

Statistical considerations and endpoints for clinical lung cancer studies: can progression free survival (PFS) substitute overall survival (OS) as a valid endpoint in clinical trials for advanced non-small-cell lung cancer?

Lothar R Pilz¹, Christian Manegold¹, Gerald Schmid-Bindert²

¹Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1, 68167 Mannheim, Germany; ²Interdisziplinäre thorakale Onkologie, Department of Surgery, University Medical Center Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1, 68167 Mannheim, Germany
Correspondence to: Lothar R Pilz. Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1, 68167 Mannheim, Germany. Email: Lothar.Pilz@medma.uni-heidelberg.de.

Abstract: In the last decades significant progress has been achieved in the biological understanding of non-small-cell lung cancer (NSCLC) and its tumor heterogeneity has become more evident. The identification of novel tumor targets with different pathways has stimulated the search for anti-tumor agents with a specific target directed mode of action, stipulating the need of testing these agents in clinical trials with an appropriate choice of the study endpoint. Gold standard as an endpoint has been so far overall survival (OS). By definition there are 3 categories of classical endpoints applied generally in clinical lung cancer studies: survival time endpoints, symptom endpoints, and endpoints relying on patients' reporting. Beside classical endpoints like OS which are tending to show the direct clinical effect of treatment, efforts have been taken to substitute these classical endpoints by surrogates. As a surrogate candidate for OS progression-free survival (PFS) should have the inherent considerable advantage, that it can detect subpopulations with longer PFS intervals early. Based on the (sub-) population treated and having in mind the risk-benefit profile of the drug under consideration, PFS can be considered for regulatory decision making. If accompanied by some independent measures like quality of life or treatment toxicity, PFS should be able to cover the clinical benefit achieved by treatment. Selecting PFS as primary endpoint in Phase III trials of advanced NSCLC may be based on a number of questions such as: Does the definition of PFS fit into the setting used by other trials? Are there accepted consensus standards? Are there consistent surveillance intervals? Is validation for each agent group planned? Is the incremental improvement of PFS big enough ($\geq 30\%$)? And are there some additional measures to confine clinical benefit? OS is still accepted as the gold standard in trials investigating advanced NSCLC. OS is easy to measure and precise but it may be difficult to interpret if treatment action takes place only in a small subinterval of overall survival. PFS with some additional measures has become attractive when it seems advisable to make study results available earlier. Candidates for supporting PFS as "additional measures" may be treatment toxicity and quality of life measures. PFS allows a more precise detection and attribution to effects of the investigational treatment without being compromised by subsequent treatments. Therefore "enriched PFS" can be considered as an alternative primary endpoint replacing OS in studies investigating advanced NSCLC. The endpoint selection process should always be performed carefully considering all true and surrogate endpoint options in respect to the hypotheses to be proven.

Keywords: Non-small-cell lung cancer; clinical trial; study endpoint; overall survival; progression-free survival



Submitted Dec 24, 2011. Accepted for publication Dec 31, 2011.

doi: 10.3978/j.issn.2218-6751.2011.12.08

Scan to your mobile device or view this article at: <http://www.tlcr.org/article/view/287/599>

Introduction

Non-small-cell lung cancer (NSCLC) is a wide spread disease with a proportion of about 80% in all 1.2 million lung cancers cases per year and is one of the leading causes of malignancy related to morbidity and mortality in the industrialized countries with increasing incidence worldwide (1). In most patients lung cancer is diagnosed in an advanced state with poor prognosis. Because in the last decade's significant progress has been achieved in the biological understanding, tumor heterogeneity of NSCLC became more evident and calls for treatment individualization (2). The identification of novel tumor targets with different pathways has stimulated the search for anti-tumor agents with a specific target directed mode of action. Taking into account subpopulations of patients and special efficacy pathways of new drugs, the gold standard overall survival as an adequate endpoint for clinical trials may have become questionable and has been challenged lately by endpoints not dependent on the growing number of subsequent therapies (3). Therefore, we will discuss in this paper the role of progression free survival as a possible alternative endpoint to overall survival in treatment of NSCLC (4-7).

Methods

Classical study endpoints

Based on the practice in the last centuries there are by definition three categories of classical endpoints generally applied in clinical studies: Survival time endpoints, symptom endpoints, and endpoints relying on patients' reporting. Adjacent to classical endpoints which are measuring single parameters and are tending to show the direct clinical effect of treatment, efforts have been taken to substitute these classical endpoints by surrogates (8-10). Motivated is the use of surrogate endpoints mainly by faster approval processes, smaller sample sizes, positive cost effects and shorter trial duration. The process of substituting a classical endpoint means to replace a rare endpoint or endpoint, describing a late event in the course of disease by a more frequent endpoint or endpoint describing an early event (11,12). It is widely accepted that such an endpoint should be ideally based on a biological rationale and should be related to the study endpoint of interest (13). It may be convenient for clinical trials to base the main hypothesis on tumor parameters, treatment toxicity, and drug pharmacology. The ultimate aim in the majority of patients however, is to prolong the remaining lifespan and reduce tumor morbidity without worsening the performance of the patient by treatment related adverse events, what is commonly understood as clinical benefit.

Survival time endpoints

Lifespan intervals such as progression-free survival (PFS) and

overall survival (OS) are the most common endpoints especially in Phase III studies (14,15). OS measures the clinical benefit of a treatment in total and is used for the regular approval. It is defined as the lifespan starting from recruitment (or first diagnosis of treated disease, or start of first treatment under consideration) to the event of death, what reason so ever. In a study with more than one treatment arm randomization of the study is essential (16). Blinding of the study is not necessary, especially in lung cancer treatments, since side effects are most known and co-medication is described. In ad-on treatment with biologics blinding may be possible (17,18). Censored are patients alive or lost to follow up (19). PFS has the same starting points as OS and includes in its endpoint measurement all deaths, whatever the reason is and all progressions of the tumor under investigation as well as new tumor manifestations. Censored are the patients without such events and those in the lost to follow up situation (20,21). Uncommon survival time endpoints include time to progression (TTP), disease free survival (DFS), recurrence free survival (RFS), and time to treatment failure (TTF) which will not be discussed in the sequel.

Tumor symptom endpoints

Leading symptoms that are directly related to the tumor mass and the extension of the disease can be used as a measure of treatment effect. In Phase III trials they are not so often employed as survival endpoints and defined by rates in a given time interval. The objective response rate (ORR) is a counting measure for the best treatment result (22). The occurrence of partial or complete remission of the tumor is counted as a success and the duration of remission is not taken into account. The number of objective tumor remissions is measured in relation to all evaluable patients in the study at a given time point (23,24). The change in tumor related symptom endpoints such as pain, weight loss, performance status, and dyspnoea can measure the clinical benefit. Studies using these endpoints should be randomized blinded trials. Here, the measurement concerns the perspective of direct clinical benefit for each patient and can also measure the toxic side effects of the study medication (25). Blinding is often difficult (skin rash from the investigated drug!) and data are frequently missing (less evident events) or arrive incomplete (because of the amount of data) in the study center (18).

Patient reported outcomes

Simplicity, instantaneous availability, and proximity to the true health condition of patients are the main advantages in using patients' self-evaluation of their health performance in structured reports for clinical outcome description (26). Because linking patients' information directly to clinical benefit, quality of life (QOL) is important (27). In Europe the questionnaire EORTC

Table 1 Drugs and approval endpoints (33)

Drugs/regimens and approval endpoints	
First-line NSCLC	
Vinorelbine monotherapy	OS, ORR
Vinorelbine + cisplatin	OS, ORR
Docetaxel + cisplatin	OS, TTP, ORR
Gemcitabine + cisplatin	OS, TTP, ORR
Bevacizumab [†] + paclitaxel/carboplatin	OS
Paclitaxel + cisplatin	TTP, ORR, OS
Pemetrexed in combination with cisplatin ^{†,‡}	OS
Pemetrexed as maintenance therapy [†]	OS
Second-line NSCLC	
Docetaxel	OS, TTP, ORR
Erlotinib	OS, TTP, ORR
Pemetrexed ^{†,‡}	Durable ORR, decreased toxicity
Third-line NSCLC	
Erlotinib	OS, TTP, ORR
Gefitinib [‡]	Durable ORR

[†]Limited to non-squamous non-small cell lung cancer;
[‡]Accelerated approval; OS, overall survival; ORR, objective response rate; TTP, time to progression; NSCLC, non-small-cell lung cancer.

QLQ C30 is a widely used instrument translated into 81 languages and employed in more than 3,000 clinical trials (28). The main questionnaire is supplemented by disease specific modules, so LC13 for lung cancer, where 13 special questions concerning lung cancer symptoms are assembled (29,30). In the United States a lung cancer symptom scale (LCSS) and a module of the American Thoracic Society are in application. LCSS was designed as a site-specific measure of QOL for particular use in clinical lung cancer trials evaluating six major tumor symptoms (31). The questionnaire FACT-G (Functional Assessment of Cancer Therapy) is distributed by the American Thoracic Society and includes the special lung cancer module FACT-L (32).

Results

Application of endpoints in clinical studies

One of the main sources worldwide for a standardized procedure of regular drug approval is the FDA. In the FDA guidance draft of June 2011 concerning recommendations for clinical trial endpoints for the approval of anticancer agents three commonly used efficacy endpoints in assessing anti-neoplastic agents are advocated in

NSCLC treatment: OS, PFS, and ORR (19). Nevertheless, the majority of approvals rely on OS. Of 16 trials which have been conducted for drug approval, 11 were based on OS as leading endpoint, 2 on PFS, 2 on durable ORR, and 1 on improvement in disease related symptoms, respectively. It was recently stated that the regular basis of approval for drug in treatment of advanced NSCLC has been OS (*Table 1*) (33). In the FDA perspective, studies have demonstrated that PFS as well as ORR may not reliably predict for a corresponding effect on survival (34,35). Not every detail is clear how to assess a risk/benefit profile that was stipulated for approving new drugs (36). However, it was claimed that an improvement in OS in randomized controlled trials has been the standard for establishing clinical benefit for drug approvals in advanced NSCLC. Not discussed in detail by FDA is the influence of subsequent treatments and special supportive care on OS. It is still hoped that the law of large numbers will rule out differences in comparisons of treatment arms in randomized trials.

Recent advances in molecular biology have led to the development of many anticancer agents targeting on specific target aberrant pathways and proteins of tumor cells. The number of agents available for testing asks for a more efficient and quicker drug testing system (12,37). Surrogate endpoints, which can replace or supplement traditional endpoints, are useful if they can be measured earlier, more conveniently or more frequently than true endpoints (38).

FDA allows accelerated approval on less-established surrogate endpoints, that are reasonably likely to predict clinical benefit and it is explicitly stated that this approval mechanism may only be used when the drug provides a benefit over available therapy (7). If a variable in a clinical study is correlated to a true clinical outcome that does not mean, that it can be used as a valid surrogate endpoint and replace the true clinical outcome (37). FDA still believes that QOL measures can be important clinical benefit endpoints, particularly in a predominantly symptomatic disease such as NSCLC (19). In these patients, with limited survival expectations, symptom palliation, quality of life, and convenience of therapy are especially important endpoints (39). The main argument for these instruments is that health care professionals are not accurate in evaluating subjective or palliative benefits associated with anti-cancer treatments, when compared with patient self-reports. But so far no approval has been based on QOL only. Contradictive facts in using QOL are blinding the trial, reliability and validity of instruments, and the handling of missing data. Since for OS the distinction between two or more treatments will require a larger study sample to show statistically significant differences, PFS is increasingly used as endpoint for randomized Phase II/III studies especially for testing schedules and regimens of targeted therapies.

PFS assessment of disease progression is based on radiologic testing at scheduled time points in cycles and the task is, to deliver reliable results for decision. In statistical models of survival

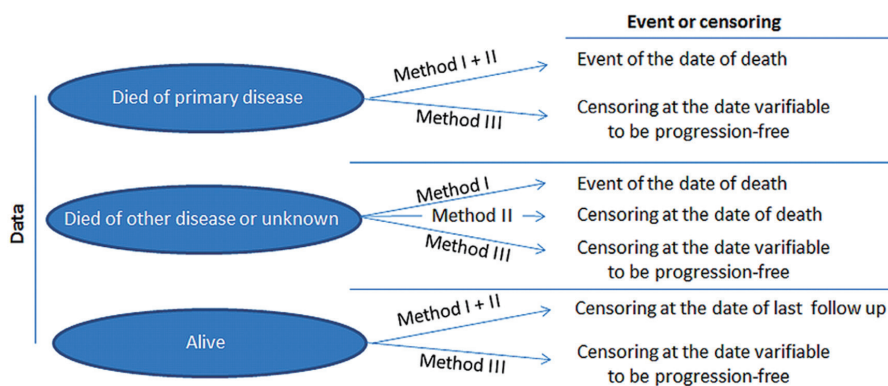


Figure 1 Methods of data handling with missing data on progression.

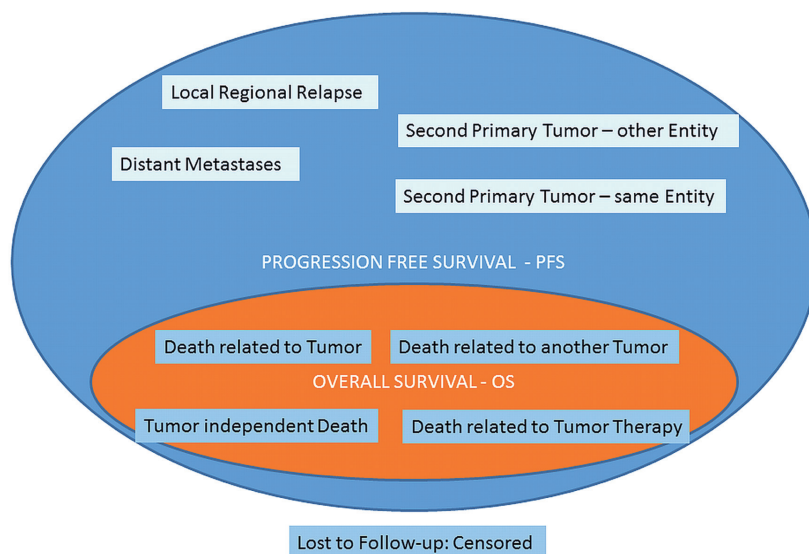


Figure 2 All events of OS are covered by PFS; PFS shows additionally the events of tumor control.

analysis as PFS and OS, not for all subjects the time-to-event will be observed, which will then be censored in the describing statistics. Censoring can be performed in several ways and in Figure 1 different handling methods are demonstrated (20,40). The graphical output of survival data are best presented by Kaplan-Meier (KM) curves where the percentage of survived individuals is plotted against time, which is more reliable than individual survival probabilities. Reliability of KM curve diminishes with increasing time. But there is generally no defined time point or number where the tail of the curve becomes uncertain in interpretation.

At least nine events are related in assessing PFS: (I) Local regional relapse; (II) Distant metastases; (III) Second primary tumor-same entity; (IV) Second primary tumor-other entity; (V) *Death related to tumor; (VI) *Death related to another tumor; (VII) *Tumor independent death; (VIII) *Death related to tumor therapy; (IX) Lost to Follow-up: Censored. All events marked by *

are also covered by OS (Figure 2) (41).

To demonstrate the impact of sample size Figure 3 illustrates calculation for trials where the surplus in overall survival varies. Sample size is calculated by the log rank test (Freedman) under the assumption that in the two groups the hazards are proportional (PASS 11, version 11.0.7, 2011) (42,43). Further on, it is assumed that the survival times are exponentially distributed. If the median survival of the control group is 12 months then the 1-year survival rate is 50%. An increase of 10% to a 1-year rate of 55% is equivalent to a survival of 14 months. A rate of 60% corresponds to 16.3 months, 65% to 19.4 months, and 70% to 23.4 months, nearly doubling the survival time. For superiority the one-sided test is used. The power was set to 90% or 80% and $\alpha=0.025$ or $\alpha=0.05$. Accrual is 1:1 for the two groups. The sample sizes vary between N=4059 and N=151 depending on the power, significance and survival proportion. A small incremental increase in survival

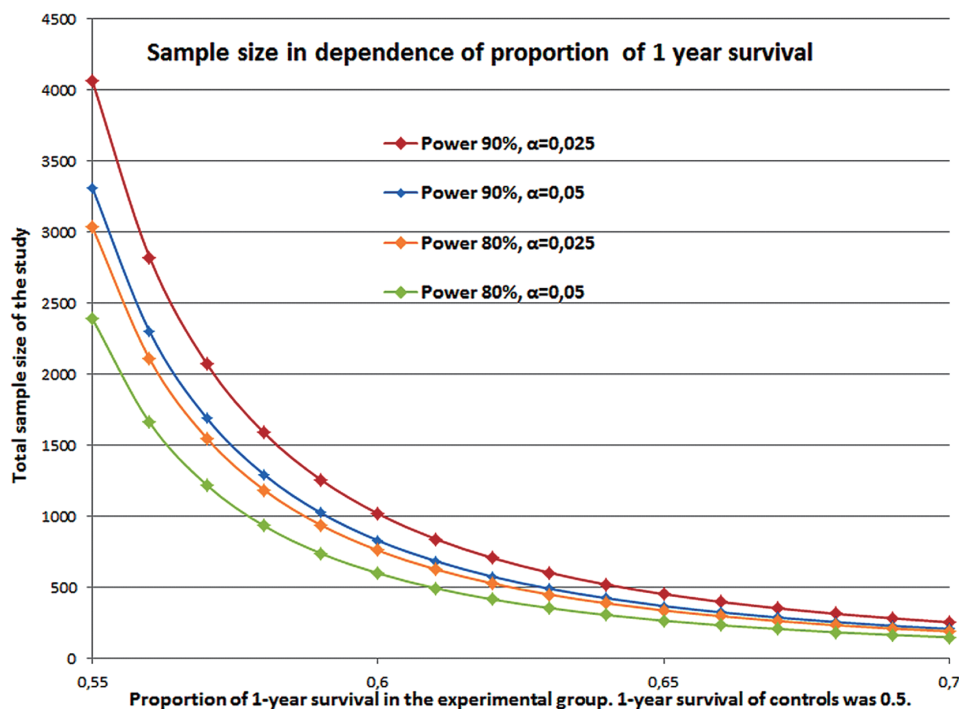


Figure 3 Sample size calculation for proportion of 1-year survival of 0.55-0.70 of experimental group versus 1-year survival of 0.5 for controls in a 1:1 recruitment and with power of 80% or 90% and $\alpha=0.025$ or $\alpha=0.05$.

time accompanied by small error rates (power and α) leads to a large study population.

Substitution of endpoints

The main factor for PFS to predict OS is the duration of PFS (43). A significant survival gain can be predicted in a trial of 750 patients if 18% of patients responding to treatment. 500 patients are needed if 21% of patients respond and if it is postulated that 30% patients respond, at least 250 patients must be enrolled.

The necessary incremental gain for PFS to predict a survival improvement was a median of 1.7 months for trials with 750 patients, 2.2 months for 500 patients, and 3.3 months for trials with 250 patients. Hence, a significant surplus on PFS can only be predictive for OS if the sample is big enough. Statistically it comes to compare PFS and OS as two failure time distributions in the presence of censoring, where the distribution is shifted due to the different time interval of measurement (44). Compared to OS in PFS more competing risks are included (Figure 2). PFS as surrogate endpoint for OS can also be investigated by the aid of a stochastic model (45). The Process was demonstrated theoretically by a statistical OS model with two independent stochastic processes defining an interim and a restart point of progression. Since it has been discovered that the interval after progression to death may have similar distribution for most of the lung cancer trials PFS as

a surrogate of OS seem to be reasonable (46,47). The strongest prediction of PFS for OS results from a high correlation of PFS and OS and a high treatment effect homogeneity (Figure 4).

Quality of life assessment represents an independent outcome in clinical research and an index of effectiveness of a treatment (48,49). The augmentation of PFS by QOL is a combination of two almost independent endpoints depicting a patient self-evaluation of personal health performance and an objective management. Subsequent chemotherapy improvement of tumor related symptoms translates to a better quality of life. Using QOL questionnaires can result in well documented health-related quality of life data and complementary efficacy information, helping in clinical decision finding (50).

Endpoints for targeted drugs

Based on the fact that conventional treatment of NSCLC has reached a plateau of effectiveness in improving the survival new targeted drugs have been developed. Testing targeted therapeutic agents in clinical trials need specific consideration for a number of reasons (51,52). The mode of action of the targeted agent depends specifically on molecular biology and therefore needs an adaption of the study design according to the tumor target concerning study endpoints, patient selection, diagnostic requirements, time frame, etc (53,54). The large number of targeted agents available for

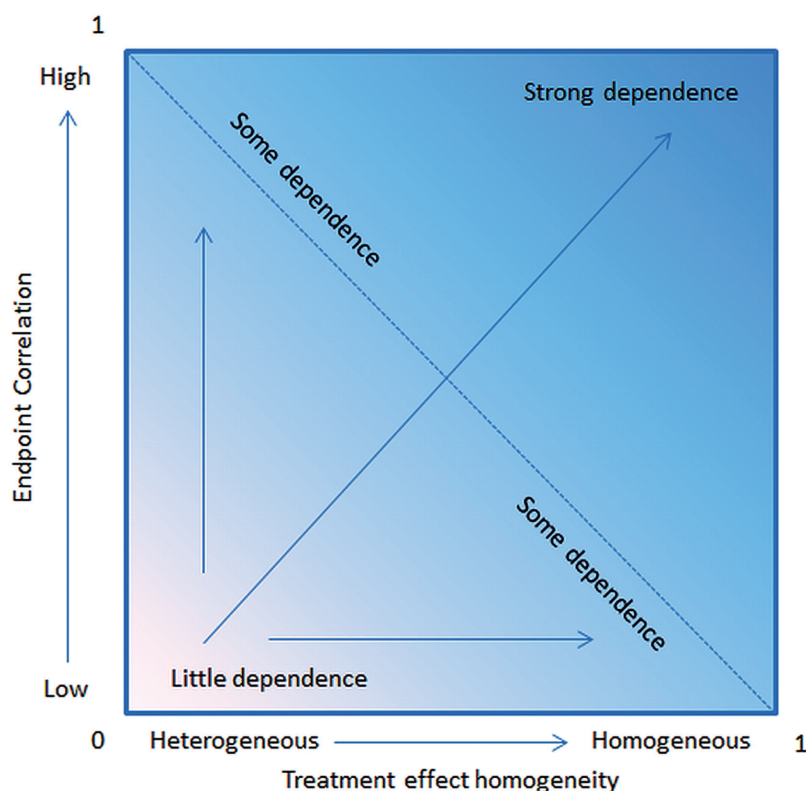


Figure 4 OS prediction depends on the treatment effect homogeneity and the correlation with PFS.

clinical testing certainly is a challenging task in clinical research, which means also to solve specific methodological, logistic, administrative, and statistical issues (55).

If a targeted agent is tested against a placebo arm, and considering a longer time interval for observing progression events, it has been recognized that the drug development is time consuming even using PFS. An outcome-based adaptive randomization trial design for patients can be employed with conjunction of an early stopping rule and can be used to identify effective agents and match those treatments with patients' biomarker profiles (56). Testing the target may cause additional costs and logistic strength (e.g., EGFR-mutation testing). Additionally, to the limitations of PFS as the main endpoint are the limitations of the diagnostic test procedure itself that may lead to misinterpretation of study outcome and in the last consequence to an un-appropriate therapy (40).

Discussion

Merits and shortcomings of study endpoints

So far, OS was used as a gold standard in Phase III lung cancer studies but that concept may not work for the approval of chemotherapeutic drugs and targeted agents. For the choice of OS as endpoint there are two main pitfalls: I. the competition of

death events others than caused by the treated lung cancer disease and the difficulty to distinguish between cancer related and not related deaths and II. the probable long time interval after end of treatment and death of the patient where many lifespan influencing events can happen (57). If the sample size is big enough the statistical rule of the “law of large number” will balance competing risks and under these assumptions overall survival can be accepted as a universal primary endpoint in lung cancer even if the potential causes of other deaths cannot be neglected. More complicated is the statistical handling of a possible long time interval after the end of treatment till death, with the determinants therapy effectiveness, subsequent therapies and salvage therapies making the process not any longer strictly depend directly on the primary treatment action taken (58). Hence, the aspect of respond to primary treatment is not depicted generally in OS. A numerical evaluation of subsequent treatments and influence correcting statistical analysis is almost not possible and depends on the tumor type. Since subsequent treatments are very divers these influences cannot be easily categorized (59-61).

Clearly, progression under treatment is a sign that the disease is not positively affected by treatment action and states the failure of treatment (62). If curative treatment is not available the absence of progression may be evaluated as partial success especially in case of mild side effects of the underlying treatment. For the patient itself

non-progression may be an important signal that the disease has come to a stop with all psychological positive side effects. PFS is the clinical measure of this process.

PFS as measure of response or shrinkage of the tumor can act as an indicator of overall survival. But interpretation remains difficult for various reasons: Many patients showing no progression and no tumor size reduction to treatment, midterm side effects disturb the enduring effect of response and diminishing a beneficial overall survival, and no adequate further treatment options may be available in case of progression (63-65). Measurement of PFS may be constricted by errors in assessing progression and the fixed time schedule of reassessment of the tumor (65). In addition, there is no standardization of measurement in the chosen time interval, tumor type, and tumor mass. As a matter of fact, most of the progressions will happen between the therapy assessments or follow up intervals (66).

PFS as an endpoint is a choice worth considering if multiple therapy options exist. Then it can be compared if only the response of the study drug is measured (67). But PFS is not statistically validated as a surrogate for survival in all settings (68). It can happen, that the results of PFS on the one side and OS on the other side are contradicting. The measurement of progression is not as precise as statement of death in OS and may be subject of assessment bias particularly in open-label studies. Additionally, the definitions in various studies may vary and so it will be complicated to compare trials with different PFS definitions. The diagnostic complexity and costs are high, frequent radiological or other assessments are necessary. Important is the balanced timing of assessment within the study and especially between treatment arms. Sample size is often smaller in PFS in comparison with OS. This fact together with earlier results is an advantage for PFS. But the time interval between progression and death might be large and influence the ability of translating PFS into OS.

Endpoints for lung cancer trials

As in any other cancers, lung cancer trials must determine the clinical benefit of a specific new treatment. The concept of survival time-not necessarily OS-can be used for hypothesis driven statistics as a primary study endpoint and for the sample size calculation. The main thought of these advisements is that all positive actions taken, will translate into a longer overall survival. In the case that the study drug was given only for a short time period all other effects will be ruled out by the law of large numbers, in comparison with a similar population without investigational treatment (control arm). These thoughts are independently true whatever survival time concept is used (69,70).

Following the concept of evidence based medicine the clinical study should have a reasonable hypothesis, which incorporates a primary endpoint to be able to demonstrate the difference between

the investigational and the standard treatment at a statistical significant level. The final evidence is simply provided by testing the given hypothesis. The choice of the endpoint definitely influences the treatment and study success (71). As a basic claim, success should always be postulated as the final result of the study. The chosen endpoint is a determinant of the sample size and has to be selected prospectively in an explicit manner and cannot be altered during the course of the trial. In advanced lung cancer the horizon of the remaining lifespan is quite limited. Thus it is not inequitable that PFS derived under the study drug will translate directly into OS (72).

Lung cancer can be taken as a chronic disease diagnosed often in an advanced stage with a remaining median lifespan of 7 to 15 months. Thus, in advanced lung cancer OS should be a quite natural measure of treatment success (73). Tumor progression is an indicator for disease worsening, a time point for new treatment considerations, and a negative signal for tumor morbidity and death. PFS is measured in the complete study population without exceptions and is an indirect measure for treatment benefit. The main advantages of PFS over OS are that the outcome measure is available earlier, that it can shorten the drug development time, and that there is no influence by subsequent therapies.

PFS as a surrogate for OS should have the inherent considerable advantage, that it can detect subpopulations with longer PFS intervals early. Based on the (sub-) population treated and having in mind the risk-benefit profile of the drug under consideration, PFS can be considered for regulatory decision making. If accompanied by some independent measures like QOL or treatment toxicity, PFS should be able to cover the clinical benefit achieved by treatment. In a recent report on surrogate endpoints in oncology by the German Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) two levels of evaluation were communicated: The validity of the surrogate endpoint should be evaluated and the conclusions should be drawn from the seen effects (74). Selecting PFS as primary endpoint in Phase III trials of advanced NSCLC may be based on a number of questions such as: Does the definition of PFS fit into the setting used by other trials? Are there accepted consensus standards? Are there consistent surveillance intervals? Is validation for each agent group planned? Is the incremental improvement of PFS big enough ($\geq 30\%$)? And are there some additional measures to confine clinical benefit?

Conclusions

OS as a concept of clinical benefit is still accepted as the gold standard in trials investigating advanced NSCLC. OS is easy to measure and precise but it may be difficult to interpret if treatment action takes place at the beginning of the measured survival interval. PFS, basically, does not overcome those difficulties of OS. However, PFS with some additional measures has

become attractive when it seems advisable to make study results available earlier. Candidates for supporting PFS as “additional measures” may be treatment toxicity and quality of life measures. PFS allows a more precise detection and attribution to effects of the investigational treatment without being compromised by subsequent treatments. Therefore “enriched PFS” can be considered as an alternative primary endpoint replacing OS in studies investigating advanced NSCLC. In comparison to OS, PFS is clearly more complex to assess, requires a very stringent definition in clinical trials and as a consequence the ability for a more intensive radiological imaging. Nevertheless, the process of endpoint selection should always be performed carefully judging all true and surrogate endpoint options in respect to the study hypotheses including the projected study design.

Acknowledgements

Disclosure: The authors declare no conflict of interest.

References

- Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin* 2011;61:69-90.
- Holt RJ, Davison T, Dibben S, et al. Development of expression based biomarkers in NSCLC: A study of intratumor heterogeneity using FFPE tissue [abstract]. *J Clin Oncol* 2010;28:e21024.
- Mandrekar SJ, Qi Y, Hillman SL, et al. Endpoints in phase II trials for advanced non-small cell lung cancer. *J Thorac Oncol* 2010;5:3-9.
- Vansteenkiste J. Is there a preferred endpoint? *Lung Cancer* 2009;64:S22.
- Pirker R. Endpoints in clinical trials. *Lung Cancer* 2009;64:S22.
- Fromer MJ. ODAC hashes out endpoints for lung cancer clinical trials. Major sticking points: Time to progression, disease-free survival. *Oncology Times* 2004;26:8-15.
- FDA Lung Cancer Endpoints Final Summary, 2003. Available online: <http://www.fda.gov/AboutFDA/CentersOffices/CDER/ucm120838.htm>
- Flandre P, O'Quigley J. A two-stage procedure for survival studies with surrogate endpoints. *Biometrics* 1995;51:969-76.
- Lara PN Jr, Redman MW, Kelly K, et al. Disease control rate at 8 weeks predicts clinical benefit in advanced non-small-cell lung cancer: results from Southwest Oncology Group randomized trials. *J Clin Oncol* 2008;26:463-7.
- Gralla RJ, Griesinger F. Interpreting clinical trials in lung cancer: impact of methodology and endpoints. *J Thorac Oncol* 2007;2:S51-8.
- Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;8:431-40.
- Adjei AA, Christian M, Ivy P. Novel designs and end points for phase II clinical trials. *Clin Cancer Res* 2009;15:1866-72.
- Ellenberg S, Hamilton JM. Surrogate endpoints in clinical trials: *Cancer. Stat Med* 1989;8:405-13.
- Klein PJ, Goel PK. eds. *Survival analysis: State of the art*. Dordrecht: Kluwer Academic Publisher, 1992.
- Klein PJ, Moeschberger ML. *Survival analysis. Techniques for censored and truncated data*. 2nd ed. New York: Springer-Verlag, 2003.
- Rosenberger WF, Lachin JM. *Randomization in Clinical Trials: Theory and Practice*. New York: John Wiley & Sons, 2002.
- Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ* 2000;321:504.
- Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002;359:696-700.
- Guidance for Industry Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics. Draft Guidance 2011. Available online: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM259421.pdf>
- Shih W. Problems in dealing with missing data and informative censoring in clinical trials. *Curr Control Trials Cardiovasc Med* 2002;3:4.
- Michiels S, Mauguen A, Fisher D, et al. Evaluation of disease-free survival as surrogate endpoint for overall survival using two individual patient data meta-analyses of adjuvant chemotherapy in operable non-small cell lung cancer [abstract]. *J Clin Oncol* 2011;29:s7004.
- Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;13:19-21.
- Baker SG. Two simple approaches for validating a binary surrogate endpoint using data from multiple trials. *Stat Methods Med Res* 2008;17:505-14.
- George SL. Response rate as an endpoint in clinical trials. *J Natl Cancer Inst* 2007;99:98-9.
- Freemantle N, Calvert M. Composite and surrogate outcomes in randomised controlled trials. *BMJ* 2007;334:756-7.
- Thongprasert S, Sanguanmitra P, Juthapan W, et al. Relationship between quality of life and clinical outcomes in advanced non-small cell lung cancer: best supportive care (BSC) versus BSC plus chemotherapy. *Lung Cancer* 1999;24:17-24.
- Patient-reported outcome measurement group, Oxford. A structured review of patient-oriented outcome measures (PROMs) for lung cancer. Report to the department of health, University of Oxford, 2010.
- EORTC QLQ C30. Available online: http://groups.eortc.be/qol/questionnaires_qlqc30.htm
- C13 of EORTC Quality of Life Department. Available online: http://groups.eortc.be/qol/downloads/modules/specimen_20qlq_lc13.pdf
- Bergman B, Aaronson NK, Ahmedzai S, et al. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC Study Group on Quality of Life. *Eur J Cancer*

- 1994;30A:635-42.
31. Lung Cancer symptom Scale (LCSS). Available online: <http://www.lcss-ql.com/>
 32. Lung cancer module FACT-L. Available online: <http://qol.thoracic.org/sections/instruments/fj/pages/fact-g.html>
 33. Malik SM, Ibrahim A, Sridhara R, et al. Use of progression-free survival (PFS) as an endpoint in advanced non-small cell lung cancer (NSCLC) trials: FDA perspective [abstract]. *J Clin Oncol* 28, 2010 (e18001).
 34. Thatcher N, Chang A, Parikh P, et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* 2005;366:1527-37.
 35. Pirker R, Pereira JR, Szczesna A, et al. Cetuximab plus chemotherapy in patients with advanced non-small-cell lung cancer (FLEX): an open-label randomised phase III trial. *Lancet* 2009;373:1525-31.
 36. Kelly K, Crowley J, Bunn PA Jr, et al. Randomized phase III trial of paclitaxel plus carboplatin versus vinorelbine plus cisplatin in the treatment of patients with advanced non-small-cell lung cancer: a Southwest Oncology Group trial. *J Clin Oncol* 2001;19:3210-8.
 37. Johnson JR, Williams G, Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J Clin Oncol* 2003;21:1404-11.
 38. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605-13.
 39. Gralla RJ. Quality-of-life considerations in patients with advanced lung cancer: effect of topotecan on symptom palliation and quality of life. *Oncologist* 2004;9:14-24.
 40. Malik SM, Mansfield E, Waxman IM, et al. Defining success in personalized therapy development: FDA perspective [abstract]. *J Clin Oncol* 28, 2010 (s151).
 41. Chakravarty A, Sridhara R. Use of progression-free survival as a surrogate marker in oncology trials: some regulatory issues. *Stat Methods Med Res* 2008;17:515-8.
 42. Machin D, Campbell M, Fayers P, et al. Sample size tables for clinical studies. 2nd ed. London: Blackwell Science, 1997.
 43. Johnson KR, Ringland C, Stokes BJ, et al. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. *Lancet Oncol* 2006;7:741-6.
 44. Lin DY, Robins JM, Wei LJ. Comparing two failure time distributions in the presence of dependent censoring. *Biometrika* 1996;83:381-93.
 45. Chen TT, Simon RM, Korn EL, et al. Investigation of disease-free survival as a surrogate endpoint for survival in cancer clinical trials. *Commun Statist-Theory Meth* 1998;27:1363-78.
 46. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst* 2009;101:1642-9.
 47. Ulm K. Primary endpoints in cancer trials. *J Thorac Dis* 2011;3:82-3.
 48. Reck M. Quality of life as endpoint of clinical trials. *Lung Cancer* 2009;64:S22-S23.
 49. Tassinari D. Surrogate end points of quality of life assessment: have we really found what we are looking for? *Health Qual Life Outcomes* 2003;1:71.
 50. de Marinis F, Pereira JR, Fossella F, et al. Lung Cancer Symptom Scale outcomes in relation to standard efficacy measures: an analysis of the phase III study of pemetrexed versus docetaxel in advanced non-small cell lung cancer. *J Thorac Oncol* 2008;3:30-6.
 51. Gridelli C, Rossi A, Maione P. Treatment of non-small-cell lung cancer: state of the art and development of new biologic agents. *Oncogene* 2003;22:6629-38.
 52. Schilsky RL. Target practice: oncology drug development in the era of genomic medicine. *Clin Trials* 2007;4:163-6; discussion 173-7.
 53. Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. *Stat Med* 2006;25:183-203.
 54. Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res* 2008;17:303-40.
 55. Di Maio M, Gallo C, De Maio E, et al. Methodological aspects of lung cancer clinical trials in the era of targeted agents. *Lung Cancer* 2010;67:127-35.
 56. Zhou X, Liu S, Kim ES, et al. Bayesian adaptive design for targeted therapy development in lung cancer - a step toward personalized medicine. *Clin Trials* 2008;5:181-93.
 57. Vincent MD. Optimizing the management of advanced non-small-cell lung cancer: a personal view. *Curr Oncol* 2009;16:9-21.
 58. Di Leo A, Buyse M, Bleiberg H. Is overall survival a realistic primary end point in advanced colorectal cancer studies? A critical assessment based on four clinical trials comparing fluorouracil plus leucovorin with the same treatment combined either with oxaliplatin or with CPT-11. *Ann Oncol* 2004;15:545-9.
 59. Tuma R. Progression-free survival remains debatable endpoint in cancer trials. *J Natl Cancer Inst* 2009;101:1439-41.
 60. Burzykowski T. Surrogate endpoints: wishful thinking or reality? *Stat Methods Med Res* 2008;17:463-6.
 61. Buyse M, Burzykowski T, Michiels S, et al. Individual- and trial-level surrogacy in colorectal cancer. *Stat Methods Med Res* 2008;17:467-75.
 62. Buyse M, Quinaux E, Hendlisz A, et al. Progression-free survival ratio as end point for phase II trials in advanced solid tumors. *J Clin Oncol* 2011;29:e451-2; author reply e453.
 63. Fleming TR, Rothmann MD, Lu HL. Issues in using progression-free survival when evaluating oncology products. *J Clin Oncol* 2009;27:2874-80.
 64. Karrison TG, Maitland ML, Stadler WM, et al. Design of phase II cancer trials using a continuous endpoint of change in tumor

- size: application to a study of sorafenib and erlotinib in non small-cell lung cancer. *J Natl Cancer Inst* 2007;99:1455-61.
65. Hunsberger S, Albert PS, Dodd L. Analysis of progression-free survival data using a discrete time survival model that incorporates measurements with and without diagnostic error. *Clin Trials* 2010;7:634-42.
 66. 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.
 67. Soria JC, Massard C, Le Chevalier T. Should progression-free survival be the primary measure of efficacy for advanced NSCLC therapy? *Ann Oncol* 2010;21:2324-32.
 68. Baker SG. Surrogate endpoints: wishful thinking or reality? *J Natl Cancer Inst* 2006;98:502-3.
 69. Vansteenkiste JF, Vandebroek JE, Nackaerts KL, et al. Clinical-benefit response in advanced non-small-cell lung cancer: A multicentre prospective randomised phase III study of single agent gemcitabine versus cisplatin-vindesine. *Ann Oncol* 2001;12:1221-30.
 70. Steer CB, Marx GM, Harper PG. Is there quality in clinical benefit? *Ann Oncol* 2001;12:1191-3.
 71. European Medicines Agency: ICH Topic E 9. Statistical Principles for Clinical Trials. Note for guidance on statistical principles for clinical trials (CPMP/ICH/363/96). Available online: <http://www.emea.eu.int>
 72. Stone A, Wheeler C, Carroll K, et al. Optimizing randomized phase II trials assessing tumor progression. *Contemp Clin Trials* 2007;28:146-52.
 73. Perrone F. Don't forget survival, please... *Lancet Oncol* 2006;7:703-4.
 74. IQWiG-Berichte-Jahr: 2011 Nr. 80: Aussagekraft von Surrogatendpunkten in der Onkologie. Auftrag: A10-05 Version: 1.0 Stand: 31.01.2011.

Cite this article as: Pilz LR, Manegold C, Schmid-Bindert G. Statistical considerations and endpoints for clinical lung cancer studies: can progression free survival (PFS) substitute overall survival (OS) as a valid endpoint in clinical trials for advanced non-small-cell lung cancer? *Transl Lung Cancer Res* 2012;1(1):26-35. doi: 10.3978/j.issn.2218-6751.2011.12.08