

Interpretation of lung cancer study outcomes

Diego Cortinovis, Marida Abbate, Paolo Bidoli, Davide Pelizzoni, Stefania Canova

Medical Oncology Unit, S Gerardo Hospital, Monza, Italy

Correspondence to: Diego Cortinovis. Medical Oncology Unit, S Gerardo Hospital, Monza, Italy. Email: d.cortinovis@hsgerardo.org.

Abstract: Lung cancer is the leading cause of cancer death in developed countries. However, in the last few years we observed an important acceleration in drug development due to oncogenic driver tumors discovery. Sharing and putting together preclinical data from benchmark and data from clinical research is the scientific paradigm that allows real breakthrough in clinical practice in this field, but only a few targeted agents are worthy and practice changing. The clinical research and proper use of statistical methodology are the pillars to continue to achieve important goals like improvement of overall survival. A good medical oncologist should be able to critically read a scientific paper and move from the observed outcomes into clinical perspective. Despite clinical improvements, sometimes the union of promising targeted agents and optimistic expectations misrepresent the reality and the value of clinical research. In this article, we try to analyze the meaning of statistical assumptions from clinical trials, especially in lung cancer, through a critical review of the concept of value-based medicine. We also attempt to give the reader some practical tools to weigh scientific value of literature reports.

Keywords: End-points; clinical trials; lung cancer

Submitted Oct 09, 2015. Accepted for publication Nov 03, 2015.

doi: 10.3978/j.issn.2072-1439.2015.11.26

View this article at: <http://dx.doi.org/10.3978/j.issn.2072-1439.2015.11.26>

Introduction

The major advances in the management of lung cancer have been achieved in the last few years. At the beginning of this millennium, a pessimistic attitude mirrored the awareness that chemotherapy reached a plateau and that the probability of new drugs with classic but very effective cytotoxic effects was very low. The discovery that some chemotherapeutic agents act differently on different histologies and that some lung cancers progress on an oncogene-addicted basis led the conception of this neoplasm to change toward a more curable disease.

The introduction in clinical practice of targeted agents like epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) as gefitinib, erlotinib and afatinib, and anaplastic-lymphoma kinase (ALK) TKIs like crizotinib dramatically changed the clinical outcome of these oncogene-addicted lung cancer patients with median overall survival (mOS) and median progression free survival (mPFS) very different compared with wild type cancers. The immunomodulating agents like antibodies against protein

death/ligand 1 (Ab antiPD1/PDL1) are the new hope and technically an important tool to harness the tumor growth triggering the immune system against the tumor immune escape mechanisms.

In general, all these drugs are more specific than antitlastic agents and usually combine the improvement of clinical outcomes with better quality of life (QoL), causing adverse events often well managed with supportive care.

These examples are extremely positive results arising from well-designed studies; however, the history of lung cancer therapy is studded with a huge number of unsuccessful drugs. In fact, many phase III trials failed to evidence significant benefits with waste of money and thousands of patients submitted to futile treatments. Many factors could explain these failures, for example bad rationale assumptions, weak statistical plans, commercial or mass media pressure, misguided interpretation of clinical studies by enthusiastic investigators and too premature reported trial results.

In this scenario, it is important to know some statistical

aspects about clinical trials in order to properly interpret the results and to be able to recognize practice-changing treatments.

How to interpret lung cancer study outcomes? When does statistical value relate to clinical significance? Why are some drugs not employed in clinical practice even if the primary study endpoint is met? Is the current clinical trial methodology suitable to study newer targeted agents?

In this article, we will review the statistical aspects that a reader should strive to recognize reading a paper, with a particular emphasis in interpretation of clinical outcomes, principal endpoints and value medicine in lung cancer research.

True and surrogate endpoints in lung cancer clinical trials

Primary and secondary endpoints should be clearly stated in the study design. Overall survival should ideally remain the primary endpoint in lung cancer phase III trials because it is unambiguous, easily measured and clinically significant. Nevertheless, it requires large sample sizes and prolonged follow-up, thus increasing costs and delaying new drugs approval. Furthermore, it is affected by subsequent therapies and crossover, and takes into account cancer-unrelated deaths.

Progression free survival has been proposed and studied as a surrogate endpoint to overcome these drawbacks. A surrogate endpoint is a measure that can replace a “true” endpoint as a predictor of clinical benefit. PFS requires more frequent disease status assessments than OS, it is affected by biases and measurement errors, but it is not affected by subsequent treatments, it needs fewer patients and can be assessed earlier than OS, thus shortening the duration of cancer trials, finally resulting in less costs and faster approval of new agents.

However, a correlation between surrogate and real endpoint is not sufficient to establish surrogacy. The validation of a surrogate endpoint should proceed through different steps, from Prentice’s criteria for a single trial (1) toward a “meta-analytic” approach using individual patient data from randomized trials, not meta-analyses of published papers (2).

Moreover, a proven surrogate endpoint for a type of cancer or a class of drugs should not be automatically extended to other malignancies or different classes of agents.

This issue is particularly important in non-small cell

lung cancer (NSCLC) trials. In fact, while PFS could be considered an adequate endpoint in the first-line setting, OS must remain the primary endpoint beyond first line. Furthermore, while there is poor evidence to consider PFS as a proper surrogate endpoint for OS in patients with advanced NSCLC treated with cytotoxic agents (3), there is lacking evidence to apply surrogacy in patients treated with targeted agents. However, Tsujino *et al.* demonstrated a strong correlation between response rate (RR) and OS in patients treated with gefitinib or erlotinib (4). It is possible that the benefit of tyrosine kinase inhibitors on survival, included post-progression survival (PPS), mask the improvements on short time. On the other hand, a long PPS can obscure a statistical significant improvement in OS, but this does not imply lack of benefit (5). This could be especially true for immunological agents that demonstrated a delayed positive effect on survival with prolonged PPS.

Interpretation of clinical outcomes in oncogene-addicted era

We will critically review some examples from the literature to give a right interpretation of clinical outcomes in clinical trials employing targeted agents in selected populations.

Today Gefitinib is considered one of the most important weapons against NSCLC harboring EGFR sensitive mutations.

The pivotal study published by Mok and colleagues showed that gefitinib was superior to carboplatin and paclitaxel in advanced chemo-naïve lung adenocarcinoma, non-smoker or previous light smoker patients, decreasing the risk of progressive disease (6).

If we consider a quantitative measure like PFS, that was the primary study endpoint, no difference was observed in terms of months for gefitinib and chemotherapy (5.7 *vs.* 5.8 respectively). However, the shape of the curves (*Figure 1*) suggested a great benefit for TKI compared with chemotherapy in a subpopulation. In fact, the hazard ratio (HR) for PFS was positive for gefitinib with a significant reduction of the risk of progressive disease of 26% (HR 0.74, 95% CI: 0.65-0.85, $P<0.001$) with respect to chemotherapy. A pre-planned analysis of tumor biomarkers like EGFR mutations explained this result. Gefitinib performed in opposite manner on these two distinct populations with a great effect evidenced only in the oncogene-addicted group (*Figure 1*).

We can analyze many other characteristics of the curves to understand the significant activity of gefitinib in this

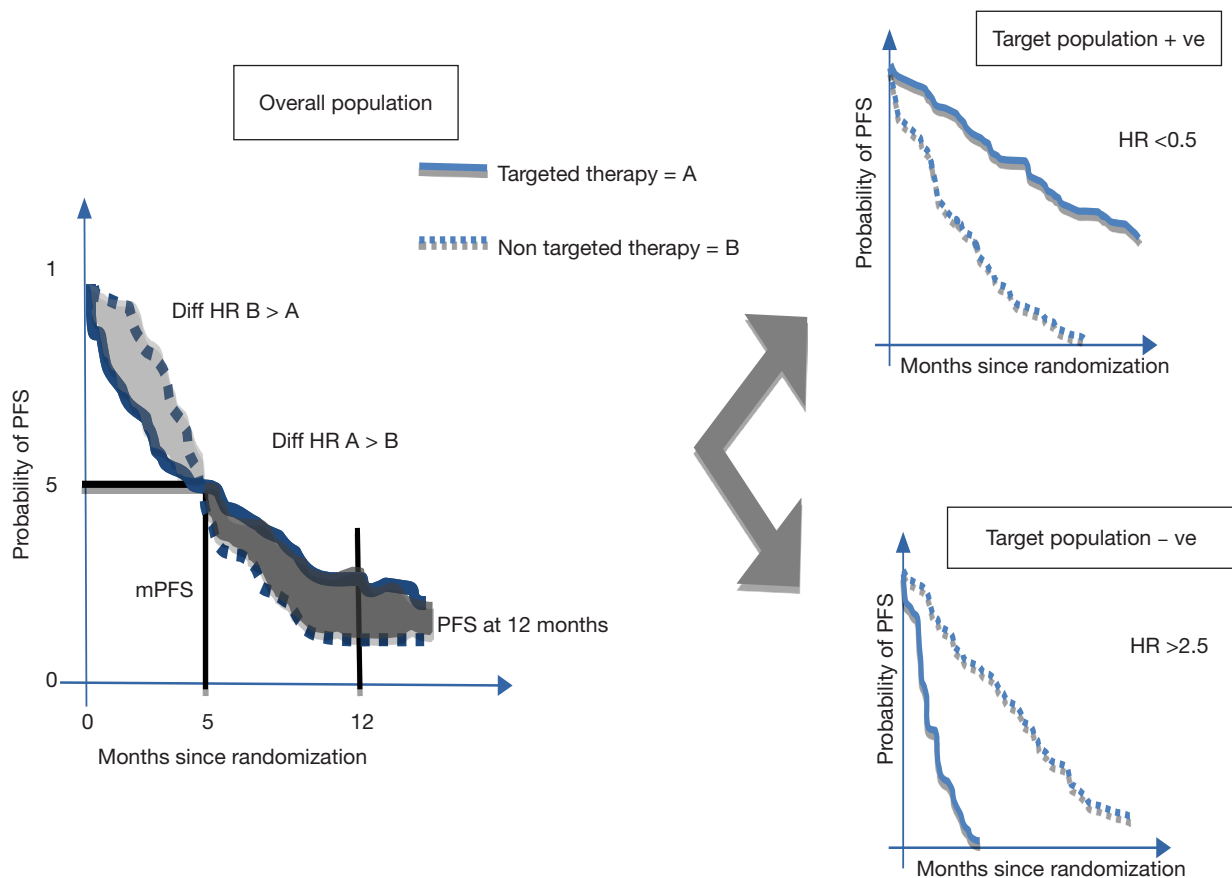


Figure 1 A peculiar shape of curves. Interpretation of targeted agents' effect on target sub population. mPFS, median progression free survival.

specific subpopulation, but the met primary endpoint well highlights the heart of the matter. In fact, the 12-month rates of PFS were 24.9% with gefitinib and 6.7% with carboplatin plus paclitaxel. Furthermore, the different probability to be free of disease progression throughout the course of the study indirectly demonstrated that there were two distinct populations. The probability was greater in the carboplatin plus paclitaxel arm in the first 6 months and greater in the gefitinib arm in the following 16 months. The tails of the curves were also important to weigh the impact of this EGFR TKI. Finally, there was a significant interaction between EGFR mutation status and treatment with respect to PFS ($P < 0.001$).

Another relevant point highlighting the difference between a conventional chemotherapy compared with a targeted agent is the different impact of the treatments on OS.

In this study, gefitinib was not different with respect to chemotherapy in terms of OS. Median OS was 18.6 months in gefitinib arm compared with 17.3 months among patients

receiving chemotherapy. These results are surprising because one NSCLC selected population had a median OS superior to 12 months for the first time. However, the lack of superiority of gefitinib over chemotherapy was apparent, because the cross over phenomenon could explain these results. In fact, more than 50% of the patients in the chemotherapy arm received gefitinib at progression with a benefit magnitude comparable also in the second line setting.

Finally, the full comprehension of this trial with a careful focus on statistical aspects allows the reader to consider the EGFR mutated NSCLC patients as a unique population and the EGFR status as a predictive and prognostic factor.

Another example is crizotinib compared with chemotherapy in ALK translocated patients (7).

In this pivotal, phase III trial crizotinib showed a significant superiority over platinum-derivatives plus pemetrexed in terms of PFS (10.9 vs. 7.0 months, HR 0.45, 95% CI: 0.35-0.6) with a more favorable toxicity profile

than chemotherapy. These impressive results were obtained in a selected population, namely ALK translocated NSCLC patients, but once again, OS was comparable probably due to allowed cross over as well explained in the Supplementary Appendix, and the low rate of death at the time of analysis.

Can we then conclude that PFS is the best endpoint to evaluate the activity of a targeted agent in the era of biomolecular markers?

Probably not, or not yet until a proven surrogacy would be demonstrated in this setting and for this type of drugs, as explained above.

Considering a different class of agents, we move toward a distinct scenario where benefit is evidenced both in terms of PFS and in terms of OS.

Antiangiogenic agents recently demonstrated activity in pretreated lung cancer patients. Ramucirumab, an antiVEGFR2 antibody, plus docetaxel showed superiority versus docetaxel plus placebo in terms of OS, that was the primary endpoint, in pretreated advanced NSCLC (8). The benefit of 1.4 months was proven in the overall population with a maximum effect in patients who progressed to first line therapy within 9 months (HR 0.65, 95% CI: 0.56-0.75). The curves of OS early diverged and this clinical benefit was maintained even at a longer follow-up.

We often observe this curve trend in studies evaluating antiangiogenic agents (9), probably due to the major dependence of tumor to angiogenesis in the early phase of disease. Unfortunately, this category of drugs lacks predictive biomarkers and the little difference respect to placebo is probably due to some unknown biomolecular aspects. These drugs seem active in unselected populations, but with an unimpressive clinical benefit.

Hence, it is as challenging as important for a young clinician to discriminate between prognostic and predictive factors (10). This is especially worthwhile in the era of targeted agents in which the hope remains to employ the customized therapy for the right patient.

Statistical versus clinical value

Nowadays, economical aspects are as important as clinical considerations to choose the right strategy for the right patient. Furthermore, regulatory agencies are more aware of economic issues with the aim to avoid approval of drugs demonstrating a statistical significant benefit of OS without a clinical relevant improvement, and to fasten approval of drugs showing differences in time to progression (TTP) or PFS “of a substantial magnitude” (11) well balanced with

toxicity profile.

The quality of clinical trials in lung cancer is also under debate.

In one original report, Sacher *et al.* demonstrated a shift in the paradigm of design and interpretation of phase III clinical trials during the last three decades (12).

Since the 80' to current days, the use of OS as primary endpoint declined in favor of PFS and the sample size of single studies increased, but the median magnitude of clinical benefit decreased. In particular an alarming trend seems to emerge as a higher number of clinical trials without significance in primary endpoint are reported as “clinical meaningful” based on significance in secondary endpoints or in pre-planned subgroup analysis.

Even if criticized (13), this article raises some considerations to take into account. The differences between therapeutic strategies often show a marginal clinical improvement, in particular in the latest years, and the clinical benefit of one agent should be considered even in the setting of post-progression therapeutic opportunities, toxicity profile and QoL.

The employment of second and further line treatments could reduce the impact of first line therapies on OS in lung cancer. In this setting an endpoint like PFS, even with all the pitfalls previously described, may be more advisable than the gold standard OS.

However, clinicians as well as patients argue about what and when consider clinically meaningful a clinical trial in lung cancer.

The most important cancer scientific societies, American Society of Clinical Oncology (ASCO) (14) and European Society of Medical Oncology (ESMO) (15), dealt with this topic and drew the bar in order to define what is noteworthy in clinical practice and which kind of drugs deserved a role in the treatment of real world population.

In particular, the perception of medicine value in NSCLC trials varies between histotypes. The American society suggest that one experimental agent in non-squamous NSCLC should be considered practice changing if it increases PFS of at least 4 months and OS of 3.5-4 months with a corresponding death risk reduction of 20-24%. Due to more severe prognosis, the desired benefit in squamous NSCLC should be of 3 months in terms of better PFS and 2.5-3 months for OS with a death risk reduction of 20-23%.

The working groups also dealt with issues about the estimation of benefit in terms of QoL, but the lack of worldwide-approved scales did not led to any firm conclusion. The challenge is that a positive clinical trial, as

Table 1 ASCO and ESMO suggestion for clinical benefit

Patient population (first line)	Current baseline mOS (months)	Clinical meaningful mOS (months)	HR	Improvement 1 year survival rate (%)	Improvement mPFS (months)
ASCO (14)					
Non SCC NSCLC	13	3.25-4	0.76-0.8	53→61	4
SCC NSCLC	10	2.5-3	0.77-0.8	44→53	3
ESMO (15)					
Overall population	>1 year	5	0.7	>10% (at 3 years)	3
	<1 year	3	0.65	>10% (at 2 years)	1.5

mOS, median overall survival; mPFS, median progression free survival; SCC, squamous cell carcinoma; NSCLC, non-small cell lung cancer; HR, hazard ratio; ASCO, American Society of Clinical Oncology; ESMO, European Society of Medical Oncology.

previously defined, could also demonstrate a good toxicity profile with a resulting improvement in QoL.

The ESMO tried to raw a scale, the ESMO Magnitude of Clinical Benefit Scale (ESMO-MCBS), with the aim to measure the relevance of clinical benefit of any new anti-cancer therapy. This scale will be employed by ESMO to grade the strength of new treatments licensed by European Medicine Agency (EMA).

If a standard treatment demonstrates a median OS inferior than 12 months, the maximum grade of benefit for a palliative therapy should show an HR of 0.65 and a gain of 3 months with an OS increase at 2 years of at least 10%. Similarly, if the standard treatment is associated with a median OS longer than 12 months, the new gold standard should reduce the risk of death of 30% (HR 0.70) with a gain of 5 months, so enhancing the 3-year survival by 10% or more. In this scale, QoL improvement is also considered (*Table 1*).

After these considerations, most of the current treatments should not be employed in clinical practice. However, the current cost/benefit ratio is more challenging than in the past and the introduction of value medicine is necessary to safeguard patients from potential dangerous treatments that do not keep the promise of real clinical benefit.

The future in lung cancer clinical trials

In the era of biomolecular medicine, it is essential to draw powerful clinical trials with a solid scientific rationale that could guarantee the best treatment to the right patient with an acceptable timeline and fast approve of newer life-saving therapies.

Considering the high incidence of lung cancer, it is

necessary to change the design of traditional clinical trials hitherto used.

Primarily, investigators and companies should simplify inclusion and exclusion criteria to enhance the accrual, shorten the duration of enrollment and accelerate drug development and approval, in particular for targeted agents in oncogenic driver selected populations.

In this setting, the ASCO Cancer Research Committee recently published some recommendations stating proposals to limit the use of eligibility criteria adopted in clinical trials designed in the non-oncogenic era and to move toward more streamlined and more appropriate criteria in molecularly driven trials (16).

Secondly and mostly important, the usual procedures concerning the development of new agents from phase I to phase IV trials should be fully revised.

Several models are currently available and employed, and basket and umbrella trials are the most reliable tools in lung cancer research.

Basket trials enroll patients with different tumor types or histotypes but with a specific molecular alteration. Therefore, it happens that patients with different tumors can be randomized in the same cohort.

One example of basket trial is the CUSTOM (Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies) trial that tried to identify molecular biomarkers in different thoracic tumors (NSCLC, SCLC, thymic cancer) and evaluated five targeted agents grouped by molecular markers (*Figure 2*) (17).

On the other side, umbrella trials focus on a single tumor type or histology and include an infrastructure for screening and identification of patients. Multiple and parallel subtrials are simultaneously performed testing different targeted

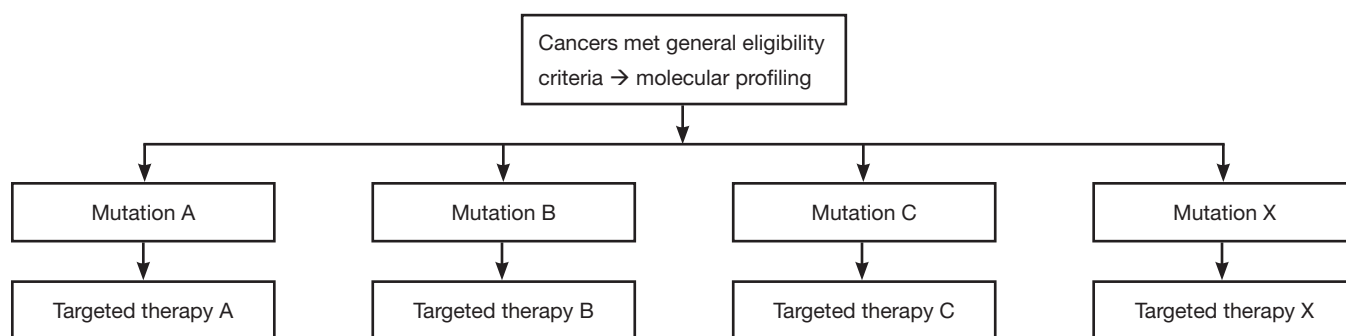


Figure 2 Basket trial schema.

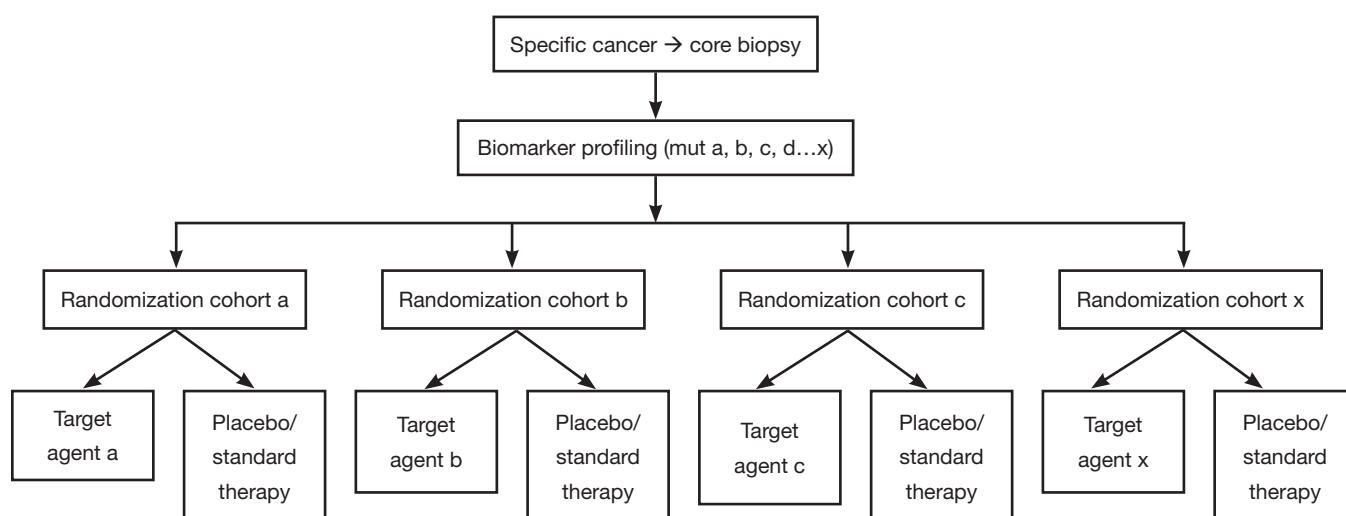


Figure 3 Umbrella trial schema.

agents in molecularly defined subsets.

An example of umbrella trial is the Biomarker-integrated Approached of Targeted Therapy for Lung cancer Elimination (BATTLE) trial in which advanced, refractory NSCLC patients were adaptively randomized to receive a targeted agent matched on relevant biomarkers analyzed in fresh core needle biopsy specimen (*Figure 3*) (18).

The goals of these recent types of study and statistical approaches are to facilitate patients screening and accrual, and to speed drug development.

Conclusions

Lung cancer is still considered a devastating disease with high incidence and mortality rate. Clinical trials are the backbone of the research and the main powerful tools to develop active and effective therapies.

The clinical outcomes in lung cancer trials are obviously the ever important PFS, OS, RR and QoL. However, the interpretation of the magnitude of benefit is not always easy for clinicians not accustomed to lung cancer research.

In particular, in the era of targeted agents the classical outcomes must be critically reviewed with the aim to understand the value of new agents before their employment in clinical practice considering different aspects linked to activity, acute and chronic toxicity, and costs.

In a world in which the economic resources are limited and public health systems risk collapse, it is mandatory to adopt only drugs that demonstrate a major and meaningful impact on population.

These goals could be achieved using new tools and statistical approaches to accurately measure the clinical benefit so as to manage to foresee treatments deserving employment in clinical practice.

The awareness of these aspects could lead the clinicians to properly use targeted agents and find out the essential of modern oncology, the precision medicine.

Acknowledgements

None.

Footnote

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

References

- Heller G. Statistical controversies in clinical research: an initial evaluation of a surrogate end point using a single randomized clinical trial and the Prentice criteria. *Ann Oncol* 2015;26:2012-6.
- Baker SG, Kramer BS. A perfect correlate does not a surrogate make. *BMC Med Res Methodol* 2003;3:16.
- 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.
- Tsujino K, Kawaguchi T, Kubo A, et al. Response rate is associated with prolonged survival in patients with advanced non-small cell lung cancer treated with gefitinib or erlotinib. *J Thorac Oncol* 2009;4:994-1001.
- Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst* 2009;101:1642-9.
- Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361:947-57.
- Solomon BJ, Mok T, Kim DW, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med* 2014;371:2167-77.
- Garon EB, Ciuleanu TE, Arrieta O, et al. Ramucirumab plus docetaxel versus placebo plus docetaxel for second-line treatment of stage IV non-small-cell lung cancer after disease progression on platinum-based therapy (REVEL): a multicentre, double-blind, randomised phase 3 trial. *Lancet* 2014;384:665-73.
- Reck M, Kaiser R, Mellemegaard A, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol* 2014;15:143-55.
- Simms L, Barraclough H, Govindan R. Biostatistics primer: what a clinician ought to know--prognostic and predictive factors. *Biostatistics primer: what a clinician ought to know--prognostic and predictive factors. J Thorac Oncol* 2013;8:808-13.
- United States, Department of Health and Human Services, Food and Drug Administration (FDA). Guidance for Industry: Clinical trial endpoints for the approval of non-small cell lung cancer drugs and biologics. Available online: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM259421.pdf>, 2015.
- Sacher AG, Le LW, Leigh NB. Shifting patterns in the interpretation of phase III clinical trial outcomes in advanced non-small-cell lung cancer: the bar is dropping. *J Clin Oncol* 2014;32:1407-11.
- Loong HH, Mok TS. Lung cancer: dropping bars or rising hoops--phase III outcomes of NSCLC. *Nat Rev Clin Oncol* 2014;11:381-2.
- Ellis LM, Bernstein DS, Voest EE, et al. American Society of Clinical Oncology perspective: Raising the bar for clinical trials by defining clinically meaningful outcomes. *J Clin Oncol* 2014;32:1277-80.
- Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol* 2015;26:1547-73.
- Kim ES, Bernstein D, Hilsenbeck SG, et al. Modernizing Eligibility Criteria for Molecularly Driven Trials. *J Clin Oncol* 2015;33:2815-20.
- Lopez-Chavez A, Thomas A, Rajan A, et al. Molecular profiling and targeted therapy for advanced thoracic malignancies: a biomarker-derived, multiarm, multihistology phase II basket trial. *J Clin Oncol* 2015;33:1000-7.
- Kim ES, Herbst RS, Wistuba II, et al. The BATTLER trial: personalizing therapy for lung cancer. *Cancer Discov* 2011;1:44-53.

Cite this article as: Cortinovis D, Abbate M, Bidoli P, Pelizzoni D, Canova S. Interpretation of lung cancer study outcomes. *J Thorac Dis* 2015;7(11):E541-E547. doi: 10.3978/j.issn.2072-1439.2015.11.26