Editorial

End Points in Cancer Clinical Trials and the Drug Approval Process[1]

Richard L. Schilsky[2]

Medicine and Associate Dean for Clinical Research, Biological Sciences Division, University of Chicago, Chicago, Illinois 60637

Abstract

The sequencing of the human genome and the elucidation of many molecular pathways important in cancer cell proliferation, apoptosis, and metastasis have provided unprecedented opportunities for development of new agents to prevent and treat cancer. The types of molecules in development are increasingly varied and include small molecules, monoclonal antibodies, antisense oligonucleotides, and ribozymes. Thus, the variety of anticancer agents in clinical development is now greater than ever before, and the number of agents currently in clinical trial for various cancer indications is estimated to exceed 400. Many of these drugs would be expected to work in only narrowly defined patient populations that must be prospectively identified. Thus, the development of the therapeutic agent must often be linked to the development of a molecular diagnostic product. Drugs that produce primarily cytostatic effects might not be expected to produce regression of tumor masses; thus, evaluation of such agents would best be done in populations of patients with low tumor burdens but high risk of disease progression. As traditional clinical end points prove more difficult to apply in evaluation of molecularly targeted therapies, a great need exists to define and validate surrogate markers of effect and of benefit. New clinical trial designs and end points are necessary to permit more efficient evaluation of putative cancer treatments. This editorial will review commonly used clinical trial end points and describe their potential advantages and disadvantages to expedite the drug approval process required in the United States.

Introduction

The sequencing of the human genome and the elucidation of many molecular pathways important in cancer cell proliferation, apoptosis, and metastasis have provided unprecedented opportunities for development of new agents to prevent and treat cancer. Aberrant genes and pathways, once validated, become drug targets, as well as potential markers for cancer diagnosis, prognosis, and response to therapy. Thus, development of new pharmaceuticals is increasingly linked to development of molecular diagnostics that are necessary to identify those patients most likely to respond to a targeted agent. The types of molecules in development are also increasingly varied and include small molecules, monoclonal antibodies, antisense oligonucleotides, and ribozymes. Contemporary drug delivery routes and systems include not only traditional oral and i.v. routes but direct intratumoral injection, regional infusion, transdermal delivery, liposome-encapsulated therapy, and tumor-activated prodrugs, among others. Thus, the variety of anticancer agents in clinical development is now greater than ever before, and the number of agents currently in clinical trial for various cancer indications is estimated to exceed 400.

The requirement that new agents be demonstrated to be safe and effective before marketing approval is granted has led to a traditional paradigm for drug development that typically requires thousands of patients, hundreds of millions of dollars, and ≥1 decades to fulfill. Unless this paradigm is changed, there will not be enough patients, money, or time available to evaluate all of the promising new anticancer agents in development. New clinical trial designs and end points are necessary to permit more efficient evaluation of putative cancer treatments so that the most promising agents can move forward quickly, while disappointing agents are rapidly identified and discarded. This editorial will review commonly used clinical trial end points and describe their potential advantages and disadvantages to expedite the drug approval process required in the United States.

Basis for New Drug Approval

United States statutes require that drugs be demonstrated to be effective with an acceptable safety profile in adequate and well-controlled clinical studies as the basis for marketing approval. Studies must also provide sufficient information to define an appropriate population for treatment with the drug and describe the safety profile and intended use of the agent. With respect to oncology drugs, safety usually implies a risk/benefit assessment, and the acceptable ratio might vary for different diseases, patient populations, or stages of disease. A higher degree of risk might be acceptable for treatment of a solid tumor for which there is little known effective therapy. An agent considered to have an acceptable safety profile for treating metastatic cancer might not be considered to be safe in the adjuvant setting or as a prevention agent where large numbers of patients might be exposed to the treatment, but only a small fraction might be expected to obtain benefit. An analysis of safety should also include both acute and chronic toxicities of treatment. A new agent with a low risk of acute, reversible toxicity might not be considered safe if a significant proportion

Received 11/12/01; accepted 1/24/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This manuscript is based on a symposium entitled "Domestic guidelines for new drug approval and their impact on clinical trial design" (Richard L. Schilsky, Symposium Chair) presented at the Annual Meeting of the American Society of Clinical Oncology, May 2001, San Francisco, California.

[2] To whom requests for reprints should be addressed, at Biological Sciences Division, University of Chicago, 5841 South Maryland Avenue MC1000, Chicago, IL 60637. Phone: (773) 834-3914; Fax: (773) 834-3915; E-mail: rschilsk@medicine.bsd.uchicago.edu.

of patients develop late onset, chronic, or irreversible toxicity. This consideration might be particularly important in the evaluation of cytostatic therapies that patients are expected to use for long periods of time. In any analysis of safety data, attempts should also be made to determine whether any particular patient population is at high risk for developing a toxic effect. This might be particularly likely to occur in the setting of a genetic polymorphism of a drug-metabolizing enzyme where patients who are deficient in enzyme activity might be at very high risk for developing a toxic effect that is considerably less prevalent in the population at large.

Demonstrating the effectiveness of a new agent usually requires a demonstration of clinical benefit in a defined patient population. The term "clinical benefit" can be interpreted in a number of ways but is commonly accepted to mean that the agent demonstrates an improvement in survival compared with no therapy or to a known effective therapy, equivalence or noninferiority to a known effective treatment, or, in some cases, a clear improvement in time to disease progression together with an improvement in symptoms or QoL[3] for the patient. Successful demonstration of clinical benefit and an acceptable safety profile are usually sufficient to obtain approval to market the agent for a specific indication.

Interpretation of noninferiority trials is particularly complex. In most circumstances, a study designed to demonstrate superiority will not have sufficient statistical power to demonstrate equivalence. Therefore, the failure of a new drug to demonstrate superiority over a standard treatment does not imply that it has equivalent effectiveness; it could be worse. The design and interpretation of noninferiority trials first require an estimate of the benefit associated with the control arm. Such information is typically derived from data published previously rather than prospectively during the pivotal trial. It is important, therefore, to consider the consistency and reproducibility of the effect of the control treatment compared with a prior standard or placebo treatment. The proportion of the control effect that must be maintained to declare noninferiority will vary depending on the proposed clinical indication, although preservation of $\geq$80% of the control effect is a common goal of noninferiority trials. Large numbers of patients are usually required to estimate such an effect with sufficient statistical precision, and even then, the possibility remains that the new therapy could be as much as 20% inferior to the control treatment. Whether such a result is clinically acceptable will depend on the patient population being studied and the overall goals of treatment.

Clinical Trial End Points

Survival. Improvement in survival is generally considered to be the gold standard for drug approval, although improvement in disease-free survival may also be acceptable when the proposed indication is use in the adjuvant setting. Survival is an unambiguous end point that is not subject to investigator bias or interpretation. It is an end point that can be assessed easily, frequently, and without reliance on tumor measurements of any kind. In many ways, survival also provides a clear risk:benefit assessment of a new therapy. Therapies with a high treatment-related mortality might fail to show a survival benefit even if tumor control is substantially better with the new treatment. The major limitations to use of survival as a clinical trial end point are the period of observation required to reach the end point and its potential to be confounded by use of effective second-line therapies. Studies that permit crossover from the control arm to the experimental arm at the time of disease progression have the greatest potential to obscure a survival benefit conferred by a new agent. This concern seems to be more theoretical than real as effective agents applied in second-line therapy are rarely as effective as when used as initial treatment. Nevertheless, the increasing availability of agents with marketing approval in the second-line or refractory setting has focused attention on alternative trial end points that are less likely to be confounded.

TTP. Increasingly, TTP of disease has been proposed as an acceptable end point for cancer clinical trials. Like survival, TTP includes all patients in the primary efficacy analysis and has the advantage of achieving the trial end point sooner. TTP is not confounded by crossover effects or use of second-line therapies. However, TTP requires a great deal more precision in its definition than other clinical trial end points and is subject to confounding by investigator bias, particularly in unblinded clinical trials. Assessment of TTP is, at best, always an estimate because the actual disease progression always occurs sometime between scheduled observations. Thus, TTP can vary simply based on the frequency of patient evaluation with longer TTP resulting from less frequent assessments. It is critically important, therefore, that the schedule of disease assessments is the same on both arms of a randomized trial. Precise evaluation of TTP also requires that all sites of possible disease be assessed at baseline and at each follow-up assessment, ideally with the same assessment tools and technology. For this reason, TTP, unlike survival, is perhaps the most expensive clinical trial end point. Difficulty is also introduced in trying to define what is meant by "progression." Death attributable to cancer, appearance of new lesions, and unequivocal increase in size of existing lesions are commonly accepted as representing progression of disease. What about an increase in metabolic activity on positron emission tomography scan, an increase in plasma level of a tumor marker, a decline in performance status, or an increase in tumor-related symptoms (1)? These are not often included in the definition of "progression" but might clearly influence physician decision making about whether to continue a patient on protocol-specified therapy. Use of TTP virtually always requires a randomized clinical trial design to control for the rate of tumor progression that would occur in the absence of a treatment effect (2). Yet unblinded clinical trials, the norm in oncology drug development, are subject to ascertainment bias when a change in the clinical status of the patient prompts an unscheduled assessment of disease status. Finally, there is little agreement on how much improvement in TTP is necessary to constitute benefit to a patient. Because most patients in oncology clinical trials are asymptomatic, or only minimally so, it is not clear that a 2–3 month improvement in TTP, accompanied by physician visits and treatment toxicity, provides true clinical benefit if it is not ultimately associated with an improvement in survival.

[3] The abbreviations used are: QoL, quality of life; TTP, time to progression; PBMC, peripheral blood mononuclear cell.

Response Rate. Spontaneous regression of cancer is a rare event, so it is reasonable to assume that tumor regression after treatment is attributable entirely to a treatment effect. For this reason, response rate is generally considered clear evidence of antitumor activity and a surrogate for clinical benefit. Response rate has the advantage of being an early clinical trial end point, generally reached within 2–3 months of initiating treatment. Response criteria must be defined prospectively and applied consistently, and validation of response by an independent review committee that is blinded to treatment assignment adds credibility to the study results (3). Assessment of response duration is at least as important as measurement of response rate. Short-lived responses are unlikely to be clinically meaningful, particularly in asymptomatic patients. As for TTP, response duration can only be estimated and will vary based simply on the frequency of patient assessment. The proportion of complete and partial responses in a study may give some insight into the mechanism of action of the agent or the biological characteristics of the tumor. Thus, tumors that express or rely to a great extent on the molecular pathway targeted by the therapy might be more likely to regress completely than tumors that fail to express the molecular target. When the target is known with certainty and can be reliably measured, it becomes possible to screen patients in advance of trial enrollment so as to enrich the study population with patients whose tumors are most likely to respond.

For patients with serious or life-threatening diseases for which no effective therapies exist, drugs can be approved for marketing based on a surrogate end point that is reasonably likely to predict that clinical benefit will be demonstrated in prospective, randomized clinical trials. For oncology drugs, the response rate has generally been accepted as such a surrogate and can be used to support "accelerated approval" of a new agent. In such circumstances, the drug sponsor is required to undertake definitive studies designed to demonstrate clinical benefit, and approval for marketing can be withdrawn if such studies fail to do so. Oncology drugs recently approved under the accelerated approval mechanism include docetaxel, irinotecan, capecitabine, gemtuzumab, and imatinib mesylate.

QoL Assessment. Assessment of patient QoL is a potentially powerful clinical trial end point, because it provides information from the patient's perspective about the clinical benefits of treatment. QoL presumably represents a global assessment that incorporates the patient's experience of disease-related symptoms, treatment-related toxicity, impact of treatment, and disease on lifestyle and sense of general well being. However, QoL is one of the most difficult clinical trial end points to use because it relies almost entirely on subjective reporting by the patient that is difficult to verify independently. Generally, QoL studies should be placebo controlled and have a double-blind design in an attempt to minimize bias. They should use appropriate and validated instruments, and protocols should clearly specify how, when, and by whom the instruments are administered. Ideally, QoL studies should be hypothesis driven and should have sufficient statistical power, independent of the primary efficacy end points of the study, to demonstrate unambiguous results. Analysis of QoL studies is further complicated by missing data points and multiple end points and comparisons that make it likely that, by chance alone, improvements in QoL will be detected for some end points. Whenever possible, it is also desirable to specify prospectively the magnitude of change in an end point that is considered clinically meaningful. Despite these difficulties, QoL analyses can provide valuable perspectives that complement information obtained from traditional clinical trial end points and contribute to the overall assessment of a new agent or therapeutic plan.

Biological End Points and Surrogate Markers

As agents are developed with the goal of inhibiting specific cellular and molecular targets, it is appropriate to consider the use of biological end points and surrogate markers in clinical trials that reflect the proposed mechanism of action of the agent being studied. Doing so, however, requires knowledge of the target to be measured, a specific and reproducible assay for target inhibition, knowledge of the distribution of the target in the tissues of interest, and accessibility of the appropriate tissue or demonstration that the tissue in which the effect is assayed is a valid surrogate for the tissue of interest. When these conditions are met, it might be possible to more quickly complete the early phases of clinical testing of new agents. Thus, Phase I studies might aim to define the dose or concentration of a drug that provides maximal or sufficient target inhibition rather than maximally tolerated toxicity (4). Such an approach might be necessary when the agent under study produces few toxic effects. Demonstration of efficacy in Phase II testing is most likely to occur when eligibility is restricted to patients whose tumors express the target of interest. If the target is expressed in only a small percentage of tumors, then large numbers of patients will need to be screened before sufficient numbers of patients can be enrolled on study. However, enriching the patient population in this way increases the chances of detecting an antitumor effect and offers a therapeutic option to those patients most likely to benefit from it.

For some new agents, it will not be possible to directly measure the target of interest in the tissue of interest. In such circumstances, it is reasonable to consider use of a surrogate marker or direct measurement of the target of interest in a surrogate tissue. In either case, there should be a clear link between the surrogate marker or tissue and the primary biological effect of the agent. It is common, *e.g.*, to assay biological effects in PBMCs rather than in tumor tissue because of the ease of accessibility of PBMCs. However, validation of PBMCs as an appropriate surrogate requires clear demonstration that the target of interest is expressed in both PBMCs and tumor tissue at similar levels and that target inhibition in PBMCs is accompanied by target inhibition of a similar magnitude and duration in tumor cells. Validation of a surrogate marker (*e.g.*, measuring an event downstream of target inhibition) requires the demonstration that the biological event being measured is directly related to the putative mechanism of action of the agent in inhibiting tumor growth or progression. Failure to establish such a link will lead to many erroneous conclusions during the clinical evaluation of the new agent. Finally, it is necessary to distinguish between surrogate markers of effect and surrogate markers of benefit. The former may provide reliable evidence that the desired biological effect has occurred but provide no assurance that such an occurrence will result in clinical benefit to patients. With the exception of response rate, surrogate mark-

ers for clinical benefit, such as change in a tumor marker over time, have not yet been accepted by regulatory agencies in the United States.

Conclusions

Rapid development of new anticancer drugs that are safe and effective is a goal shared by clinical investigators, physicians, patients, industry, and regulatory agencies. Achieving that goal is increasingly complicated as new agents with novel mechanisms of action continue to enter clinical trials. Many of these drugs would be expected to work in only narrowly defined patient populations that must be prospectively identified. Thus, the development of the therapeutic agent must often be linked to development of a molecular diagnostic product. Drugs that produce primarily cytostatic effects might not be expected to produce regression of tumor masses; thus, evaluation of such agents would best be done in populations of patients with low-tumor burdens but high risk of disease progression (5). Such populations are difficult to define and recruit for clinical trials. As traditional clinical end points prove more difficult to apply in evaluation of molecularly targeted therapies, a great need exists to define and validate surrogate markers of effect and benefit. The former are best defined using appropriate preclinical model systems, whereas the latter can only be defined by well-designed clinical trials that link the proposed surrogate measure to a traditional measure of clinical benefit. Surrogate markers are likely to be specific to both the agent being studied and the tumor type being treated and are, therefore, optimally developed in parallel with the clinical evaluation of new agents. The force driving the development of new therapies for cancer is the continuing need to improve outcomes for patients with virtually all kinds of cancer, whether through improvement in survival, reduction in symptoms, or amelioration of acute and chronic toxicities of treatment. Increased patient accrual to clinical trials is necessary to achieve each of these goals, and the burden is on all involved in cancer drug development to work toward broadening knowledge about and increasing access to clinical trials around the world.

References

1. Eisenhauer, E. A. Phase I and II trials of novel anti-cancer agents: endpoints, efficacy and existentialism. Ann. Oncol., *9:* 1047–1052, 1998.

2. Korn, E. L., Arbuck, S. G., Pluda, J. M., Simon, R., Kaplan, R. S., and Christian, M. C. Clinical trial designs for cytostatic agents: are new approaches needed? J. Clin. Oncol., *19:* 265–272, 2001.

3. Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., *et al.* New guidelines to evaluate the response to treatment in solid tumors. J. Natl. Cancer Inst. (Bethesda), *92:* 205–216, 2000.

4. Gelmon, K. A., Eisenhauer, E. A., Harris, A. L., Ratain, M. J., and Workman, P. Anticancer agents targeting signaling molecules and cancer cell environment: challenges for drug development? J. Natl. Cancer Inst. (Bethesda), *91:* 1281–1287, 1999.

5. Von Hoff, D. D. There are no bad anticancer agents, only bad clinical trial designs. Twenty-first Richard and Hinda Rosenthal Foundation Award Lecture. Clin. Cancer Res., *4:* 1079–1086, 1998.

Clinical Cancer Research

End Points in Cancer Clinical Trials and the Drug Approval Process

Richard L. Schilsky

Clin Cancer Res 2002;8:935-938.