these questions yet, it seems a sensible course of action to continue to use OS as (one of the) primary endpoint(s) in immunotherapy trials.

At face value, the discrepancy between patient-level and trial-level results seems paradoxical: indeed, if response to treatment predicts OS, and if immunotherapy improves

OR rates, why do improvements in OR rates not predict corresponding improvements in OS? One explanation might be that response to treatment might merely capture prognostic information over and above that measured at baseline. Of note, Figure 1 in Mushti *et al.* (1) did not use a landmark approach to avoid guarantee-time bias (3), and did not adjust the OS comparison for prognostic factors measured at baseline, so the difference between the OS curves of responders and non-responders may have been overestimated. Even so, there is little doubt that OR is a powerful prognostic factor for OS. But another story may lie behind the curves of Figure 1 in Mushti *et al.* (1). The PFS curves (Figure 1B) show a much better outcome for responders in the experimental arm than in the control arm, and a slightly worse outcome for non-responders in the experimental arm than in the control am. This suggests that the experimental treatment "pushed" more patients of good prognosis to OR than did the control arm. In contrast, the OS curves (Figure 1A) show a much better outcome for responders in the experimental arm than in the control arm, and also a slightly better outcome during the entire observation period for non-responders in the experimental arm than in the control arm. This suggests that both responders and non-responders have benefited from the experimental treatment, so that the observed survival benefit is not (entirely) explained by the higher OR rates obtained with the experimental treatment. In other words, different mechanisms may be at play for the effect of immunotherapy on response and on OS, such that an improvement in OR rate does not causally induce a corresponding improvement in OS. Techniques of causal inference, in particular mediation analysis, can be used to estimate the causal effect of OR on OS with adjustment for baseline prognostic factors. Such an analysis was carried out, for instance, in patients with resectable breast cancer receiving neo-adjuvant chemotherapy (4). The causal analysis confirmed the results of a previous meta-analysis, also carried out under the auspices of the FDA, in showing that complete pathological response was an individual-level but not a trial-level surrogate for OS (5). One limitation of the causal inference approach is that it rests on the very strong and untestable assumption that there are no unmeasured confounders, an assumption that is also required for the Prentice criteria to be applicable (6). If one is not willing to make such an assumption, then the meta-analytic approach can be used, with a strong association between the treatment effects at the trial level suggesting that a claim of surrogacy is plausible. A mediation analysis of the immunotherapy datasets analyzed by Mushti *et al.* would nicely complement and likely confirm their conclusions that neither OR nor PFS are acceptable trial-level surrogates for OS.

Mushti *et al.* (1) point out that their analyses combine data from patients with different tumor types, and as such may have caused real associations in some specific tumor types or patient subsets to have been missed. While this is theoretically possible, there is accumulating evidence in a wide range of settings that OR and PFS are generally not acceptable surrogates for OS in patients with metastatic disease (2). At the same time, there is good evidence from patient-level meta-analyses that in the adjuvant setting, DFS may be a valid surrogate for OS across a range of tumor types, including colorectal cancer, gastric cancer, non-small cell lung cancer and head and neck squamous cell cancer (2,6). Likewise, in patients with follicular lymphoma, an indolent disease, a patient-level meta-analysis showed that complete response rate at 30 months after initiation of induction therapy could be used as a surrogate for PFS (7). Hence it is possible to identify surrogate endpoints in cancer, using a meta-analytic approach based on patient-level data.

Mushti *et al.* (1) also stress that their analyses are limited by the fact that all trials showed positive effects of the immunotherapies tested (otherwise the trials would not have been submitted to the agency for approval). This is, indeed, a drawback of using only registration trials to validate surrogate endpoints, since the trial-level correlation analysis requires a wide range of treatment effects on both the surrogate and OS to be informative and reliable. HRs ranged from 0.43 to 0.91 for PFS and from 0.42 to 0.86 for OS. The odds ratios for OR ranged from 1.05 to 13.6. Hence the meta-analysis included few trials with a small or no treatment effect, and none favoring the control group over the experimental group. In fact, the regression lines of the trial-level analyses (Figures 2,3), in addition to showing no association between the treatment effects on the surrogates and OS, do not pass close to the origin (1,1), a condition required of a valid surrogate, for which no treatment effect on the surrogate would also imply no treatment effect on OS (8). The absence of trials with no treatment effect or a negative treatment effect is unlikely to be the reason why these analyses failed to identify acceptable surrogates for OS in trials of immunotherapy. This limitation could be addressed—albeit at the cost of collecting further data from negative trials. In an ideal world, such data should be readily available for re-analysis,

in order to avoid potentially misleading conclusions based on a biased subset of trials showing statistically significant results on at least one of the endpoints considered. We commend the FDA for conducting and publishing thoughtful surrogacy analyses, but we contend that these analyses are not definitive and that data from all randomized clinical trials should be available to allow the agency and other interested parties to conduct similar analyses using the totality of the evidence (9).

Acknowledgements

None.

Footnote

Conflict of Interest: The authors have no conflicts of interest to declare.

References

1. Mushti SL, Mulkey F, Shridhara R. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. Clin Cancer Res 2018;24:2268-75.
2. Prasad V, Kim C, Burotto M, et al. The strength of association between surrogate end points and survival in oncology. A systematic review of trial-level meta-analyses. JAMA Intern Med 2015;175:1389-98.
3. Giobbie-Hurder A, Gelber RD, Regan MM. Challenges of guarantee-time bias. J Clin Oncol 2013;31:2963-9.
4. Vandenberghe S, Duchateau L, Slaets L, et al. Surrogate marker analysis in cancer clinical trials through time-to-event mediation techniques. Stat Methods Med Res 2017. [Epub ahead of print].
5. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014;384:164-72.
6. Buyse M, Molenberghs G, Paoletti X, et al. Statistical evaluation of surrogate endpoints with examples from cancer trials. Biom J 2016;58:104-32.
7. Shi Q, Flowers CR, Hiddemann W, et al. Thirty-Month Complete Response as a Surrogate End Point in First-Line Follicular Lymphoma Therapy: An Individual Patient-Level Analysis of Multiple Randomized Trials. J Clin Oncol 2017;35:552-60.
8. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Stat Med 1997;16:1965-82.
9. Bertagnolli MM, Sartor O, Chabner BA, et al. Advantages of a truly open-access data-sharing model. N Engl J Med 2017;376:1178-81.

Cite this article as: Buyse M, Burzykowski T, Saad ED. The search for surrogate endpoints for immunotherapy trials. Ann Transl Med 2018;6(11):231. doi: 10.21037/atm.2018.05.16

