Surrogate clinical endpoints to predict overall survival in non-small cell lung cancer trials—are we in a new era?

Jeffrey M. Clarke, Xiaofei Wang, Neal E. Ready

Department of Medicine, Duke Cancer Institute, Duke University Medical Center, Durham, NC, USA
Correspondence to: Jeffrey M. Clarke, MD. Department of Medicine, Duke Cancer Institute, Duke University Medical Center, DUMC 3841, Durham, NC 27710, USA. Email: jeffrey.clarke@duke.edu.

Abstract: Surrogate endpoints for clinical trials in oncology offer an alternative metric for measuring clinical benefit, allowing for shorter trial duration, smaller patient cohorts, and single arm design. The correlation of surrogate endpoints with overall survival (OS) in therapeutic studies is a central consideration to their validity. The Food and Drug Administration (FDA) recently published an analysis of fourteen clinical trials in advanced non-small cell lung cancer (NSCLC), and discovered a strong association between response rate and progression free survival. Furthermore, a correlation between response rate and OS is demonstrated when analyzing the experimental treatment arm separately, minimizing bias from patient crossover. We also highlight multiple, important considerations when using response as an endpoint in clinical trials involving NSCLC patients.

Keywords: Lung cancer; surrogate; endpoint; response; survival

During the past 2 years, the Food and Drug Administration (FDA) has approved no less than five new targeted agents for the treatment of metastatic non-small cell lung cancer (NSCLC). Furthermore, novel 3rd generation EGFR inhibitors (including CO-1686 and AZD9291) and new indications for checkpoint immunotherapies (nivolumab and pembrolizumab) are anticipated to become available in the near future for the treatment of patients with NSCLC. The development of these new therapies has spurred renewed interest and debate on the topic of surrogate clinical endpoints for clinical trials in NSCLC. While overall survival (OS) remains the gold-standard metric of benefit for clinical trials involving therapeutic agents, there are multiple limitations to using this as an endpoint. Foremost, successive lines of therapy, patient crossover, and increased post-progression survival can all mask or 'dilute' treatment effects (1). Given the urgent need for new cancer therapies, the FDA offers an accelerated approval mechanism that is based on surrogate endpoints reasonably likely to predict clinical benefit. Surrogate endpoints including objective response rate (ORR) and progression-free survival (PFS)

potentially offer more feasible options to measure clinical benefit, allowing for shorter trial duration, smaller patient cohorts, and single arm design. The correlation of ORR and PFS endpoints with OS remains a central consideration in determining their validity.

Recently, the Office of Hematology and Oncology Products at the FDA investigated the association of ORR and PFS with OS in patients with advanced NSCLC. Blumenthal *et al.* published a meta-analysis using both trial and patient level data from fourteen clinical trials submitted to the FDA from 2003 to 2013, including over 12,000 patients (2). Importantly, the clinical trials used in this analysis included three smaller randomized studies of targeted therapies in ALK rearrangement or EGFR mutation positive patients. The remaining eleven trials were larger and incorporated mainly unselected treatment-naïve or refractory lung cancer patients. Using weighted linear regression methods, the authors compared the association of hazard ratios of PFS and OS with ORR odds ratio. Interestingly, Blumenthal and colleagues found a strong association between ORR and PFS ($R^2=0.89$), but
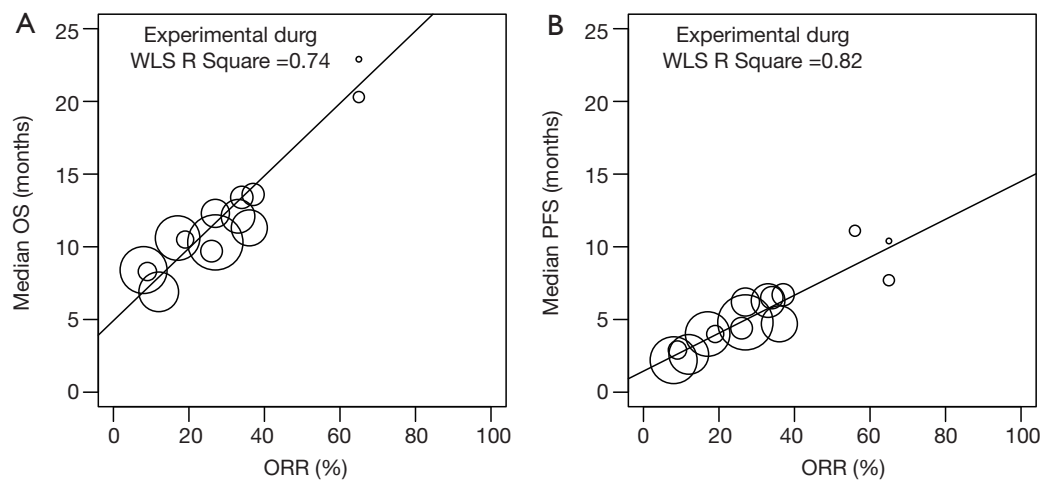
Figure 1 Weighted, linear correlation between ORR and OS (A) and PFS (B) in the experimental treatment arms of phase III clinical trials in metastatic NSCLC. ORR, objective response rate; OS, overall survival; PFS, progression-free survival; NSCLC, non-small cell lung cancer.

no association between OS and ORR or OS and PFS using trial data. A responder analysis using patient-level data was also performed demonstrating a strong association between response and PFS and OS independent of treatment arm. Based on these data, the authors conclude that therapies with a significant ORR would likely have a large impact on PFS. Importantly, this is the first analysis demonstrating ORR as a strong predictor of clinical benefit in NSCLC.

We applaud the authors and the FDA for publishing their experience over the past decade and addressing the important topic of clinical endpoints for NSCLC trials. While the authors discovered a strong association of ORR with PFS and OS in patient-level data, post-treatment bias with crossover and post-progression survival are acknowledged as likely explanations for the lack of trial-level association between ORR and OS. Indeed, others have also attempted to correlate PFS and OS in phase III clinical trials using molecularly targeted therapies without success secondary to significant treatment crossover effect (3). We agree with this explanation, given the trial level analysis were performed using hazard ratios and would be potentially subject to post-progression bias. However, we postulate that a strong relationship between ORR and OS might be demonstrated by controlling for potential patient crossover. Analysis of ORR and OS in the experimental treatment arm alone would not depend on outcomes in the control arm of the study. Therefore, this analysis would eliminate potential bias from crossover between treatment arms in the trials, while also minimizing bias from post-progression survival in

the control arm.

A paradigm of interest as a pathway to FDA approval for novel therapy for uncommon biomarker defined subsets of cancer is single arm trials with a primary endpoint of ORR, and/or PFS. We performed an analysis of the association between ORR, PFS and OS of the experimental treatment arms of the trials analyzed by Blumenthal *et al*. in order to simulate a series of single arm trials and minimize the confounder of crossover therapy. We performed a weighted linear correlation of the trial level data reported by Blumenthal *et al*. using ORR with both PFS and OS as measured in months. The analysis was performed using data from the experimental arm separately from the control arm of the studies. Indeed, we discovered a strong linear correlation with ORR and PFS ($R^2=0.82$), as well as with OS ($R^2=0.74$) by including patients only in the experimental treatment arms (*Figure 1*). Furthermore, no association was found between ORR and OS, when the same analysis was performed on patients in the control/standard care arms (data not shown). These data suggest that when controlling for post-progression treatment and especially crossover to effective molecular therapy such as in EGFR mutated lung cancer, a strong correlation exists between ORR and OS, thus further supporting ORR as a robust surrogate predictor of OS and clinical benefit.

There are multiple important implications based on the findings by Blumenthal *et al*. regarding the use of ORR as a surrogate clinical endpoint to predict survival. Firstly, given the shifting molecular landscape of oncology, particularly in

the arena of NSCLC, the use of OS as the primary endpoint for large randomized clinical trials may be prohibitive. Simple histologic classification of lung cancers is insufficient for modern patient care; multilocus genotyping of oncogenic drivers by next generation sequencing is now routinely incorporated into standard care in many clinical settings, resulting in small, narrowly-defined patient populations. For example, data published in 2014 by the Lung Cancer Mutation Consortium of patients with advanced lung adenocarcinoma found ten different actionable, oncogenic driver alterations in 64% of samples (4). While KRAS, EGFR, and ALK alterations were most common, ERBB2, BRAF, PIK3CA, and MET alterations among several others were also detected, each with frequency <5%. In order to conduct clinical trials of targeted therapeutic agents in these populations, smaller single arm clinical trials are most feasible, while accrual for large randomized controlled trials from rare patient populations may be problematic. Furthermore, historical control data on PFS and OS in these molecularly defined subgroups is either nonexistent or exceedingly challenging to obtain. In this context, the use of ORR and/or PFS as a clinical endpoint is advantageous over alternative clinical endpoints. Notable examples recently in lung cancer include both crizotinib and ceritinib which were granted accelerated approval by the FDA for the treatment of ALK rearrangement positive NSCLC based on ORR in single-arm, phase II and I studies, respectively (5,6). Finally, molecular mechanisms modulating TKI resistance in EGFR mutation positive lung cancer have been well described. These mechanism involve commonly the T790M mutation of EGFR, along with MET or ERBB2 amplification, further defining small patient populations in the refractory setting (7). The FDA has granted Breakthrough Designation within the past year to two agents, CO-1686 and AZD9291, demonstrating high response rates approaching 60% in patients harboring the T790M mutation following standard anti-EGFR treatment in phase I clinical trials (8,9). These examples demonstrate the challenges in conducting clinical trials with the endpoint of OS and using small patient populations based on molecular profiles, and highlight the role of ORR as an important surrogate endpoint.

The recent use of umbrella clinical trials in lung cancer adds further debate regarding optimal clinical endpoints in the clinical trials involving novel, targeted therapeutics. For example, the Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE)-1 trial utilized an adaptive randomization protocol based on biomarker positivity to one of four treatment arms with

targeted agents (10). The primary endpoint of this study was 8-week disease control rate (comprised of response plus stable disease) and the trial enrolled 255 patients. Indeed, this trial has contributed to understanding the efficacy of several targeted agents in small, biomarker-defined populations (10,11). BATTLE-2 is currently evaluating agents targeting EGFR, PI3K/AKT and MEK with the same endpoint of 8-week disease control rate (12). Most recently, the Lung Master Protocol (Lung-MAP) is a multi-arm, biomarker driven clinical trial evaluating novel therapies in patients with advanced lung squamous cell carcinoma (13). Biomarker positive patients are randomized within separate cohorts to experimental therapy or standard care. With the exception of PDL1, initial estimates of the frequency of each biomarker are less than 15%. Contrary to the BATTLE protocols, the primary endpoint chosen for the phase II and III components of Lung-MAP are co-primary PFS and OS. Accordingly, 2,500-5,000 patients approximately will need to be enrolled over the five treatment cohorts (13). The key advantage of this structure is the ability to move directly to registration with the FDA by industry partners from phase III. However, this design is logistically complex and requires an immense accrual population with enrollment open at hundreds of clinical sites (13).

The success of immune checkpoint agents in oncology over the past five years also highlights multiple complexities in using clinical endpoints, such as OS, with immunotherapies. Most recently, nivolumab (anti-PD1 monoclonal antibody) was approved by the FDA for the treatment of refractory lung squamous cell carcinoma after demonstrating superior OS compared to docetaxel in a phase III trial (Checkmate 017) and safety in a single arm phase II trial (Checkmate 063) (14). Likewise, pembrolizumab has been granted Breakthrough Therapy Designation by the FDA after demonstrating safety and promising efficacy in an expanded phase I clinical trial (Keynote-001) (15). While response rates have ranged from approximately 15-45% depending on PDL1 expression, there is accumulating evidence demonstrating prolonged duration of response in many patients. Therefore, while impact on OS or PFS maybe modest in large clinical trials, a subset of patients can have dramatic and durable clinical benefit. Further complicating the issue, novel patterns of immunologic response are well described, including pseudo-progression, durable stable disease, and late response (16). These clinical observations have necessitated the development of immune related Response Evaluation Criteria in Solid Tumors (irRECIST) criteria to account for variable tumor regression

patterns (17). Finally, delayed separation of survival curves is commonly encountered in immunotherapy trials and can lead to underestimate of survival benefit if measured prematurely (16). Early assessment of PFS and OS may therefore be suboptimal for analyzing benefit in novel immunotherapy trials. Further data is required to assess the correlation of response with prolonged response duration to survival in both biomarker-selected and unselected patient populations treated with immune checkpoint therapies.

While prior analyses leveraging response as a predictor of survival in NSCLC are limited, multiple publications have explored the correlation of PFS or disease free survival (DFS) with OS. In 2013, the Surrogate Lung Project Collaborative Group published a re-analysis of six meta-analyses, including 60 randomized trials and 15,000 patients (18). The analysis included trials employing both single and combined modality therapy in the adjuvant setting and for patients with locally advanced lung tumors. The authors reported a strong correlation between DFS and OS in the adjuvant setting, irrespective of whether radiotherapy was used in conjunction with chemotherapy. Additionally, PFS strongly correlated with OS in patients with locally advanced disease using both patient- and trial-level data. Notably, the trials included in this analysis involved only cytotoxic chemotherapy agents, and the authors caution against extrapolating to targeted therapies in different diseases settings. In contrast, in treatment naïve patients with advanced lung cancer, Laporte *et al*. performed a meta-analysis of five randomized clinical trials using cytotoxic chemotherapy agents and found only a modest association of PFS and OS (19). Finally, many investigators have advocated for the use of major pathologic response in the neoadjuvant setting of resectable lung cancer, given a reported strong association of pathologic complete response with OS (20). Together, these studies highlight the importance of disease and treatment setting with surrogate clinical endpoints, and the need for validation.

Surrogate clinical endpoints predicting survival in oncology trials are highly context dependent, and contingent on disease, stage, patient population, and therapy. A fundamental problem of surrogate endpoint validation stems from the reliance on retrospective analysis of completed clinical trials. Extrapolation of surrogate endpoints to alternate disease settings or novel therapies with different mechanisms of action, mechanisms of acquired resistance, and toxicity profiles is problematic and can limit incorporation into subsequent clinical trials. Nonetheless, the utilization of clinical endpoints including ORR, response duration, and PFS remains an attractive solution to the

challenges of prolonged, large randomized controlled trials and post-progression bias associated with OS. The findings reported by Blumenthal *et al*. represent an important step in validating a relationship between ORR and PFS in trials treating patients with metastatic cancer using both cytotoxic and targeted agents. Additionally, by controlling for potential crossover in our analysis, a relationship is also suggested between ORR and OS. In the current era of clinical trials for uncommon biomarker defined patient populations, surrogate endpoints of clinical efficacy are increasingly needed. Likewise, the rapid rise of checkpoint immunotherapies poses unique challenges in measuring clinical benefit for lung cancer patients. The continued validation of surrogate endpoints for use in trials of NSCLC patients must remain a priority.

Acknowledgements

Footnote

Comment on: Blumenthal GM, Karuri SW, Zhang H, *et al*. Overall response rate, progression-free survival, and overall survival with targeted and standard therapies in advanced non-small-cell lung cancer: US Food and Drug Administration trial-level and patient-level analyses. J Clin Oncol 2015;33:1008-14.

References

1. Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. Lancet Oncol 2015;16:e32-42.
2. Blumenthal GM, Karuri SW, Zhang H, et al. Overall response rate, progression-free survival, and overall survival with targeted and standard therapies in advanced non-small-cell lung cancer: US Food and Drug Administration trial-level and patient-level analyses. J Clin Oncol 2015;33:1008-14.
3. Hotta K, Suzuki E, Di Maio M, et al. Progression-free

survival and overall survival in phase III trials of molecular-targeted agents in advanced non-small-cell lung cancer. Lung Cancer 2013;79:20-6.

4. Kris MG, Johnson BE, Berry LD, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA 2014;311:1998-2006.

5. Kim DW, Mehra R, Tan DS, et al. Ceritinib in advanced anaplastic lymphoma kinase (ALK)-rearranged (ALK+) non-small cell lung cancer (NSCLC): Results of the ASCEND-1 trial. ASCO Meeting Abstracts 2014;32:abstr 8003^.

6. Crinò L, Kim D, Riely GJ, et al. Initial phase II results with crizotinib in advanced ALK-positive non-small cell lung cancer (NSCLC): PROFILE 1005. ASCO Meeting Abstracts 2011;29:abstr 7514.

7. Camidge DR, Pao W, Sequist LV. Acquired resistance to TKIs in solid tumours: learning from lung cancer. Nat Rev Clin Oncol 2014;11:473-81.

8. Sequist LV, Soria JC, Gadgeel SM, et al. First-in-human evaluation of CO-1686, an irreversible, highly selective tyrosine kinase inhibitor of mutations of EGFR (activating and T790M). ASCO Meeting Abstracts 2014;32:abstr 8010^.

9. Janne PA, Ramalingam SS, Yang JC, et al. Clinical activity of the mutant-selective EGFR inhibitor AZD9291 in patients (pts) with EGFR inhibitor-resistant non-small cell lung cancer (NSCLC). ASCO Meeting Abstracts 2014;32:abstr 8009^.

10. Kim ES, Herbst RS, Wistuba II, et al. The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov 2011;1:44-53.

11. Blumenschein GR Jr, Saintigny P, Liu S, et al. Comprehensive biomarker analysis and final efficacy results of sorafenib in the BATTLE trial. Clin Cancer Res 2013;19:6967-75.

12. Papadimitrakopoulou V, Wistuba II, Lee JJ, et al. BATTLE-2 program: A biomarker-integrated targeted therapy study in previously treated patients with advanced non-small cell lung cancer (NSCLC). ASCO Meeting Abstracts 2013;31:abstr TPS8118.

13. Herbst RS, Gandara DR, Hirsch FR, et al. Lung Master Protocol (Lung-MAP)-A Biomarker-Driven Protocol for Accelerating Development of Therapies for Squamous Cell Lung Cancer: SWOG S1400. Clin Cancer Res 2015;21:1514-24.

14. Rizvi NA, Mazières J, Planchard D, et al. Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. Lancet Oncol 2015;16:257-65.

15. Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med 2015;372:2018-28.

16. Hoos A, Eggermont AM, Janetzki S, et al. Improved endpoints for cancer immunotherapy trials. J Natl Cancer Inst 2010;102:1388-97.

17. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009;15:7412-20.

18. Mauguen A, Pignon JP, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. Lancet Oncol 2013;14:619-26.

19. Laporte S, Squifflet P, Baroux N, et al. Prediction of survival benefits from progression-free survival benefits in advanced non-small-cell lung cancer: evidence from a meta-analysis of 2334 patients from 5 randomised trials. BMJ Open 2013;3.

20. Hellmann MD, Chaft JE, William WN Jr, et al. Pathological response after neoadjuvant chemotherapy in resectable non-small-cell lung cancers: proposal for the use of major pathological response as a surrogate endpoint. Lancet Oncol 2014;15:e42-50.