Clinical Development of Anticancer Agents—A National Cancer Institute Perspective

Silvia Marsoni* and Robert Wittes[1]

Since the first report that a chemical could cause significant tumor shrinkage (1), the development of clinically useful antineoplastic drugs has grown from the preoccupation of a few investigators to a major international effort. Over the past 35 years or so, about 30 drugs have been defined as active in one or more tumor types. When used alone in patients with disseminated malignancy, these drugs cause reduction in bulk of measurable neoplasm in a significant percent of cases; for most tumor types, however, ample evidence of residual cancer usually persists and after a few months, tumor regrowth occurs. More striking successes have been achieved with combinations of drugs; as is well known, for several kinds of disseminated human cancers, a high frequency of clinical complete remissions, with substantial long-term disease-free survival rates, is now possible (2–5). For other cancers which may not be curable by chemotherapy once they have disseminated, combinations of drugs appear to result in a higher total remission rate and a greater prolongation of life than single drugs (6,7). Perhaps more significant in the long run is the apparent effect of chemotherapy when it is used as part of a planned multimodality effort (8,9).

By some perverse quirk of fate, chemotherapy seems to chiefly exert a major impact in rare tumors, while the common epithelial neoplasms of adulthood have thus far resisted satisfactory solutions. Therefore, the central problem of drug development, the identification of effective agents with reasonable therapeutic index, is as pertinent for oncology now as at any time in the past.

The idealized outlines of the successive steps in drug development are familiar to all oncologists. In phase I trials, we define a dose suitable for use in studies of the drug's activity across a spectrum of human tumors. Increasing awareness of the importance of patient- and disease-related parameters has effectively led to the replacement of the broad phase II trial with a series of disease-oriented activity studies. Having defined settings in which the new drug is active, investigators then proceed to compare the new treatment with standard therapy (phase III) and to further explore the drug's therapeutic potential in other ways, such as in combination with other agents or by alternate routes of administration.

Anyone familiar with the actual workings of this process over the past two decades knows that despite its successes, it has not functioned as systematically or efficiently as the above description might imply. In addition, many of the assumptions on which the process was based are in need of re-examination. Since there are no reliable laboratory predictors of efficacy for specific human cancers, drug development will continue to require extensive testing in human subjects, an endeavor that is never without ethical dilemmas, however thoughtfully it is carried out. In addition, because human cancers vary widely in sensitivity to anticancer drugs, the apparatus required to sustain the clinical trials effort is necessarily large and very expensive. For these reasons alone, another look at the drug development program of the National Cancer Institute (NCI) seems to be worthwhile.

Phase I

Phase I trials of antineoplastic compounds are conducted in patients with disseminated malignancies for whom standard treatment either does not exist or has proved ineffective. Drugs are given in a phase I trial with therapeutic intent; the main scientific goal is to define the qualitative and quantitative characteristics of the drug's acute toxicity, and in so doing, to determine a biologically active dose which is tolerable for every patient. The maximum tolerated dose (MTD) is usually higher in children than in adults (10), probably because of better organ function and possibly because of different pharmacokinetics (11,12). Also, since the MTD for patients with acute leukemia is often substantially higher than that for solid tumors, at least four

[1]Cancer Therapy Evaluation Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD.

*Reprint requests to: Silvia Marsoni, MD, Drug Evaluation and Reporting Section, Investigational Drug Branch, Cancer Therapy Evaluation Program, National Cancer Institute, Landow Bldg, Rm 4C09, 7910 Woodmont Ave, National Institutes of Health, Bethesda, MD 20814.

phase I trials should be conducted for each drug. In practice, of course, three to four phase I trials testing different schedules of the drug are usually performed in adult patients with solid tumor alone, and trials in children do not start until substantial experience is accumulated from the adult trials.

The assumption underlying all phase I escalation procedures is that anticancer compounds must be given at or near the MTD; therefore, the job of a phase I trial is to define the highest dose that can be safely delivered to a patient, since this dose will also be the one that has the best chance of being active. This approach reflects the difficulties in establishing a clear-cut measurable endpoint for drug activity against cancer. Since induction of response is not usually an event that is immediately recognizable, attainment of toxicity is the only assurance that, if a response is not obtained, at least a biologically active dose was delivered. Underlying this assumption is the more fundamental one that the dose-response curve for human cancers is monotonically increasing throughout the range of tolerated doses and is to the right of the dose-toxicity curve. Needless to say, the details of this assumption have not been generally verified for antitumor agents, chiefly because rigorously defining the shape of a clinical dose-response curve is a laborious task, requiring a large number of patients treated at each of several dose levels. Where the relationship between dose and response has been examined, however, most of the data are at least consistent with the conclusion that the higher the administered dose, the more probable an antitumor effect (13-17) or the longer the duration of remission (18). On the other hand, recent trials in small cell lung cancer suggest that the probability of response does not continue to increase linearly as the dose approaches the MTD (19).

For most of the clinically useful compounds, the bone marrow is dose-limiting. The dose-toxicity curve for myelosuppression is quite reproducible, and the status of marrow reserve is the major source of interpatient variability. Generally, treatment of six to ten patients at or near the MTD is sufficient to establish a safe phase II dose when myelosuppression is the dose-limiting toxicity.

However, major problems may arise when other toxic effects which are less easily quantifiable are dose-limiting. For example, in a phase I study of escalating doses of carmustine with autologous bone marrow support (20), major organ toxicity (liver, central nervous system, and lung) surfaced abruptly at a dose of 1500 mg/m^2. Because of the sudden appearance of these side effects in the escalation scheme and the long interval from the beginning of treatment to onset of toxicity (6-9 weeks), the overall mortality rate for patients entered at a dose of $\geq$ 1500 mg/m^2 was approximately 30%.

Experience suggests that whenever a drug has dose-limiting side effects other than myelosuppression, its transition into phase II has often been compromised. An analysis of 31 drugs entered in phase I evaluation by the NCI shows that whenever the drug had myelosuppression alone as the dose-limiting toxicity, it had a high probability of undergoing full phase II study; however, when other organ toxicity was dose-limiting, only 25% of the drugs proceeded to full phase II study (table 1). Evaluation of the remainder of drugs was restricted by either the NCI or lack of investigator interest. The main reason for these difficulties relates largely to the uncertainty regarding reversibility of acute major organ damage. In addition, even if organ damage should turn out to be reversible, medical support during periods of severe organ failure is either extremely intensive and costly (kidney, CNS) or technically unsatisfactory (liver), and is not seen as feasible or justifiable by most investigators in the context of a clinical experiment.

TABLE 1.—Phase II evaluation as a function of the dose-limiting toxicity of 31 cytotoxic compounds (1975-1982)

Phase II evaluation	Dose-limiting toxicity		
	Myelosuppression *	Myelosuppression and organ toxicity†	Organ toxicity‡
Full	10	3	3
Restricted	0	3	4
Dropped			
Toxicity	0	0	4
No interest	1	1	1
No drug supply	1	0	0
Total	12	7	12

* Bisantrene, diaziquone, aclarubicin (aclacinomycin), mitoxantrone, PCNU, amsacrine, zorubicin, chlorozotocin, carboplatin, ICRF 187, 5-methyltetrahydrohomofolate, and 3-deazauridine.

†Acivicin, maytansine, anguidine, teroxirone, dihydro-5-azacitidine, indicine *N*-oxide, and DON.

‡Pyrazofurin, L-alanosine, homoharringtonine, TCN-P, *N*-methylformamide, pentostatin, hycanthone, dichloroallyl lawsone, pyrazole, aminothiadiazole, bruceantin, and spirogermanium.

This dilemma appears to have no easy solutions. One alternative might be to utilize data from pharmacologic studies to define the relationship between dose, plasma level, tissue level, and clinical activity. For example, in the case of pentostatin, knowledge concerning the amount of drug needed to abolish activity of the target enzyme adenosine deaminase has helped to establish a phase II dose independent of the attainment of clinical toxicity. Unfortunately, however, this is an exceptional situation and, under most circumstances, specific intracellular targets of drug action either have not been identified or are not so susceptible to analysis.

A somewhat more empiric approach is exemplified by the current plans for developing *N*-methylformamide, a drug which had been introduced into the clinic in the 1950s and was subsequently dropped while in phase I because of hepatotoxicity (21). Interest in the drug has recently been revived because of its activity against human tumor xenografts (22) and its capacity to induce differentiation in vitro (23). Phase I trials in both the United Kingdom and the US confirm that, at a dose of 1000 mg/m^2, the reversible hepatotoxicity of *N*-methylformamide is dose-limiting and myelosuppression is completely absent.[2,3] This dose has been defined as the MTD, at which phase II trials have just begun. If activity is observed in any tumor type, a repeat phase II study in one or more susceptible tumors will be performed at a level immediately below the MTD. This procedure will define whether the attainment of toxicity is necessary for activity.

Phase II

In a phase II trial, the main goal is to assess the activity of the drug in a variety of disseminated malignancies and to further define the patterns of acute toxicity in patients who are homogeneous in diagnosis and in better general medical condition than patients in phase I. Since large numbers of patients are treated during phase II, rarer acute toxic effects often surface for the first time (24). Also, since cumulative drug doses may be appreciable in responding or stable patients, phase II provides an appropriate setting for initial assessment of chronic toxic effects.

Several vexing problems are inherent in this process. In the first place, since patient numbers and resources are finite, it is impossible to explore the activity of each drug in each tumor type. A method must be found to focus the effort of drug development in a way that will minimize the chance of overlooking active agents. Accordingly, the NCI decided to evaluate all experimental drugs in selected types of cancer. The NCI Human Tumor Panel was created in 1975 and included lung, colon, and breast carcinomas, and lymphoma, leukemia, and melanoma. The original intention was to match tumors in the human panel with those in the preclinical panel, thereby providing information for the validation of the preclinical screening program. In addition, these classes of human cancer represent the two extremes of chemotherapy sensitivity and might be expected to exhibit both high sensitivity and high selectivity. Finally, the inclusion of the most common causes of cancer deaths (breast, colon, and lung cancers) permits the study of large numbers of patients and assures that results will have immediate implications for the treatment of prevalent cancers. Needless to say, evaluation of individual drugs is also carried out in tumors other than those in the panel, particularly if there is a specific reason to do so. For example, diaziquone was chemically designed to cross the blood-brain barrier and therefore has been extensively evaluated in brain tumors with encouraging results.

How has activity in the prelinical panel correlated with clinical activity? Thus far, we have analyzed the results with 13 experimental drugs for which clinical and experimental data are available. The correlation of activity in each model tumor system with activity in the corresponding human cancer is shown in figure 1. Prediction of true-negative results (resistance) seems fairly reliable across most of the rodent and xenograft systems. On the other hand, the probability of predicting true-positive results (sensitivity) is very low. Because of the small number of active drugs in humans for which complete data are available, no definitive conclusions can be drawn. However, even if the preclinical panel should not turn out to be an accurate predictor of response in individual tumor types, overall activity in prelinical screening may still serve as a general predictor of activity in at least one human cancer. The aggregated data are, in fact, consistent with this notion. This has obviously been the general premise on which antitumor screening programs have operated for years. Its validity has been widely assumed but has not been subjected to direct test, since drugs are not brought to the clinic if screening data are not positive. An assessment of the validity of the assumption will be afforded by the use of the human tumor stem cell assay as a screening tool. The plans are to bring selected compounds which are positive in the human tumor stem cell assay to clinical trial, even if they are negative in the P388 prescreen (26). Obviously, more data are needed to determine the ultimate usefulness of the panel.

How has the human panel fared as a predictor of clinical efficacy in human tumors other than those of the panel? Since 1971, 62 cytotoxic agents have been introduced into clinical trials under the sponsorship of

[2]McVie JG, ten Bokkel Huinink WW, Simonetti G, et al. Phase I trial of *N*-methylformamide (NSC 3051) (NMF). Manuscript submitted to *Cancer Treatment Reports*.

[3]Minutes of the Phase I Working Group Meeting, NCI, Bethesda, MD, July 1983.

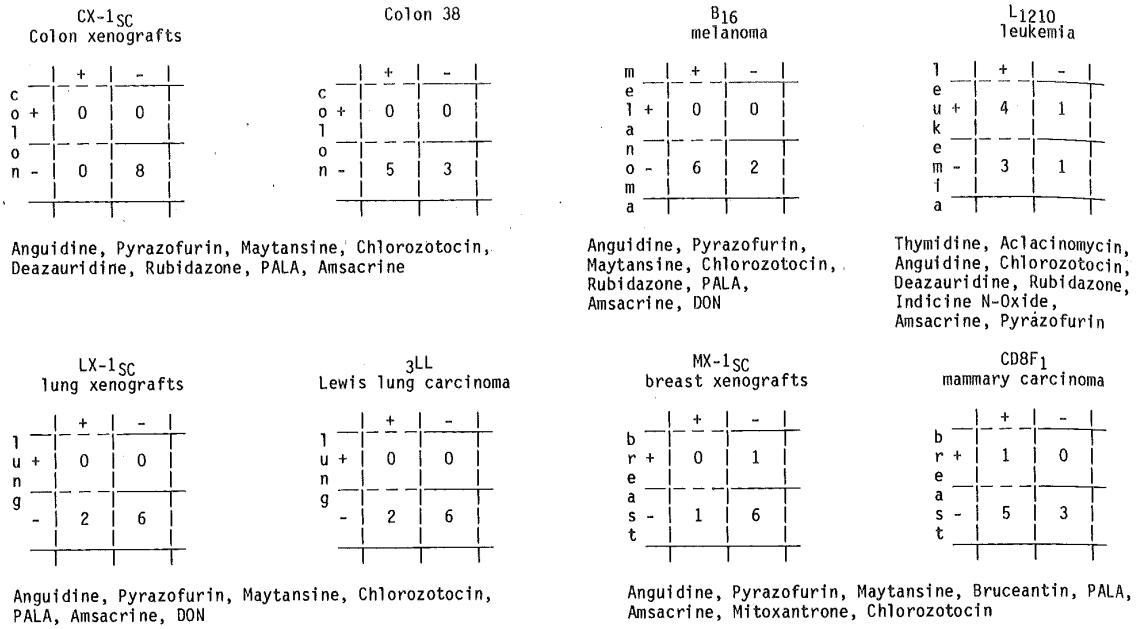

FIGURE 1.—Correlation of activity of 10 antitumor agents in murine model tumor systems with activity in human cancer. Activity in murine tumors was judged according to NCI Decision Network 2 criteria (Goldin A, et al. Eur J Cancer 17:129-142, 1981). Activity in human tumors is defined as a 20% response rate in at least 1 clinical trial with ≥ 14 evaluable patients.

NCI. Results of an interim analysis of phase II results are available for 13 drugs, which were studied in 180 protocols (table 2). Although these data represent only a fraction of the large NCI experience, certain trends are apparent. First, significant activity of ≥ 20% was seen only in the lymphomas, leukemia, and breast carcinoma (50%, 29%, and 14% of the studies, respectively); most of the results were in the 20%-30% range. No drug showed > 20% activity in colon carcinoma and melanoma, and only 6% of the lung cancer trials showed positive activity. Even with only those drugs which have shown activity in at least one tumor type, the overall response rate in colon and lung cancers and melanomas is still consistently < 10%.

It is well known that colon cancer and melanoma are highly resistant diseases, and that these diseases which are consistently refractory to all therapies are of no value in screening. Although more data are needed, the results to date suggest that inclusion of colon cancer and melanoma in the panel may not be useful for screening. However, it would seem reasonable to continue testing new agents for activity in those common and refractory neoplasms until truly reliable screens for activity have been defined. Such screens may emerge from further analysis of data for clinical trials or from advances in the use of in vitro or in vivo laboratory methods.

Second, even in intrinsically sensitive diseases like

TABLE 2.—Outcome of phase II studies in human cancer

Disease	Total No. of studies	Response rate		
		0%	< 20%	≥ 20%
Breast cancer	30	13 (43%)	13 (43%)	4 (14%)
Colon cancer	38	25 (66%)	13 (34%)	0
Leukemia	21	6 (29%)	9 (48%)	6 (29%)
Lung cancer	47	27 (57%)	17 (36%)	3 (6%)
Lymphoma	18	4 (22%)	5 (28%)	9 (50%)
Melanoma	26	16 (62%)	10 (38%)	0
Total	180	91 (51%)	67 (37%)	22 (12%)

breast or small cell lung cancer, the number of drugs showing activity turned out to be very small. As seen in table 3, of the 11 drugs considered in breast cancer, only bisantrene showed an overall response rate of > 20%. As has been true throughout the history of medical oncology (27), the estimates of activity vary widely from trial to trial. For example, with mitoxantrone, response rates ranged from 5% to 28%. This observation suggests the well-known importance of factors other than drug dose and schedule as major influences on estimates of response rate. Again, for mitoxantrone, the deleterious effect of prior therapy on response seems to be fairly clear (table 4).

In fact, the success of the human tumor panel in predicting (or ruling out) general patterns of efficacy for other human cancers will depend to a large extent on what kinds of patients with "panel cancers" are chosen for entry in the study. A negative trial of a new drug in 20 patients with breast cancer who have failed multiple prior regimens tells us nothing about either the potential of this drug in a more favorable breast cancer population or drug activity in other tumors. Moreover, data from earlier eras of cancer chemotherapy cannot be used reliably to decide which tumors may be usefully included in a panel without extensive consideration of how shifting patterns of practice may have altered important patient characteristics.

The phase II effort also needs certain administrative refinements. Table 5 shows a breakdown by disease of patient accrual patterns for negative phase II studies, ie, trials yielding a < 10% response rate. Even allowing for the histologic heterogeneity of certain primary sites such as lung, the extent of overaccrual in some of these categories suggests the need for much earlier review of the data by investigators and a tighter system of control by the statistical offices of cooperative groups. Indeed, several groups have already implemented procedures which should minimize the chances that patients will be entered in treatments already shown to be inactive.

TABLE 3.—Activity of 11 NCI drugs in patients with breast cancer (1975-1980)

Drug	No. of responding patients/ total evaluable	Response rate (%)
Aclarubicin	1/48	2
Amsacrine	12/173	7
Anguidine	1/37	3
Acivicin	0/15	0
Bisantrene	13/50	26
Bruceantin	0/15	0
Diaziquone	2/63	3
Mitoguazone	4/104	4
Mitoxantrone	16/182	9
PCNU	0/45	0
Piperazinedione	3/47	6

TABLE 4.—Responses to mitoxantrone in carcinoma of the breast trials according to previous treatment

Response rate (%)	No. of previous regimens	Institution *
3	3	SWOG
10	3	SECSG
5	2.6	ECOG
22	3	M. D. Anderson Hospital and Tumor Institute
21	1	Ohio State University
19	1	EORTC
28	0	The Royal Marsden Hospital

* SWOG = Southwest Oncology Group; SECSG = Southeastern Cancer Study Group; ECOG = Eastern Cooperative Oncology Group; and EORTC = European Organization for Research on Treatment of Cancer.

Phase III

Once the activity of a compound is established in one or more diseases, subsequent development of the drug proceeds along two separate lines. One of these lines is to establish the role of the drug in the disease for which activity was demonstrated. The endpoints of such studies, which are designed to compare the drug alone or in combination against standard treatment in a randomized fashion, are not only relative activity (eg, response rate), but also response duration, survival, and toxicity; the ultimate goal is to define the specific contribution of the drug in the treatment of a particular cancer. The data from such trials may be used by pharmaceutic firms seeking New Drug Application (NDA) approval from the Food and Drug Administration (FDA) for marketing purposes.

In this connection, the intense interest in chemical analogs of existing active agents poses special challenges to clinical drug development. Of 31 drugs developed by NCI since 1975, eight have been analogs of commercially available or experimental drugs. Until very recently, the development of analogs proceeded along essentially the same lines as that of novel structures. Formal prospective comparisons of analog versus parent were rarely carried out (28). As a result, little direct comparative data exist on the relative merits of the various bifunctional alkylating agents, nitrosoureas, anthracyclines, or epipodophyllotoxins.

Surely, if parent and analog have borderline activity in a certain cancer, such direct comparisons are probably not worth undertaking. Moreover, when such comparisons are worth doing, the trials need to be quite

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.