Preclinical Toxicology of Anticancer Agents

Philip S. Schein
Division of Medical Oncology, Vincent T. Lombardi Cancer Research Center, Georgetown University School of Medicine, Washington, D.C. 20007

The use of animal toxicology as a potential predictive system for drug toxicity in man is particularly important to the field of cancer chemotherapy because of several major clinical considerations (11). The nature of the growth characteristics of tumors in man requires that repeated courses of treatment be given over extended periods of time; this approach carries with it an increased risk for the development of drug toxicity. Cancer chemotherapy agents, as a group, fail in varying degrees to discriminate efficiently between normal and target tissues. A wide range of qualitative toxicities are encountered; with many compounds, administration of moderately toxic doses may be required in order to achieve a maximum therapeutic response. In addition, a proportion of patients to whom new compounds are to be given are debilitated by disease and/or previous therapy and thus are potentially less tolerant of drug toxicity. Many of the drugs act upon DNA as their principal mechanism for cytotoxic action. This derangement of genetic information has been correlated with teratogenesis, mutagenesis, and carcinogenesis in lower animals and probably in man.

The principal objective of animal toxicology is the generation of information about the new agent that may forewarn the physician about potential drug hazards that may be encountered during the initial trial. In this process, toxicity data are extrapolated from one species to another. This involves 2 implicit assumptions: toxicity tests in a particular animal species have significant predictive value for man; important toxicity will not go unpredicted. The scientific community has remained somewhat skeptical of this basic assumption of predictivity, and the overall value of animal toxicological studies has been seriously questioned. Nevertheless, this problem has received remarkably little systematic study. The pharmaceutical industry continued to generate extensive descriptive information relating to the toxicity of a new compound in animals as required by federal regulatory agencies, without an apparent attempt by either party to analyze these data for the larger question of its overall usefulness.

In recognition of the potential importance of animal toxicological testing as a predictive system, the Laboratory of Toxicology of the NCI has developed an extensive protocol for use by its contractors (9). The Food and Drug Administration has accepted this protocol and offers it as a model for preclinical evaluation of anticancer agents. The major goals of this program include the following (3). There is an attempt to define drug doses that produce distinct levels of toxicity in dogs and monkeys. These studies include a determination of the highest dose at which no hematological, chemical, or morphological drug toxicity is observed. In addition, the lowest dose that produces toxicity, a high toxic dose, and the lethal dose are to be delineated. The program attempts to indicate the major organ toxicity in both experimental species and to evaluate the predictability and reversibility of observed toxicity. The consistency of qualitative and quantitative toxicity within and between experimental species is compared. The influence of dosage schedules on drug toxicity is evaluated. The schedules currently being tested include single-dose studies in dogs, intermittent 5-day courses in dogs and, for particular drugs, selected schedule dependency studies in dogs. The animals are held for periods of prolonged observation to determine the potential for delayed drug toxicity. Elucidation of the etiological aspects of drug-induced toxicity is attempted by analysis of the hematological, chemical, and histopathological findings.

Protocols such as that developed and used by the NCI go beyond the simple purpose of safe clinical introduction of a new agent, but they also strive to advance the state of the art of drug safety evaluation. If this protocol is to be used as a model for the entire pharmaceutical industry, then certain basic questions should be raised regarding the necessity of such extensive animal testing and the validity of the observations.

1. What is the justification for each phase of the program? In essence, which of the many studies are actually required for safe introduction of a new compound into clinical trial, and which aspects of the protocol represent "research"? Is the pharmaceutical industry required to carry out a similarly elaborate toxicological evaluation before the value of each facet of the program is well documented?

2. What is the cost/benefit ratio of such extensive animal testing? Some of the costs are clear: (a) approximately $100,000 per compound under the current guidelines of the protocol, and (b) costs in time. A protocol study takes 8 to 12 months for completion. There is great concern on the part of clinicians that an excessively elaborate protocol may unnecessarily delay completion of the preclinical evaluation and as a result increase the time before a drug can be introduced into clinical trial.
3. Does such a protocol constitute the most rational and efficient use of dogs and monkeys?

The field of drug development of anticancer agents is in a unique position to outline an efficient and rapid system for toxicity evaluation. Unlike many other important classes of pharmacological agents, a great deal of data relating to the toxicity of anticancer agents have been developed and analyzed that allow for the formulation of reasonable guidelines. It has been demonstrated that information derived from animal toxicology can be useful in 5 specific ways.

1. Determination of a Safe Dose for Phase 1 Clinical Trials. Freireich et al. (4) compared the quantitative toxicity of 18 anticancer drugs in 6 species, after correcting the data to a uniform schedule of daily treatment for 5 consecutive days. This analysis demonstrated that mouse, rat, dog, monkey, and man have essentially the same MTD when compared on a basis of mg/sq m of body surface area. The study suggested that Phase 1 clinical trials could be safely initiated at a dose one-third that of the animal MTD. The mouse was as useful as any other species in this regard. Homan (8) applied regression analysis to the MTD’s of 25 anticancer drugs for dog and monkey. It was estimated that one-tenth of the dog MTD in mg/sq m carried a 1.3% risk of exceeding the human MTD in Phase 1 testing; the corresponding risk for monkey data was 1.1%. If the initial clinical dose is increased to one-third of the mg/sq m animal MTD, the risk for exceeding the human MTD was 10% for each of the animal species and 6% if the more sensitive species was used. Goldsmith et al. (5) have extended these observations in a retrospective analysis of the ability of mouse, dog, and monkey to predict for quantitative toxicity of 30 anticancer agents. The schedules of administration included single doses and repeated daily treatment. The mouse proved to be as good as, if not a better, predictive system as the dog and monkey for ensuring a safe starting dose for man. Using one-third of the mg/sq m of the lethal dose of 10% of the normal mice, this species overpredicted the human MTD for only 2 of 29 (1) drugs tested. Using one-third of the mg/sq m lowest toxic dose in dogs and monkeys, 21 to 24% of initial doses in man would have been toxic.

Each of the 3 studies dealing with the quantitative prediction of drug prediction have been criticized because of 1 or more ways in which the data were handled. These have included the range of animal strains within each given species, the specific dose designated as the MTD for the animal species, and the normalization of the data to confirm to a standard schedule. Nevertheless, the results of each of 3 independent studies are similar; the introduction of a new drug at one-tenth to one-third of the MTD of the more sensitive animal species is a safe method for initiating a Phase 1 trial. This approach remains the standard policy for new drug investigation.

2. Prediction of Qualitative Toxicities. A detailed retrospective analysis of drug and monkey prediction of human organ system toxicity has been conducted for 25 anticancer compounds of diverse chemical and functional classification (10). The study used data accrued from an earlier NCI toxicity protocol that required only the determination of acute lethal toxicity in dogs and subacute repeated-dose toxicity in dogs and monkeys (1). Nevertheless, the observations furnish a valid benchmark for assessing the potential usefulness of newly designed animal toxicology.

The dog and monkey proved equal in the prediction of leukopenia with 60 and 61% true positives, and 20 and 17% false negatives, respectively. The dog system was more accurate and efficient in the prediction of anemia with a high percentage of true positives and fewer false-positive and false-negative predictions when compared to the monkey. The monkey proved to be quite resistant to the development of thrombocytopenia, with 56% false-negative predictions compared to 19% for the dog, and is therefore a less acceptable model for this important clinical problem. In the case of gastrointestinal toxicity the dog was the superior predictive species with 92% true positives and no false negatives compared to 74% true positives and 19% false negatives, respectively, for the monkey. Both species grossly overpredicted for liver toxicity with 44% false positive predictions for the dog and 35% for the monkey. A similar overprediction was noted for the estimation of renal toxicity; the dog presented a 56% false-positive index with 20% false negatives, while the monkey produced a 48% false-positive prediction with 10% false negatives.

Overall, the combined large animal screen did predict a useful population of the total spectrum of qualitative toxicities inherent in a new compound. However, the correct predictions are accomplished at the cost of a high percentage of false positives, particularly for renal and hepatic toxicity. There are several possible explanations of the high incidence of overprediction. For determination of all possible qualitative toxicities inherent in a given compound, the animals had to be given severely toxic, sometimes lethal-dose levels. This was particularly true for the prediction of hematological, liver, and neurological toxicity. This stands in contrast to a Phase 1 clinical trial where the study is discontinued when the 1st dose-limiting toxicity is encountered. In addition, there appear to be definite differences in organ system sensitivity to drug toxicity among the 3 species; man may demonstrate toxicity in an organ system predicted susceptible by the animal studies; however, this toxicity may be expressed in a different specific clinical or chemical parameter. It was also observed that the adverse reaction may appear in man at a greater or lesser dose level than in the animal, or it may follow a different order of appearance in relation to the total spectrum of qualitative toxicity inherent in the compound.

We have recognized an increased sensitivity of certain animal organ systems, in particular the liver and kidney, to drug effects. The question may arise as to whether one can justify not taking a compound into clinical trial based on evidence that the drug produces serious hepatic dysfunction in dogs. The principal criterion for testing a new anticancer agent in man should remain the expectation of therapeutic efficacy, not the toxicological findings in a drug-sensitive organ system of the animal. At the same time it must be recognized that the animals may fail to predict important toxicities that are subsequently observed in man. There are many such examples, including the cases of...
pancreatitis, observed during the clinical trials with L-asparaginase and daunomycin-induced cardiac toxicity, which were wholly unappreciated in dogs and monkeys. These examples of underprediction by the animals may ultimately represent inherent biological differences among the species, as was clearly demonstrated by the resistance of the monkey to the development of thrombocytopenia and gastrointestinal toxicity. In some cases, a false negative may reflect the relatively small numbers of animals used in toxicity evaluation compared to the large number of patients who receive the new drugs. In addition, normal healthy animals are used in toxicological studies, in contrast to the patient population with which they are being compared; many of the patients have preexisting organ failure resulting from disease and prior therapy, and such patients are almost always receiving concurrent analgesics, hypnotics, and antiemetics that complicate interpretation of clinical data.

3. Prediction of Delayed Toxicity. With extended observation periods after treatment, the animal screen was able to predict those compounds for which toxicity was delayed in onset. This has been demonstrated for the chloroethylnitrosoureas and the imidazolecarboxamide group of compounds, although the prediction of delayed toxicity did not correlate directly with the same organ system affected in man. For the nitrosoureas, dogs predicted delayed hepatic toxicity, whereas in man the target organ is the bone marrow (2, 7).

4. Prediction of Bone Marrow Toxicity. The animal toxicology screen has, in general, correctly identified drugs that lack bone marrow toxicity, as in the case of bleomycin and streptozotocin. This is a particularly important feature of any new anticancer compound, in view of the myelosuppressive features of most current chemotherapeutic agents.

5. Exploitable Toxicity. Toxicological studies have on occasion identified a specific toxic feature of a new compound that may be exploited clinically. This has been the case with streptozotocin, in which the diabetogenic property of this compound has been used effectively in treatment for malignant insulinoma. Similarly, the observation that 1,1-dichloro-2-(o-chlorophenyl)-2-(p-chlorophenyl)ethane selectively destroyed the adrenal cortex of the dog led to the eventual successful use of this compound in the treatment of adrenocortical carcinoma.

The data derived from animal toxicity have not been systematically analyzed to determine whether it can be used to predict a schedule of administration that might produce the lowest degree of toxicity. This problem has been examined for the antimetabolite 5-fluoro-2-deoxyxuridine. When this agent is administered by 24-hr i.v. infusion in man, both hematological and gastrointestinal toxicity is accentuated as a result of a decreased catabolism and increased phosphorylation to form the "S-phase" active deoxynucleotide. In contrast, the dog failed to predict for the marked increase in toxicity of 5-fluoro-2-deoxyxuridine when the antimetabolite was administered by continuous infusion as compared to single injections (6).

On the basis of the analysis of available data and a realistic appraisal of animal prediction, the following guidelines can be offered for preclinical toxicological testing for initiating a Phase 1 trial of a new anticancer agent. (a) The lethal dose in 10, 50, and 90% of the mice, and MTD in dogs, should be determined in single-dose studies so as to provide a safe initial clinical dose. This is the single most important step of any Phase 1 trial, since once this has been achieved, subsequent doses can be cautiously escalated to a MTD or a maximal effective dose schedule. There appears to be no reason to conduct such studies in monkeys, since this species offers no specific advantage over mouse and dog. (b) Qualitative organ system toxicities should be delineated in dogs, both in the single-dose study cited above, and in a subacute study. I would recommend a daily dose for 14 days, with one-half the animals held for 60 days of observation for possible delayed toxicity. This would ensure that all toxicities requiring either high dose or accumulative injury for expression would be demonstrated and would furnish dosage data for multiple-dose trials. Additional schedules of administration are not required, since the present basis for selecting a dose schedule in Phase 1 studies is arbitrary and in practice is not influenced by schedules used in animal toxicology. Monkey data do not add significantly to the prediction of qualitative toxicities in man and there is no reason to require that this species be routinely included. A detailed analysis of the predictive ability of the mouse and rat should be conducted to determine whether either of these rodent species could substitute for the monkey for prediction of qualitative toxicity. It can be assumed that many of the newly developed anticancer agents will eventually be demonstrated to be both teratogenic and carcinogenic in laboratory animals. Clearly, such data should be obtained to determine the full biological and toxicological potential of each compound prior to NDA approval. However, completion of these additional studies is not and should not be requirement for initiation of a Phase 1 trial in a patient with active cancer and a limited life expectancy. Obviously, careful selection of candidate patients for Phase 1 trial is paramount, and pregnant women and patients with benign disease must be excluded from study.

These recommendations apply only to the direct application of animal toxicological data for initiation of Phase 1 clinical trials of new anticancer agents. The animal systems are imperfect, and those who wish to use them as if they have direct relevance for man will be sorely disappointed. Much of the animal qualitative toxicity data do not correspond closely to man, and as a result many of the original goals of the NCI protocol may be unrealistic. There is considerably less published information that relates to the value of animal prediction of chronic drug effects. Such data are almost certainly available in the files of the Food and Drug Administration, NCI, and pharmaceutical companies and should be expeditiously analyzed to determine whether chronic animal toxicology has sufficient merit to warrant the time and expense. In view of the extreme complexity in determining the predictability of animal systems for individual anticancer agents, it is unlikely that combination chemotherapy regimens can be reliably tested prospectively for toxicity.
There remains a definite need for continued research in the field of drug safety evaluation. In particular, the efforts of the NCI Laboratory of Toxicology to improve methodology through the testing of new approaches in the current protocol, and in ensuring objective observation and analysis of data, have added immeasurably to our ability to make effective use of information being generated from animal screens.

Ultimately, the initiation and conduct of a safe and informative Phase 1 trial requires the realistic utilization of animal toxicological data coupled with careful monitoring, judgment, and expectation by a trained and experienced clinical pharmacologist.

REFERENCES

Preclinical Toxicology of Anticancer Agents

Philip S. Schein