EDITORIALS

Antineoplastic Drug Screening Belongs in the Laboratory, Not in the Clinic

Larry M. Weisenthal*

Drug-resistant strains of Mycobacterium tuberculosis have increasingly begun to emerge (1). Many authorities agree that this situation has the potential to become a problem of serious proportions. One hopes that scientists in both public and private sectors are hard at work studying these resistant strains in the laboratory to discover new drugs and methods to circumvent the resistance to the old drugs.

What if it were impossible to study these resistant strains in the laboratory? What if the only models for Mycobacterium tuberculosis were species of Escherichia coli? What if we didn’t have a clue whether or not a series of new drugs had any chance of working until we had tested them in thousands of dying human beings with advanced tuberculosis? Finally, how would we feel if most of the drugs which emerged from the Escherichia coli screen as potential antituberculosis candidates routinely inflicted life-threatening toxic effects in these dying human beings? One would think that the first order of business would be to make a high priority of developing better model systems. More precisely, one would look to develop laboratory methods to study directly the drug-resistant mycobacteria which are causing all of the trouble.

Of the leading causes of mortality in this country, cancer has arguably the most inadequate predictive model systems for preclinical drug development. Valid model systems are most critically needed at four stages in the drug development and cancer treatment process. First, a high-capacity screening system is required to select from the tens of thousands of interesting compounds which become available for consideration each year. Second, a more specific, second-stage screening system is needed to test the thousand or so compounds which either come out of the high-capacity screen or which are judged worthy of bypassing the high-capacity screen. Third, a disease-specific screen is needed to direct phase I and II trials of new drugs more rapidly into the specific forms of cancer in which the new agents are most likely to be effective. Fourth, flexible laboratory systems are needed to individualize drug selection for phase III and phase IV trials and also for noninvestigational cancer treatment in the community.

I doubt whether many would quarrel over the need for the first two of the four needs stated above, and many opinions and reviews are available from U.S. Government public records and from the literature (2-9). For many years, until the mid-1980s, preclinical screening systems within the National Cancer Institute (NCI) were largely based on a small panel of transplantable rodent tumors, supplemented in later years by a small panel of human tumor xenografts (3). This screen was abandoned in the mid-1980s, because it was not perceived to be producing drugs with solid tumor activity (6).

As a result of dissatisfactions with the murine tumor-based screen, the NCI switched in the mid-1980s to a screen based on the use of a human tumor cell line panel. This panel is composed of more than 60 different established cell lines, representing a number of major forms of human cancer and studied in vitro in nonclonogenic assay systems (8,10-13). As a member of an international advisory committee which provided intermittent reviews of this program, I was concerned that the known clinical patterns of activity of established agents were often not reflected by the activity of these agents in the cell line panels (14,15).

On the other hand, the new screen has clearly proven capable of processing large numbers of materials and has proven to be a "cell line specific," if it has not been entirely proven to be "disease-specific." The activity of a test agent in the screen characteristically produces a unique pattern when tested over the 60-cell line panel (8,11). This pattern can be used as a "fingerprint" to characterize the agent as, for example, being related to alkylating agents, topoisomerase II inhibitors, or dac tinomycin. Despite the fact that known activity patterns of standard drugs are not always reflected in the cell line panel, this panel has identified some interesting new compounds with some apparent selective activity in brain and melanoma lines (Shoemaker R: personal communication). Further work will be required to determine if these are real findings, which could be generalized to large numbers of additional cell lines from the same tumors, or if they represent a chance statistical grouping from the screening of thousands of compounds. As a high-capacity, first-stage screen with some novel and powerful capabilities, the NCI cell line panel is certainly worthy of further study (7), but it falls far short of the model systems available to investigators who develop drugs against Mycobacterium tuberculosis.

It is common sense to consider using in vitro cultures of fresh human tumors as potential tools for drug development. The NCI in 1979 initiated a contract research program which attempted to do just that. In that year, the NCI solicited investigator participation in a project to test the "human tumor colony assay" (HTCA) as a drug development tool. At the time, I was critical of the project (Weisenthal LM: personal communication to V. T. DeVita, Jr., et al.), because it had the goal of studying and using only a single in vitro assay technology, rather than making a

Received February 28, 1992; accepted March 2, 1992.

I thank Robert A. Nagourney and Robert H. Shoemaker for reviewing (though not necessarily endorsing) the manuscript. Any remaining, unrevised inadequacies are strictly the author’s responsibility.

*Correspondence to: Larry M. Weisenthal, M.D. Ph.D., Weisenthal-Dill Cancer Resource Group, 15201 Springdale St., Huntington Beach, CA 92649.
head-to-head comparison of different technologies, based on different principles and having different capabilities. Nonetheless, by 1986, some 1200 compounds had been tested, with a later emphasis on testing compounds that had been negative in the rodent panels (5,16,17).

Two drugs were selected by the HTCA screen for clinical trials in non–small-cell lung cancer—dihydroenolperoxide and chloroquineoxaline sulfonamide (5). Dihydroenolperoxide does not appear to be an active agent, possibly owing to the fact that drug concentrations tested in vitro could not be achieved clinically (18). Chloroquineoxaline sulfonamide may, on the other hand, turn out to be an active agent in the treatment of non–small-cell lung cancer (19). If these early indications hold up, one could not argue that the resources devoted to this program were entirely misspent, although an opportunity to learn about different assay end points was undeniably lost. Additionally, those NCI investigators with full access to the results of the project were apparently not convinced of the value of the HTCA for drug development, inasmuch as this assay has not been incorporated into either second-stage screening (as defined above) in the current NCI drug development program and contract initiatives to use the assay for directing clinical trials have not been advanced.

An unanswered question, 14 years and probably at least 30 000 assays after the first report of Salmon et al. (20) of clinical correlations with the HTCA, concerns how well the assay really does in reflecting the clinical activity spectra of antineoplastic drugs. This is a separate question from that of determining the clinical predictive accuracy of the assay in individual patients, which has been extensively addressed elsewhere (9,21-24). The assay certainly has been well documented to be one of a number of assays capable of identifying the presence of clinical drug resistance in tumors from individual patients (9,21-24). However, for many reasons, all available in vitro assays are much better at detecting drug resistance than they are at detecting drug sensitivity (24). The cell proliferation-based HTCA may be theoretically less capable of accurately identifying clinical patterns of activity for many drugs than are assays measuring drug effects in the entire (proliferating and nonproliferating) tumor cell population (24,25).

So, how about it? Can the assay tell us in what diseases new or old drugs are most likely to work? A review published in this Journal concluded that the case had not been proven, citing some notable disparities between previous predictions involving bisantrene and mithoantrone and the ultimate results of phase II trials with these agents (26). Among Sydney Salmon, Dan von Hoff, David Kern, the NCI HTCA contractor group, and a NCI publication (56) and added up results listed for the drug chloroquineoxaline sulfonamide may, on the other hand, turn out to be an active agent in the treatment of non–small-cell lung cancer (19). If these early indications hold up, one could not argue that the resources devoted to this program were entirely misspent, although an opportunity to learn about different assay end points was undeniably lost. Additionally, those NCI investigators with full access to the results of the project were apparently not convinced of the value of the HTCA for drug development, inasmuch as this assay has not been incorporated into either second-stage screening (as defined above) in the current NCI drug development program and contract initiatives to use the assay for directing clinical trials have not been advanced.

An unanswered question, 14 years and probably at least 30 000 assays after the first report of Salmon et al. (20) of clinical correlations with the HTCA, concerns how well the assay really does in reflecting the clinical activity spectra of antineoplastic drugs. This is a separate question from that of determining the clinical predictive accuracy of the assay in individual patients, which has been extensively addressed elsewhere (9,21-24). The assay certainly has been well documented to be one of a number of assays capable of identifying the presence of clinical drug resistance in tumors from individual patients (9,21-24). However, for many reasons, all available in vitro assays are much better at detecting drug resistance than they are at detecting drug sensitivity (24). The cell proliferation-based HTCA may be theoretically less capable of accurately identifying clinical patterns of activity for many drugs than are assays measuring drug effects in the entire (proliferating and nonproliferating) tumor cell population (24,25).

So, how about it? Can the assay tell us in what diseases new or old drugs are most likely to work? A review published in this Journal concluded that the case had not been proven, citing some notable disparities between previous predictions involving bisantrene and mithoantrone and the ultimate results of phase II trials with these agents (26). Among Sydney Salmon, Dan von Hoff, David Kern, the NCI HTCA contractor group, and a myriad of other investigators, tens of thousands of tumors have been tested and hundreds of papers have been published. What remains to be published is a paper in which a series of well-known drugs (cisplatin, doxorubicin, etoposide, melphalan, carmustine, and fluorouracil, etc.) are tested, under identical conditions in a series of well-known forms of human cancer (melanoma, lymphoma, leukemia, and ovarian, breast, small-cell lung, and non–small-cell lung cancers, etc.). Ideally, one would wish to see at least 20 different specimens tested per tumor type, and one would preferably also wish to see comparisons of specimens from previously treated versus untreated patients, wherever possible. The absence of these data from the literature makes it difficult to evaluate how well suited the assay would be for identifying the most promising diseases in which to carry out initial phase II trials of new drugs.

It is in the above context that the article by Taylor et al. (27) in this issue of the Journal provides some interesting information and makes some interesting predictions. The authors tested the investigational drug suramin in relatively large numbers (n = 17-115) of individual specimens in each of five different types of human cancers (melanoma and endometrial, kidney, non–small-cell lung, and ovarian cancers). The median extrapolated IC50 (i.e., suramin concentration reducing growth to 50% of control) in endometrial carcinoma was almost two logs lower than that in melanoma and that in kidney cancer. Endometrial carcinoma, ovarian cancer, and non–small-cell lung cancer all had median extrapolated IC50s significantly lower than those in melanoma and kidney cancer. These results may be taken as bold and clear prospective predictions of the clinical activity pattern of an interesting drug just beginning phase II trials. It will be important to revisit the present data after results of clinical trials are known. With this in mind, the authors might be encouraged to provide supplementary data concerning whether specimens from previously treated patients were more or less resistant to suramin than were specimens from previously untreated patients.

Additionally, Taylor et al. (27) compared the activity of suramin in the HTCA with suramin’s activity in a similar assay developed by Tanigawa et al. (28), in which thymidine labeling is substituted for colony counting as the assay end point. Taylor et al. (27) found that the thymidine assay detected lower cell killing at each tested drug concentration than did the colony assay. This fact, by itself, does not indicate which, if either, of the two assays is “right,” inasmuch as each assay must be calibrated and controlled differently and each assay is subject to its own unique artifacts (23,24,29). Although a number of previous studies (29-35) reported a reasonably good correlation between the two assay end points, the end points are not identical. The NCI program that evaluated the HTCA as a drug-screening tool used both end points at different times during the course of the program. Chloroquineoxaline sulfonamide was identified during the period when the colony end point was being used.

Do we need assays like these to direct phase II, III, and IV trials and noninvestigational cancer therapy? For the past 20 years, the ascendant clinical research paradigms have been the empiric phase II trial (meaning random drug screening in patients) and the prospective, randomized phase III and phase IV trials to identify the best treatment to administer to the average patient. Are these paradigms “broke,” and should they therefore be “fixed?” I recently contacted the NCI to ask for tabular data listing drugs studied in phase II trials, number of diseases studied, number of patients treated, number of responses, incidence of serious toxic effects, and treatment-related deaths. I was told that these data were not available. I therefore turned to an earlier NCI publication (36) and added up results listed for the drug amscarine. I calculated that 3544 patients were treated on phase II trials with this drug until 1985. The partial response rate was 2.5% for 2478 solid tumor patients and 9.9% for 1066 patients with hematologic neoplasms. The complete response rate was...
0.2% for patients with solid tumors and 10.5% for patients with hematologic neoplasms. In the case of small-cell lung cancer, there were no responses in 164 patients treated; in the case of pancreatic cancer, there were no responses in 114 patients treated. Amsacrine was introduced into clinical trials in 1977, and patients are still being accrued in the search for an indication. Most patients volunteer to participate in such trials because they are motivated by a desire for self-preservation and a desire to contribute to science and help others. Would as many patients volunteer for phase II trials if they really understood the likelihood that they would derive meaningful personal benefit or that they would contribute to a meaningful advance in medical knowledge? The value of the knowledge obtained from prospective, randomized phase III and phase IV trials during the last 20 years as a function of resources expended (including 20,000 patients per year (37)) is an interesting topic beyond the scope of this discussion (38,39). It can, however, be noted that the concept of trying to identify a single best drug regimen to administer to the average patient with a given form of cancer runs contrary to many important tenets of modern tumor biology.

Eleven years ago, I wrote (4) that "there are good reasons to think that in vitro screening tests using established cell lines would not be more accurate than whole animal screening tests ... as an alternative ... in vitro drug screening assays could be based on the use of fresh human tissue biopsies and primary cell cultures made from these biopsies ... a productive approach might well be a commitment to 'dig in' and devote more time and resources to a methodological comparison of different types of screening tests, even at the expense of slowing down the screening of new compounds for a period of time." Eleven years later, we still don't know if fresh-tumor assays can tell us anything important that cell lines can't and we still don't know which types of fresh-tumor assays can give us the best information.

In my view, there are a number of nonclonogenic fresh-tumor assay systems available for drug development and clinical practice which years ago were (and still are) worthy of support, funding, interassay comparisons, and integration into clinical trials (4,23,24,29). However, cell culture assay studies on fresh tumors have not been well supported by the peer-review and clinical trials systems (21). In the absence of alternatives, current clinical investigators are forced to screen drugs in humans in the clinic, rather than in the laboratory. Yet, it is hard to imagine that it will not one day be possible to study drugs in valid, revealing laboratory systems, rather than in patients. Much more effort should be devoted toward hastening that day. In other words, the study of systems for studying drugs is, in many ways, as important as the study of the drugs themselves. In this regard, the current study by Taylor et al. (27) represents a small but worthwhile step, and funding and publication of similar studies should be encouraged.

References

Genetic Heterogeneity in Breast Cancer

Sandra R. Wolman, Gloria H. Heppner*

A recurrent theme in recent cancer research is that genetic changes accompany, and presumably drive, the development of a neoplasm and its ensuing progression to an increasingly malignant phenotype. These changes fall broadly into two categories: those which add new potential to the behavioral repertoire of cancer cells and those which remove a regulatory influence, allowing cells to exhibit behavior that would otherwise be restrained. This latter group of changes is attributed to the loss of "suppressor genes," a designation which presupposes that cells are potentially cancerous but are somehow prevented from growing unrestrained. This latter group of changes is attributed to the loss of "suppressor genes," a designation which presupposes that cells are potentially cancerous but are somehow prevented from growing unrestrained. This latter group of changes is attributed to the loss of "suppressor genes," a designation which presupposes that cells are potentially cancerous but are somehow prevented from growing unrestrained.

Another recurrent theme in cancer research, both recent and not so recent, is the variability or heterogeneity of the neoplastic phenotype. Heterogeneity is evident among histologically similar cancers from different patients (inter-tumor heterogeneity) and among different cells of the same cancer at a single time (intratumor heterogeneity), as well as at different points in time (progression). The mechanisms underlying neoplastic heterogeneity are many; they include classical genetic (5), adaptive (6), and (micro-) environmentally determined events (5,7). Both themes - heterogeneity and genetic loss - appear in the report of Chen and co-workers (/), specifically in relation to LOH in breast cancer.

Chen et al. analyze the frequency and distribution of LOH at multiple alleles within a series of primary breast cancers and lymph node metastases. Their results should be considered in the context of a rapidly growing literature on the same topic. For example, although 98 tumors might suffice with highly informative probes, many of the probes cited here are expected to be informative in one third or fewer of the cases. Furthermore, in their brief report, it is not possible to discern, case by case, the association of LOH in primary cancers with clinical or pathological characteristics, with intratumor heterogeneity, or with LOH in metastases. It would be fascinating to know whether intratumor heterogeneity for LOH in a primary cancer is recapitulated in its metastases or, indeed, whether such heterogeneity occurs in metastases at all. What is the relationship between LOH in primary cancers and the incidence of metastases? Are there informative primary tumors without LOH that have metastasized? Is there an association between LOH and clinical stage? The limited clinical and pathologic descriptions presented compromise the biological significance of the molecular data.

The relevance of intratumor heterogeneity in terms of LOH should only be evaluated in relation to data on the percentage of cancer cells per individual tumor sample. This variable can influence the detection of LOH, particularly if LOH is present in only a minor subpopulation. "Simulation" experiments, in which cancer samples with LOH at different loci are deliberately contaminated with varying amounts and types of adjacent normal tissues, could perhaps aid in setting the technical limitations of the analyses. Additional insight might also be gained if LOH heterogeneity were associated with other evidence of intratumor heterogeneity - for example, of morphology, tumor grade, or estrogen receptor expression.

Received February 25, 1992; accepted February 27, 1992
S. R. Wolman, M.D., Department of Pathology, Wayne State University School of Medicine, Detroit, Mich.
G. H. Heppner, Ph.D., The Breast Cancer Program, Meyer L. Prentis Comprehensive Cancer Center of Metropolitan Detroit.
*Correspondence to: G. H. Heppner, Ph.D., Michigan Cancer Foundation, 110 East Warren Ave., Detroit, MI 48201.