PERGAMON

European Journal of Cancer 38 (2002) 1817–1823

European Journal of Cancer

www.ejconline.com

Position Paper

Measuring the clinical response. What does it mean?

P. Therasse*

EORTC Data Center, Av. E. Mounier, 83/11, 1200 Brussels, Belgium

Received 30 May 2002; accepted 30 May 2002

Abstract

The clinical response to treatment is an important indicator of the therapeutic effect of anticancer agents. Its value and interpretation has to be carefully considered within the context that it is used. In daily practice, response assessment is combined with other indicators of the patient's condition to contribute to the decision-making process. In clinical trials, it is widely used to identify and quantify the anti-tumour activity of new agents. In this context, response evaluation is conducted on the basis of strict predefined criteria such as the World Health Organization (WHO) or Response Evaluation Criteria In Solid Tumors (RECIST) criteria. The RECIST criteria have recently been proposed and offer a detailed guidance to perform a response evaluation. Clinical response is also used as an indicator of therapeutic efficacy in combination with other indicators. Its value as a surrogate indicator of a survival benefit remains unclear in most instances and can hardly be established within the framework of a single randomised trial. With the development of new anticancer agents that behave differently to cytotoxics, clinical benefit will have to integrate concepts of disease stabilisation or time to progression. Over the next decade, oncologists will be able to assess the biological response before the clinical response, and a lot of work and energy will have to be dedicated to assess the predictive and, possibly, the prognostic value of the biological response with regard to the clinical response, as well as more definitive measures of clinical benefit.

Keywords: Clinical trials; Clinical response; RECIST; Surrogate endpoints

1. Introduction

In cancer management, evaluating the therapeutic effect of anticancer therapy is a process undertaken every day by oncologists. In most instances, decisions to continue, change or stop systemic therapy are driven by the response to treatment recorded for each patient.

The same approach is also applied under predefined conditions to test new anticancer agents in clinical trials and quantify their level of antitumour activity. Sometimes, it also contributes with other indicators to define the real clinical benefit (the efficacy) provided to the patients with new therapeutic strategies.

Because the clinical response is the only indicator readily available to evaluate the therapeutic effect of anticancer treatment, many oncologists are tempted to use this indicator as a surrogate of long-term clinical benefit for the patients. Unfortunately, such correlation between response and long-term benefit has rarely been demonstrated.

The methodology used to evaluate the response to treatment has also substantially evolved over the past decades, starting from a complete subjective evaluation reported by the treating physician [1] to move to a complex set of objective criteria attempting to standardise the response evaluation process [2–6].

More recently, several new classes of anticancer agents have been discovered. These new drugs often operate through different mechanisms than those previously developed inducing massive cell kill. Accordingly, the methodology used to evaluate clinical response will not only require adaptation to use new tools and techniques to monitor response to treatment, but may also require a subtle different approach to monitor the therapeutic effect of these new classes of anticancer agents.

* Tel.: +32-2-774-1614; fax: +32-2-772-6197.
E-mail address: pth@eortc.be (P. Therasse).

0959-8049/02/$ - see front matter
PII: S0959-8049(02)00182-X

2. Measuring response to treatment in daily practice

In the daily practice of oncologists, the clinical response reported after each patient's visit results from the combination of different indicators out of which the most important are the response to treatment of the anatomical indicators (tumour lesions), the biological indicators (tumour markers and biochemistry) and the patient's condition.

The clinical response so reported, directly contributes to the evaluation of the risk/benefit ratio procured by a certain treatment which also takes into account the subjective opinion of the physician and the patient's performance (Fig. 1).

In this setting, criteria used to determine the clinical response must be adapted to the real life, taking into account the socio-economical constraints (costs, insurance and resources) and the comfort of the patient. The global risk/benefit ratio of a certain standard treatment is supposed to be known (from previous clinical trials) and the role of the physician is to adjust the treatment to the specific conditions of each patient.

In this setting, it is obvious that rigorous criteria to define the clinical response should not be applied systematically. It is important, however, to rely on robust and reliable assessments to support important decisions such as initiating or changing systemic therapy.

3. Measuring the clinical response to determine the antitumour activity of new anticancer agents

The evaluation of the response rate to determine the level of antitumour activity of new anticancer agents or new combination of existing agents is performed in clinical trials. These are usually phase II clinical trials with the determination of the response rate being the main endpoint. Several statistical standard designs (Gehan, Simon, Fleming) [7] are used to identify and also quantify the biological antitumour activity of anticancer agents. Such evaluation can be done qualitatively against the natural history of the disease or quantitatively against a known level of antitumour activity provided by standard therapeutic strategies.

Measuring the clinical response to determine the antitumour activity of new anticancer agents requires the application of rigorous criteria for several reasons mentioned below.

1. The investigation is conducted on a small cohort of patients who decrease the chance of treating a large cohort of patients with an inactive drug, but increase the probability of uncontrolled biases.
2. The therapeutic window of anticancer agents is usually narrow and the conditions under which a new treatment is tested should be optimised to avoid disregarding a potentially active drug.
3. The evaluation of the response rate in phase II clinical trials is a critical step in the development of new agents and just precedes the large clinical trials involving thousands of patients and the submission of a registration to regulatory authorities.
4. Often, several parallel identical investigations are conducted. It is expected that the results of these clinical trials will all show the same outcome. Reproducibility of the results will be assessed by regulatory authorities.
5. Finally, the methodology on which the evaluation of clinical response has been developed is fragile, complex and relatively imprecise with significant interobserver variability. Rigorous criteria should be used to decrease the probability of errors, misinterpretation and potential harm to future patients.

It is indeed true that the methodology used to evaluate the clinical response under this setting does not cor-

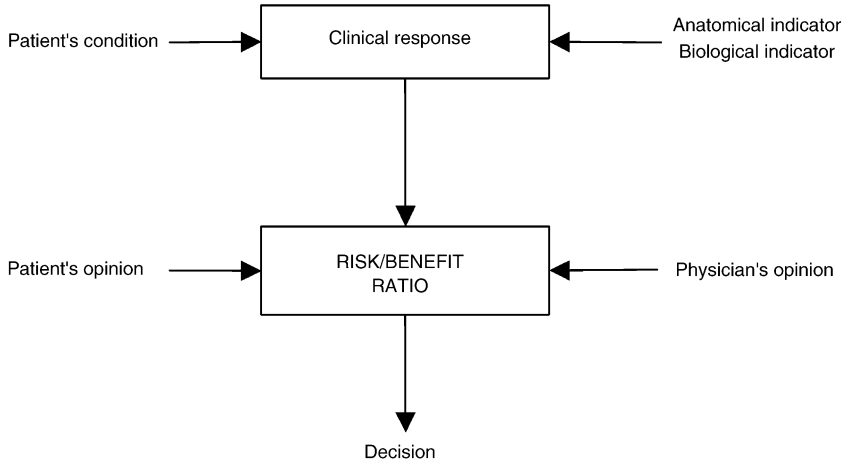

Fig. 1. The role of clinical response in daily practice.

Table 1
Response rate reported by different authors before and after independent review

Author [Ref.]	Tumour type	Patients (*N*)	RR before review (%)	RR after review (%)	% of responses downgraded
Gwyther [8]	NSCLC	374	30	21	NA
Biganzoli [9]	MBC	564	36	30	19
Thiesse [10]	RCC	489	17	13	23

RR, response rate; NSCLC, non-small cell lung cancer; MBC, metastatic breast cancer; RCC, renal cell carcinoma; NA, not available.

respond to 'real life' conditions. Nevertheless, it is, so far, the best methodology that has been used successfully to screen anticancer agents and select those to be further investigated in a large cohort of patients to determine their efficacy. The weakness of the methodology as mentioned above can be illustrated by several examples.

Table 1 illustrates, for three different tumour types, the variation in response rate reported by the investigator compared with the response rate reported after response review. It is interesting to note that not only the response rate is systematically lower after response review, but also that the percentage of response downgraded is substantial. Moreover, Thiesse and colleagues have identified the causes leading to errors which could be attributed, in his series of 489 patients with renal cell carcinomas, to error in tumour measurement in 45% of the cases, error in the selection of target lesions in 45% of the cases, and in 10% of the cases errors that were relatively independent of the evaluator.

Table 2 illustrates the variation reported by Schrijvers and Vermorken [11] for phase II clinical trials investigating single agent or combination of agents in head and neck tumours. It is clear that these findings illustrate not only random variations that cannot be easily controlled for in phase II, trials but also variations as to how the response rate and the individual response for each patient have been evaluated.

Table 3 illustrates the small cohort effect of phase II clinical trials. Although the examples mentioned are both related to breast cancer, the conclusion that response rates seen in phase II trials are systematically higher than those reported in phase III trials (under the same conditions) can be generalised to all tumour types.

Table 2
Reported ranges of response rate for single agent and combination chemotherapy in head and neck cancer

Single agent		Combination	
Agent	RR (%)	Agent	RR (%)
Paclitaxel (Taxol)	20–43	Paclitaxel + cisplatin	33–77
Docetaxel (Taxotere)	37–45	Paclitaxel + carboplatin	23–62
Topotecan	0–14		
Vinorelbine (Navelbine)	11–25		

RR, response rate.

These differences in response rates between phase II and phase III trials can be attributed to many different causes. Some of these causes can probably be controlled for such as the selection criteria (ensuring the homogeneity of the population being studied) and the compliance with the protocol (precision in measurement and follow-up). However, these factors alone can hardly explain the large differences observed between response rates.

3.1. The methodology to evaluate the antitumour activity

In the first clinical trial in solid tumours initiated in the 1950s, tumour response was already taken as an endpoint based on the subjective evaluation reported by the physicians [1].

By the end of the 1970s, a group of breast cancer specialists, under the auspices of the International Union against Cancer (UICC), set the principles under which response to treatment in breast cancer should be evaluated [2]. This work was subsequently adopted and integrated into the recommendations set by the World Health Organization (WHO) for the evaluation of cancer treatment in solid tumours [3]. The principles of response evaluation, which are still valid today, can be summarised as follows:

- The overall cancer burden can be characterised by a quantitative evaluation of tumour lesions, which are measurable, and a qualitative evaluation of tumour lesions, which are not measurable.
- The combination of the quantitative and qualitative evaluations provides an estimation of the treatment effect characterised by one of the following four categories: complete response, partial response, stable disease and progression.

The specificity of the WHO/UICC criteria, amongst others, is:

- Measurable lesions are characterised either by their surface (bidimensional lesions, product of longest perpendicular diameters) or by the longest diameter (unidimensional lesions) when only one dimension could be accurately measured.

Table 3
Response rate reported by different authors for phase II and phase III investigating the same regimens

Author [Ref.]	Development phase	Patients (*N*)	Response rate (%)
	Doxorubicin (Adriamycin) + Paclitaxel in metastatic breast cancer		
Gianni [12]	Phase II	35	94
Biganzoli [13]	Phase III randomised	138 (1 arm)	58
	Epirubicin + cyclophosphamide + G-CSF in locally advanced breast cancer		
Piccart [14]	Phase II	29	87
Therasse [15]	Phase III randomised	220 (1 arm)	57

G-CSF, granulocyte-colony stimulating factor.

- The tumour load is evaluated for each organ independently and the overall response to treatment results from the combination of the response observed in each organ.
- Partial response is attributed when a decrease of 50% of the entire tumour burden (objectively for measurable lesions and subjectively for others) is recorded. Progression status is assigned when there is an increase of 25% of the entire tumour burden (based on the same principles as for PR).

After 1981, many non-anticancer drugs have been developed, and many researchers have also started to investigate different combinations of treatments. The experience acquired over the years and the lack of details in the WHO recommendations has stimulated the development and the use of the amended version of the WHO criteria. For example, the South West Oncology Group (SWOG) published their version of the WHO criteria in 1990 [4], with a different cut-off point to define the progression status. In addition, the European Organization for Research and Treatment of Cancer (EORTC) published its version of the WHO criteria (5) defining minimum sizes for lesions from different organs to be considered as measurable.

Over the years, the use of the different versions (published and unpublished) of the original WHO criteria have rendered the accuracy of the comparison of results of identical investigations very unreliable. In addition, the evolution of cancer imaging, the importance given to the response rate endpoint and the increasing number of new anticancer agents to be tested, required a coordinated effort to review the existing criteria and attempt to 're-harmonise' the methodology throughout the entire oncology community.

This difficult exercise started in 1996 under the leadership of three research organisations: EORTC, National Cancer Institute of the United States (NCI US) and National Cancer Institute Canada—Clinical Trials Group (NCIC CTG). A comprehensive revised version of the WHO criteria was published by the group in February 2000 [6] under the acronym of RECIST (Response Evaluation Criteria In Solid Tumors). Within 2 years, these criteria have been adopted by most research groups, the pharmaceutical industry and the regulatory agencies.

The specificity of the RECIST criteria can be summarised as follows:

- All measurements of tumour lesions are based on the longest diameter only (unidimensional measurement).
- Cancer lesions are measurable when their longest diameter is $\geqslant 2$ cm when measured with conventional techniques or $\geqslant 1$ cm when measured by spiral computed tomography (CT) scan.
- Precisions are given as to which method of measurement can be used and how it can be used.
- The overall tumour burden is represented by selected target lesions and all other lesions are recorded, but not measured.
- Changes in the sum of the longest diameters of all target lesions will define the status of partial response and stable disease.
- Partial response status is defined when the sum of the longest diameters of all target lesions has decreased by 50% or more.
- Response status should be confirmed with a minimum interval of 4 weeks. Stable disease is defined following an interval between two measurements that is protocol-specific (depending on the disease being studied).
- Progression status is defined by an increase of 20% of the sum of the longest diameters of all target lesions or by a non-equivocal progression in non-target lesions or by the appearance of a new lesion.
- Precisions are provided as to how to combine the results of the evaluation of target and non-target lesions and define the overall response.
- Precisions are provided as to how to interpret successive evaluations to define the best overall response.
- CA125 can be used as an indicator that alone may determine a progression of the disease after first-line treatment in advanced ovarian cancer.

The use of one dimension only to measure tumour lesions has been based on the work published by James and colleagues [16]. Retrospective analysis using a cohort of patients from 14 different studies ($\geqslant$4000 patients) demonstrated that using two dimensions or one dimension for tumour lesions measurement did not change the response rate of each individual study. The rate of progression may be slightly different (lower with the RECIST criteria) since a larger difference in tumour growth is required to define disease progression.

Although the RECIST criteria have been launched in 2000, the harmonisation process of the response criteria has continued through the set-up of a Questions and Answers section on the web (eortc.recist.be). In addition, proposals to modify the existing criteria or add new criteria are considered regularly by the RECIST working group. Adaptation of the RECIST criteria are currently studied for specific tumour types such as brain tumours, mesothelioma and pelvic tumours.

Beside the RECIST criteria developed to be applicable to most solid tumours types, standard criteria have been developed for evaluating response and progression in specific settings such as non-Hodgkin's lymphoma [17] and prostate cancer [18].

4. Clinical response as an indicator of treatment efficacy in phase III clinical trials

How can we demonstrate the efficacy of a new treatment in oncology? In other words, do we need to demonstrate an improvement in long-term survival? An improvement in time to progression? An improvement of quality of life? A better control of the symptoms of the patient or perhaps simply an improvement in the clinical response rate?

All these endpoints (either primary or secondary objectives) of clinical trials are potential valid indicators of treatment efficacy when they are directly related to an improvement of the risk/benefit ratio for the patients.

One of these endpoints can be preferred to the others according to the context of the study (early disease versus advanced disease, symptomatic versus non-symptomatic disease and so on...). In the advanced setting, when patients usually present measurable lesions, time to progression and/or clinical response rate is often taken as primary endpoint of the trial. The latter, of course, provides a certain advantage over the other in as much as it can provide an answer relatively quickly to the question being investigated.

In this context, one may question the real meaning of clinical response in terms of improvement of the risk/benefit ratio. This is not an easy question and it requires, *a priori*, some clarification over the relationship between clinical response and long-term benefit or in other words to what extent is clinical response a surrogate indicator of another measure of treatment efficacy? It is well known that, in the long-term, those who usually respond to the treatment will do better than the others. This is what oncologists observed in their daily practice and this is also what subset analysis of clinical trials can demonstrate (comparing the survival of the responders with the non-responders). Indeed, responders do better, but does it make clinical response a valid surrogate of survival? The answer to that question is no. Making such a correlation between clinical response and survival is a well known pitfall [19]. In this particular case, selecting responders to analyse and compare their survival with the group of patients that does not respond to treatment is in fact equivalent to selecting a group of patients with specific characteristics including known and unknown prognostic factors that can influence the outcome, both in terms of response to treatment and in terms of survival for this subgroup of patients [20]. However, the effect produced by the treatment on the tumour can be mediated through different bio-molecular mechanisms for response and survival.

One specific outcome measure (such as clinical response) can be considered as a true surrogate indicator of another outcome measure (such as survival) only when the effect of the treatment on the surrogate can reliably predict the effect of the treatment on the final clinical outcome [21]. In other words, the treatment effect on the disease globally should be entirely mediated through the effect seen on the surrogate marker. Such a correlation has rarely been demonstrated in all disciplines of medicine. First of all, it requires very large data-sets of clinical data of patients treated under relatively identical conditions. Moreover, the statistical methodology deployed to prove such correlation and remove all potential confounding factors is extremely complex.

In oncology, several groups have attempted to demonstrate such a correlation between response and long-term outcome in breast [22], ovarian [23], non-small cell lung [24] and colorectal cancers [25]. Only in one study [25] could response to treatment be identified as an independent prognostic factor that affects survival. That study by Buyse and colleagues also highlights a number of important points such as:

- The relationship between response and survival may depend on the drug, the schedule and the dose for the same disease.
- Large improvements in clinical response are needed to achieve a meaningful improvement in survival. In metastatic colorectal cancer, a two-fold increase in response rate corresponded to a 1.12 increase in survival. In addition, the overall correlation between clinical response and survival may be substantially influenced by the rate of complete response (and even more by the rate of complete pathological response), but also by the

efficacy of second-line treatments (the better it is, the more difficult it is to establish the relationship).

The main conclusion of this exercise was that, for individual trials, the response rate alone cannot realistically predict the benefit for survival. However, a good clinical response rate should trigger phase III trials with more definitive outcome measures and should also encourage an extension of the drug development in the adjuvant setting, even without a survival benefit in the advanced setting.

5. Clinical response as a measure of efficacy

Using clinical response as a direct measure of treatment efficacy may be relevant under specific conditions [26]. The assumption that clinical response may indicate a certain clinical benefit will not only depend on the observed response rate and the degree of improvement over the existing standards of treatment, but should also take into consideration other characteristics of the responses observed and the drug studied.

The average duration of response, the rate of complete response (and in particular complete pathological response) and the localisation of sites responding to treatment are important characteristics. The pharmacological profile of the drug (and in particular the toxicity profile) together with the previous experience observed with the same class of drugs and in the same population are also important.

Finally, the reproducibility of the response rate in other clinical trials should confirm the overall trend observed with a particular drug. It is clear that in this setting no firm rules can be established and the opportunity to use clinical response alone or in association with other indicators of clinical benefit should be considered on a case by case basis.

6. Measuring clinical response in phase III clinical trials

In phase III clinical trials attempting to demonstrate a definitive efficacy advantage of the treatment being studied, the response rate is usually used as a secondary endpoint which may on the one hand support the primary endpoint and, on the other hand, may also be used to adjust the response rate reported from the phase II data under conditions which are usually closer to real life.

Under these conditions, the evaluation of response in phase III clinical trials may not require the same rigour as for phase II trials aiming at determining the degree of antitumour activity. More flexibility could be considered for specific requirements such as the necessity for confirmation, the number of selected target lesions, and the necessity for response review... However, when the response rate is used as a primary endpoint (which usually implies that sample size calculations are driven by a target difference in response rate between the two treatments) a rigorous methodology, as used for phase II trials should be used.

Trials with this main objective can indeed provide a rapid answer in comparison with those using time to event endpoints. However, they require much more resources and time from all of the involved parties and they are usually under-powered to assess more definitive objectives listed as secondary endpoints. The literature is overloaded with small inconclusive phase III trials that were developed and conducted using clinical response as their main endpoint.

7. Future development

Changes in the methodology developed to evaluate the clinical response will not only depend on progress in imaging technology, but will also be affected by the 'new' classes of anticancer agents that are under development [27,28]. Amongst these new drugs, those having a biological antiproliferative effect inducing delays in tumour growth should be carefully evaluated. It is conceivable that these agents might not systematically generate rapid tumour regression (and therefore measurable response), but may simply result in stabilisation of the disease or even may decrease the rate of tumour growth. For these agents, more attention will have to be given to stable disease and time to progression than to a pure clinical response depending on tumour shrinkage. However, although the endpoints of clinical trials may have to be revised, shifting from response/survival towards progression/stabilisation, the methodology used to assess response remains valid for assessing stabilisation and/or progression.

In the future, molecular responses will first be considered based on functional imaging providing early indicators of antitumour activity. Techniques for documenting anticancer effects will be targeted to follow the mechanism of the anticancer agents tested such as: Positron Emission Tomography (PET) scanning to monitor the glucose metabolism and indirectly the proliferative activity; Magnetic Resonance Spectroscopy (MRS) to analyse the cellular energetics and the membrane turnover; Magnetic Resonance Imaging (MRI) to document tumour perfusion, vascularity and permeability and molecular imaging to follow the intracellular signalling pathways, as well as to monitor gene and drug delivery.

However, even the most developed of these techniques still requires validation with regards to the interpreta-

tion of the results in terms of 'response to treatment'. Large correlative studies with the current criteria and standard clinical outcome measures will be required. When all of these issues have been solved, cost issues will have to be addressed to confirm the cost-effectiveness of these new techniques before they can be implemented in clinical trials and routine care.

References

1. Zubrod CG, Schneiderman SM, Frei E III, *et al*. Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and thio phosphamide. *J Chronic Dis* 1960, **11**, 7–33.
2. Hayward JL, Carbone PP, Heuson JC, *et al*. Assessment of response to therapy in advanced breast cancer: a project of the Programme on Clinical Oncology of the International Union Against Cancer, Geneva, Switzerland. *Cancer* 1977, **39**, 1289–1294.
3. Miller AB, Hogestraeten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981, **47**, 207–214.
4. Green S, Weiss GR. Southwest Oncology Group standard response criteria, endpoint definitions and toxicity criteria. *Invest New Drugs* 1992, **10**, 239–253.
5. Tumor eligibility and response criteria for phase II and III studies. In *EORTC Data Center Manuel*. EORTC, Brussels, 1992 (monograph).
6. Warr D, McKinney S, Tannock I. Influence of measurement error on assessment of response to anticancer chemotherapy: proposal for new criteria of tumor response. *J Clin Oncol* 1984, **2**, 1040–1046.
7. Kramar A, Potvin D, Hill C. Multistage designs for phase II clinical trials: statistical issues in cancer research. *Br J Cancer* 1996, **74**, 1317–1320.
8. Gwyther SJ, AApro MS, Hatty SR, Postmus PE, Smith IE. Results of an independent oncology review board of pivotal clinical trials of gemcitabine in non-small cell lung cancer. *Anticancer Drugs* 1999, **10**, 693–698.
9. Biganzoli L, Lohrisch C, Paridaens R, Therasse P, Piccart M. Analysis of two EORTC trials in metastatic breast cancer (MBC) support the recommendation of external response review (ERR) in trials with response rate (RR) as primary endpoint. *Eur J Cancer* **36**(Suppl. 5), S90 (abstr 227).
10. Thiesse P, Ollivier L, Di Stefano-Louineau D, *et al*. Response rate accuracy in oncology trials: reasons for interobserver variability. *J Clin Oncol* 1997, **15**, 3507–3514.
11. Schrijvers D, Vermorken JB. Update on the taxoids and other new agents in head and neck cancer therapy. *Curr Opin Oncol* 1998, **10**, 233–241.
12. Gianni L, Munzone E, Capri G, *et al*. Paclitaxel by 3-hour infusion in combination with bolus doxorubicin in women with untreated metastatic breast cancer: high antitumor efficacy and cardiac effects in a dose-finding and sequence-finding study. *J Clin Oncol* 1995, **13**, 2688–2699.
13. Biganzoli L, Cufer T, Bruning P, *et al*. Doxorubicin (A)/Taxol (T) versus doxorubicin/Cyclophosphamide (C) as first line chemotherapy in metastatic breast cancer (MBC): a phase III study. *Proceedings ASCO* 2000 (abstr 282).
14. Piccart MJ, Bruning P, Wildiers J, *et al*. An EORTC pilot study of filgrastim (recombinant human granulocyte colony stimulating factor) as support to a high dose-intensive epiadriamycin-cyclophosphamide regimen in chemotherapy-naive patients with locally advanced or metastatic breast cancer. *Ann Oncol* 1995, **6**, 673–677.
15. Therasse P, Mauriac L, Welnicka M, *et al*. Neo-adjuvant dose intensive chemotherapy in locally advanced breast cancer (LABC): an EORTC-NCIC-SAKK randomized phase III study comparing FEC (5FU, epirubicin, cyclophosphamide) vs high dose intensity EC + G-CSF (Filgrastim). *Proceedings ASCO* 1998 (abstr 472).
16. James K, Eisenhauer E, Christian M, *et al*. Measuring response in solid tumors: unidimensional versus bidimensional measurement. *J Natl Cancer Inst* 1999, **91**, 523–528.
17. Cheson BD, Horning SJ, Coiffier B, *et al*. Report of an international workshop to standardise response criteria for non-Hodgkin's lymphomas. *J Clin Oncol* 1999, **17**, 1244–1253.
18. Bubley GJ, Carducci M, Dahut W, *et al*. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the prostate-specific antigen working group. *J Clin Oncol* 1999, **17**, 3461–3467.
19. Weiss GB, Bunce H III, Hokanson JA. Comparing survival of responders and nonresponders after treatment: a potential source of confusion in interpreting cancer clinical trials. *Controlled Clin Trials* 1983, **4**, 43–52.
20. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1966, **125**, 605–613.
21. Buyse M. On the relationship between response to treatment and survival time. *Stat Med* 1996, **15**, 2797–2812.
22. A'Hern RP, Ebbs SR, Baum MB. Does chemotherapy improve survival in advanced breast cancer? A statistical overview. *Br J Cancer* 1988, **57**, 615–618.
23. Torri V, Simon R, Russek-Cohen E, Midthune D, Friedman M. Statistical model to determine the relationship of response and survival in patients with advanced ovarian cancer treated with chemotherapy. *J Natl Cancer Inst* 1992, **84**, 407–414.
24. Chen TT, Chute JP, Feigal E, Johnson BE, Simon R. A model to select chemotherapy regimens for phase III trials for extensive-stage small-cell lung cancer. *J Natl Cancer Inst* 2000, **92**, 1601–1607.
25. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumor response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. *Lancet* 2000, **356**, 373–378.
26. Pazdur R. Response rates, survival, and chemotherapy trials. *J Natl Cancer Inst* 2000, **92**, 1552–1553.
27. Buthiau D, Khayat D. New imaging techniques and clinical applications: visualizing pathology's course. *Helix* 2000, **3**, 4–13.
28. Rowinsky EK. The pursuit of optimal outcomes in cancer therapy in a new age of rationally designed target-based anticancer agents. *Drugs* 2000, **60**(Suppl. 1), 1–14.