JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

Shifting Patterns in the Interpretation of Phase III Clinical Trial Outcomes in Advanced Non–Small-Cell Lung Cancer: The Bar Is Dropping

Adrian G. Sacher, Lisa W. Le, and Natasha B. Leighl

Processed as a Rapid Communication manuscript

All authors: Princess Margaret Cancer Centre/University Health Network, University of Toronto, Toronto, Ontario, Canada.

Published online ahead of print at www.jco.org on March 3, 2014.

Supported by the Division of Medical Oncology/Hematology, Princess Margaret Cancer Centre/University Health Network.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Natasha B. Leighl, MD, MMSc, Department of Biostatistics, Division of Medical Oncology/Hematology, Princess Margaret Cancer Centre/University Health Network, University of Toronto, 610 University Ave, 5-105, Toronto, Ontario, Canada M5G 2M9; e-mail: Natasha.Leighl@uhn.ca.

0732-183X/14/3299-1/$20.00

DOI: 10.1200/JCO.2013.52.7804

ABSTRACT

Purpose
Despite multiple trials of new agents in advanced non–small-cell lung cancer (NSCLC), outcomes remain poor. This study explores how the design and interpretation of randomized trials in advanced NSCLC has changed over time.

Methods
Phase III randomized controlled trials of systemic therapy for advanced NSCLC between 1980 and 2010 were identified, and their primary end point, outcome, statistical significance, and conclusions were recorded.

Results
Of 245 trials identified, 203 were eligible for study inclusion. Although overall survival remains the most common primary end point of phase III trials, more trials from the last decade have used progression-free survival instead (none in 1980 to 1990, 13% in 2001 to 2010; $P = .002$). The percentage of trials meeting their primary statistical end points remained stable over time; however, the percentage of trials reporting a positive outcome without meeting that end point increased (30% in 1980 to 1990, 53% in 2001 to 2010; $P < .001$). A trend toward decreasing magnitude of survival gain in positive trials was seen over time (3.9 months in 1980 to 1990, 2.5 months in 2001 to 2010; $P = .11$), with a concomitant increase in the sample size of clinical trials over the same time period (median: 152 patients in 1980 to 1990, 413 in 2001 to 2010; $P < .001$). Only studies predating 1990 reported negative results as a result of insufficient magnitude of survival benefit despite statistical significance.

Conclusion
A significant shift has occurred over the past three decades in the design and interpretation of phase III trials in advanced NSCLC. The use of survival as the primary measure of benefit is declining, as is the magnitude of benefit deemed clinically relevant.

J Clin Oncol 32. © 2014 by American Society of Clinical Oncology

INTRODUCTION

Advanced non–small-cell lung cancer (NSCLC) is among the most intractable solid tumors. Its aggressiveness paired with a high prevalence have rendered it the leading cause of cancer-related death.[1,2] As such, it has been the subject of extensive investigation over several decades in order to develop effective treatments. These efforts have yielded first- and second-line chemotherapy[3-11] and targeted therapy for selected subpopulations[12,13] or as third-line after chemotherapy failure.[14] However, the clinical benefit of these treatments has been modest.

Phase III clinical trials have been the primary vehicle by which these treatments have been shown to have a clinically meaningful and statistically significant impact on patient outcomes. As interest in NSCLC from the oncology community and pharmaceutical industry has grown, changes in both the design and interpretation of these trials have begun to emerge with respect to statistical power, sample size, and the primary end points used. There has been ongoing debate regarding oncology clinical trial design, including the use of progression-free survival (PFS) in lieu of overall survival as a primary study end point.[15-18] The use of PFS as a primary study end point has been promoted on the basis of an increasing use of crossover in clinical trials as well as an increasing number of lines of subsequent treatment. The US Food and Drug Administration has recently accepted PFS as a potential primary study end point for drug approval in advanced NSCLC but

has noted reservations regarding the accuracy, reproducibility, and clinical relevance of this end point.[19] Critics of PFS as a primary study end point have questioned the potential subjectivity of this end point with respect to the frequency of radiographic surveillance as well as its relationship to clinically relevant outcomes such as quality of life and overall survival.[17,19] The definition of clinically meaningful benefit in trials of new agents in advanced cancer continues to be debated, particularly with respect to emerging end points such as PFS.[20-24] The increasing number of systemic therapies available paired with escalating treatment costs have further fueled this discussion.

Apart from changes in the design of clinical trials, the manner in which these trials are interpreted may be changing as well. An increasing willingness to interpret the results of clinical trials as positive despite extremely modest or questionable benefit of a new treatment has been previously reported. This trend appears to be influenced by numerous factors, including potential trial funding sources and statistical knowledge.[25,26]

Thus, we have sought to investigate these trends in clinical trial design and interpretation in advanced NSCLC. Here we report a review of all phase III trials conducted in advanced NSCLC over the past three decades, examining changes over time in the primary end point, sample size, interpretation, and magnitude of benefit described in advanced NSCLC clinical trials.

METHODS

An extensive literature search for all phase III randomized controlled clinical trials involving systemic therapy for advanced NSCLC was performed by using the PubMed clinical queries application (Data Supplement). Professional oncology society guidelines were screened to identify studies missed in the aforementioned search strategy (Data Supplement). Studies were eligible for inclusion if they were phase III, involved systemic therapy only, studied patients with advanced NSCLC, and were conducted between 1980 and 2010. For eligible studies (Data Supplement), the following were abstracted: publication date, size, primary and secondary outcomes, trial outcome (positive or negative), statistical significance, and the magnitude of clinical benefit in terms of the primary outcome measure (ie, absolute difference in benefit between treatment arms). Trials were deemed positive if the authors recommended the new treatment arm be adopted on any basis or studied further. This was determined by the investigators on the basis of direct scrutiny of the study text for an explicit statement of either a recommendation for adopting the study treatment into practice or recommendation of further study of the experimental treatment. Trials were scored as positive if either of these explicit statements was found. Positive trial interpretation was recorded as a distinct entity from trial outcome with respect to statistical significance. Trials were further classified as asserting noninferiority if the authors explicitly stated that the study agent was either equal in efficacy or noninferior to the control arm. Trials were categorized by publication year, including those published between 1980 and 1990, 1991 and 2000, and 2001 and 2010.

Descriptive statistics were used to summarize the patterns in trial design, primary outcome, reported statistical significance, and magnitude of net survival benefit by each time period. The Cochran-Armitage trend test was used to test the change over the time periods in reporting the primary outcomes, trial conclusions, and statistical significances. The F test was used to test the change over time in reported net survival benefit and trial size. The net survival benefit is defined as the difference in median survival time between the experimental arm and the control arm. A log transformation was used to achieve approximate normal before the F test. Analysis was carried out by using SAS v9.3 (SAS Institute, Cary, NC).

RESULTS

In all, 248 trial reports were identified, and their full text was reviewed by the authors. Forty-five were excluded because they involved small-cell lung cancer (n = 3), earlier-stage NSCLC (n = 5), surgery or radiotherapy (n = 6), were not phase III trials (n = 13), or were subgroup/interim analyses of studies already included (n = 18). A total of 203 unique trials were included in our study.

Patterns in Trial Design: Sample Size and Primary End Point

Detailed review of the 203 trials revealed several trends over time. The number of phase III trials conducted in advanced NSCLC has increased dramatically—32 trials were conducted in the 1980s compared with 53 trials in 1990s and 118 trials in the most recent decade (Table 1). At the same time, the sample size of advanced NSCLC trials has also increased, from a median of 152 patients (range, 38 to 743 patients) in the 1980s to 184 patients (range, 58 to 680 patients) in the 1990s and 413 patients (range, 58 to 1,725 patients) in 2001 to 2010 ($P < .001$). The distribution of treatments being studied has also shifted from a heavy emphasis on multiagent triplet chemotherapy in the 1980s to predominantly doublet therapy in the 1990s and a sizeable increase in studies of targeted therapy in 2001 to 2010 (Table 1).

There has been a significant shift in the primary end point of phase III advanced NSCLC trials. Before 2000, almost all studies were designed with overall survival as the primary end point, most commonly reported as the median overall survival time (97% in 1980 to 1990; 96% in 1991 to 2000). Significantly fewer trials used overall survival as the primary end point in the period between 2001 and 2010 (81%; $P < .002$), with a clear increase in the use of PFS as the primary study end point instead (13%).

Interpretation of Trials

Interpretation of trial results in advanced NSCLC is also shifting over time. This is exemplified by the significant increase in the number of trials reporting a positive result, as defined by recommendation for treatment adoption or for further study (Table 1). In the period from 1980 to 1990, only 31% of studies were reported as positive compared with 53% in 2001 to 2010. Although the percentage of trials with statistically significant improvement in the primary outcome has not changed over time (29% in 1980 to 1990 and 31% in 2001 to 2010), the percentage of trials reporting a positive outcome despite not achieving statistical significance in their primary outcome has increased (30% [three of 10] of positive trials in 1981 to 1990 *v* 24% [nine of 37] in 1991 to 2000 *v* 53% [47 of 88] in 2001 to 2010; $P < .001$). These trials were recommended on the basis of improvements seen in secondary trial end points (PFS, toxicity; n = 24), asserting noninferiority despite lack of a statistically appropriate noninferiority design (n = 26), or recommending further study on the basis of a nonsignificant trend in primary outcome (n = 9). Interestingly, the period from 1980 to 1990 was the only one in which studies that met their statistical primary end point were declared negative because of the insufficient magnitude of clinical benefit (n = 2).

Clinical Benefit in Positive Trials

The final dimension examined was the magnitude of clinical benefit using the median survival time as the primary outcome. In the

Table 1. Characteristics of Studies Reviewed

Characteristic	Publication Date: 1980-1990 No.	1980-1990 %	1991-2000 No.	1991-2000 %	2001-2010 No.	2001-2010 %
No. of phase III advanced NSCLC trials identified	32		53		118	
Trial sample size						
Median	152		184		413	
Range	38-743		58-680		58-1,725	
Mean overall survival across trials, months	6.7		7.9		9.5	
Primary end point						
OS	31	97	51	96	96	81
PFS	—	—	—	—	15	13
Other (RR, QOL)	1	3	2	4	7	6
Agent investigated						
Chemotherapy	31	97	43	81	82	69
Singlet	0	0	6	11	16	14
Doublet	12	38	20	38	51	43
Triplet or more	19	59	17	32	15	13
Targeted	0	0	3	3	25	21
Other*	1	3	7	7	11	9
Trials reported as positive	10	31	37	70	88	75
Trials reported as positive for statistically significant improvement in primary survival end point	9	28*	28	53*	38	32*
Trials reported as positive, but did not demonstrate survival benefit	3	9*	9	17*	47†	40*
Trials reported as positive on the basis of:						
Nonsignificant trend in OS	1		—		8	
Secondary end point (eg, PFS, toxicity)	2		7		15	
Assertion of noninferiority despite lack of a strict noninferiority design	—		2		24	

Abbreviations: NSCLC, non–small-cell lung cancer; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate.
*Includes novel nontargeted chemical compounds, hormonal compounds, and immunotherapy.
†Three trials in 2001-2010 were reported as positive with appropriate noninferiority design.

60 trials reporting statistically significant survival improvement, the median net survival benefit was 3.9 months (range, 1.7 to 7.2 months) in 1981 to 1990 compared with 2.4 months in 1991 to 2000 (range, 0.3 to 7.0 months) and 2.5 months from 2001 to 2010 (range, 1.1 to 4.0 months; $P = .11$). When all trials deemed positive were considered, this changed from 3.9 months in 1980 to 1990 (range, 0.25 to 7.2 months), 2.0 months in 1991 to 2000 (range, 1.0 to 7.0 months), and 0.9 months from 2001 to 2010 (range, 2.5 to 4.0 months; $P < .001$; Fig 1). Interestingly, despite the decreasing magnitude of clinical benefit in these trials over time, the average median survival across trials in each time period has increased from 6.7 months in 1980 to 1990 to 9.5 months in 2001 to 2010 (Table 1).

DISCUSSION

This study reveals important changes that have occurred in phase III clinical trials in advanced NSCLC over the past 30 years. The size of trials has increased significantly alongside a decline in the use of overall survival in favor of PFS as the primary study end point. The average overall survival of patients in these trials has modestly but progressively increased over time. The number of trials reporting a positive outcome has also increased significantly. This is largely because of a growing number of studies recommending new agents on the basis of statistically nonsignificant trends in the primary end point, such as noninferiority without appropriate statistical testing, or benefit in secondary end points like toxicity. A similar trend toward increasingly positive interpretations of clinical trial outcomes has been identified in other areas of oncology-related clinical research and in research not related to oncology. Previously, this has been attributed to lack of statistical rigor and knowledge, increasing pressure for clinical trials to produce some positive result, and the influence of industry-related funding.[25,26]

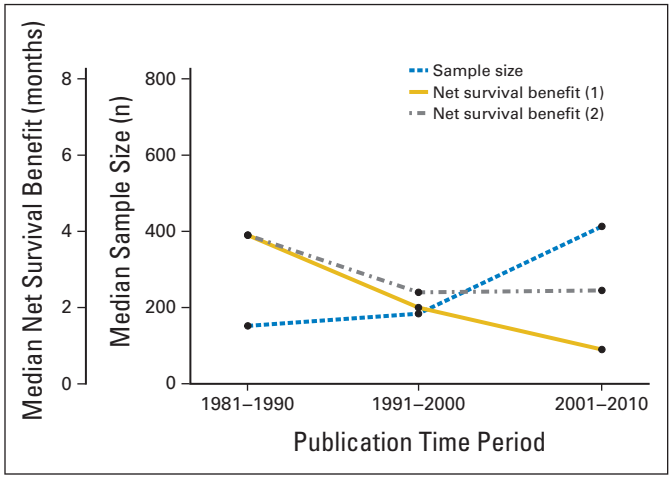

Fig 1. Median sample size over time and net survival benefit reported in (1) all positive trials and (2) those that met their primary statistical end point.

The increasing use of PFS as a surrogate for overall survival in advanced NSCLC trials also has the potential to further obfuscate the clinical benefit of newer agents in terms of those outcomes most important to people with advanced cancer: overall survival, quality of life, and symptom control.[15-18] The increasing popularity of PFS as a primary study end point can be attributed to both an increasing willingness on the part of regulatory authorities to accept it and an increasing use of cross-over trial designs and subsequent lines of treatment that may complicate the demonstration of an overall survival benefit.[19] However, the importance of demonstrating that PFS translates into meaningful clinical benefit for patients cannot be understated, given concerns regarding the potential subjectivity and variability of this end point, particularly when the magnitude of PFS benefit is small.[17,19]

Our findings confirm a disquieting shift in phase III trial design in advanced NSCLC. Despite the advances in design through larger size and statistical power over the past three decades, the bar for judging clinical benefit of new agents in lung cancer appears to be falling. Although larger trials are more adept at detecting small differences in survival benefit between treatment arms, the oncology community relies on investigators to define the minimum magnitude of clinical benefit, or minimum clinically important difference, to ascertain the value of a new agent in advanced cancer. It appears that investigators are progressively less stringent in deeming treatments of value even when the statistically significant advantage achieved is no longer clinically relevant. This distinction is particularly important in light of the increasing cost of conducting larger trials and approving new agents with modest benefit, as well as the overall poor prognosis of patients with advanced NSCLC.[21-24]

The number of trials that recommend adoption or further study of a new agent despite the initial trial not meeting its primary statistical end point has clearly increased over time. Although this may be in part related to modest benefits in secondary end points, including toxicity or presumed noninferiority, there may be additional biases at play. With advances in personalized medicine, there is a belief that an agent with a clinically insignificant benefit is a reflection of an insufficiently selected population rather than insufficient drug activity. Publication bias or trial sponsorship may also play a role in investigator conclusions regarding trials. Positive trials remain more likely to be published, no matter how important the negative result may be to the oncology community.

These findings raise questions regarding the design and interpretation of phase III trials in advanced NSCLC. The intention of multiphase clinical trials is to focus resources on promising agents, with each phase eliminating either toxic or minimally active agents. Our results suggest that we are progressively less efficient in accomplishing this task in advanced NSCLC trials. Clear standards of acceptable minimal clinical differences in advanced NSCLC and appropriate clinically relevant outcomes should be established and endorsed. Previous recommendations on the minimal relevant clinical benefit have varied greatly.[24,27] We believe the proposals to raise the bar of minimal clinical benefit to a 50% improvement in the current overall survival and doubling of PFS to be particularly promising.[24] A renewed focus on earlier-phase trials designed to detect clinically important differences in treatment before proceeding to large phase III trials and greater study of the validity of surrogate end points such as PFS may help.[21,24] Encouraging smaller earlier-phase trials designed to detect clinically significant benefit from new agents will speed development of new, more effective drugs. This will also encourage a greater emphasis on early biomarker development and patient selection where appropriate in order to optimize benefit in these trials.[24] If we can alter these trends in trial design and interpretation, there is great potential to minimize trial and drug development costs on marginally effective, or ineffective, agents.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Adrian G. Sacher, Natasha B. Leighl
Financial support: Natasha B. Leighl
Administrative support: Natasha B. Leighl
Collection and assembly of data: Adrian G. Sacher, Natasha B. Leighl
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Howlader N, Noone AM, Krapcho M, et al (eds): SEER Cancer Statistics Review, 1975-2009. Bethesda, MD, National Cancer Institute, 2012
2. Siegel R, Naishadham D, Jemal A: Cancer statistics, 2012. CA Cancer J Clin 62:10-29, 2012
3. Le Chevalier T, Brisgand D, Douillard JY, et al: Randomized study of vinorelbine and cisplatin versus vindesine and cisplatin versus vinorelbine alone in advanced non-small-cell lung cancer: Results of a European multicenter trial including 612 patients. J Clin Oncol 12:360-367, 1994
4. Shepherd FA, Dancey J, Ramlau R, et al: Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. J Clin Oncol 18:2095-2103, 2000
5. Belani CP, Pereira JR, von Pawel J, et al: Effect of chemotherapy for advanced non-small cell lung cancer on patients' quality of life: A randomized controlled trial. Lung Cancer 53:231-239, 2006
6. Hanna N, Shepherd FA, Fossella FV, et al: Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J Clin Oncol 22:1589-1597, 2004
7. Scagliotti GV, Parikh P, von Pawel J, et al: Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 26:3543-3551, 2008
8. Rosell R, Gatzemeier U, Betticher DC, et al: Phase III randomised trial comparing paclitaxel/carboplatin with paclitaxel/cisplatin in patients with advanced non-small-cell lung cancer: A cooperative multinational trial. Ann Oncol 13:1539-1549, 2002
9. Schiller JH, Harrington D, Belani CP, et al: Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. N Engl J Med 346:92-98, 2002
10. Ardizzoni A, Boni L, Tiseo M, et al: Cisplatin-versus carboplatin-based chemotherapy in first-line treatment of advanced non-small-cell lung cancer: An individual patient data meta-analysis. J Natl Cancer Inst 99:847-857, 2007
11. Le Chevalier T, Scagliotti G, Natale R, et al: Efficacy of gemcitabine plus platinum chemotherapy compared with other platinum containing regimens in advanced non-small-cell lung cancer: A meta-analysis of survival outcomes. Lung Cancer 47:69-80, 2005
12. Mok TS, Wu YL, Thongprasert S, et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 361:947-957, 2009
13. Kwak EL, Bang YJ, Camidge DR, et al: Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 363:1693-1703, 2010
14. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al: Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med 353:123-132, 2005
15. Garon EB: Issues surrounding clinical trial endpoints in solid malignancies with a focus on

metastatic non-small cell lung cancer. Lung Cancer 77:475-481, 2012

16. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al: Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol 26:1987-1992, 2008

17. Soria JC, Massard C, Le Chevalier T, et al: Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? Ann Oncol 21:2324-2332, 2010

18. Sargent DJ, Hayes DF: Assessing the measure of a new drug: Is survival the only thing that matters? J Clin Oncol 26:1922-1923, 2008

19. U.S. Department of Health and Human Services, Food and Drug Administration: Guidance for Industry Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics. Rockville, MD, 2011

20. Carney DN: Lung cancer: Time to move on from chemotherapy. N Engl J Med 346:126-128, 2002

21. Ocana A, Tannock IF: When are "positive" clinical trials in oncology truly positive? J Natl Cancer Inst 103:16-20, 2011

22. Rothenberg ML, Carbone DP, Johnson DH: Improving the evaluation of new cancer treatments: Challenges and opportunities. Nat Rev Cancer 3:303-309, 2003

23. Roberts TG Jr, Lynch TJ Jr, Chabner BA: The phase III trial in the era of targeted therapy: Unraveling the "go or no go" decision. J Clin Oncol 21:3683-3695, 2003

24. Sobrero A, Bruzzi P: Incremental advance or seismic shift? The need to raise the bar of efficacy for drug approval. J Clin Oncol 27:5868-5873, 2009

25. Bariani GM, de Celis Ferrari AC, Hoff PM, et al: Self-reported conflicts of interest of authors, trial sponsorship, and the interpretation of editorials and related phase III trials in oncology. J Clin Oncol 31:2289-2295, 2013

26. Turner EH, Matthews AM, Linardatos E, et al: Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 358: 252-260, 2008

27. Breathnach OS, Freidlin B, Conley B, et al: Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: Sobering results. J Clin Oncol 19:1734-1742, 2001

GLOSSARY TERMS

non–small-cell lung cancer (NSCLC): a type of lung cancer that includes squamous cell carcinoma, adenocarcinoma, and large-cell carcinoma.

overall survival: the duration between random assignment and death.

Acknowledgment

We thank the *Journal of Clinical Oncology* reviewers for their thoughtful suggestions and advice.

