PAPER NO. 6

CHEMICAL CARCINOGENESIS TESTING AND RELATED ISSUES - SUBCHRONIC STUDIES AND RELATED ISSUES

Andrew Sivak

Arthur D. Little, Inc.
Cambridge, MA 02140

INTRODUCTION

The prechronic studies in the context of their use in the carcinogenesis bioassay program of the National Toxicology Program (NTP) were examined in terms of the general use and design of this type of study as a toxicological tool. As the information in Table 1 reveals, the general protocols and end points for prechronic studies as described by several agencies, including one external to the United States, exhibit a remarkable similarity. Moreover, the use of the data by agencies as part of the overall toxicological dossier of a chemical suggests that NTP should consider the prechronic study as it now is formulated or could be modified to provide direct information on toxic hazards of chemicals, in addition to its initially stated purpose.

TABLE 1. PRECHRONIC STUDY GUIDELINES

	EPA Pesticide Assessment Guidelines (1982)	FDA "Red Book" (1982)	FDA IND/NDA Pharmacology Review Guidelines (1981)	OECD (1981)	EPA Health Effects Test Guidelines (1982)	NTP (1978)
Purpose	Pesticide registration support	Food additives and color additives safety assessment	IND/NDA Pharmacology review guidelines (1981)	Assessment and evaluation of toxic characteristics	Select chronic dose levels	Predict dose range for chronic study
Purpose	No-observed-effect-level	No observed adverse effects, no-effect-level	Characterize pharmacology, toxicology, pharmacokinetics and metabolism of drugs for precautionary clinical decisions	Select chronic dose levels	Establish safety criteria for human exposure	
				Useful information and permissible human exposure	No-observed-effect-level	
Species	Rat (20)[a], Dog (8)	Rat (40), Dog (8)	Rat/Mouse/Other Rodents; Dog/Monkey/Other non-rodents	Rat (20), Dog (8)	Rats (20), Dogs (8)	Fischer 344 Rats (100), B6C3F1 Mice (100)
Doses	3 dose levels	3 dose levels	3 dose levels	3 dose levels	3 dose levels	3 dose levels
End Points	Opthalmology Hematology	Opthalmology Hematology	Opthalmology Hematology	Opthalmology Hematology	Opthalmology Hematology	
End Points	Clinical Chemistry	Clinical Chemistry	Clinical Chemistry	Clinical Chemistry	Clinical Chemistry	
End Points	Histopathology (50)[b] Rat (30), Dog (33)	Histopathology Rat (41)	Histopathology	Histopathology (41)	Histology (41)	Weight Loss Histopathology
End Points	Target Organs	Target Organs	Target Organs Behavioral and Pharmacological Effects	Target Organs	Target Organs	Target Organs

[a] (N) Number of animals per test group.
[b] (N) Number of tissues per animal.

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

TITLE: Proceedings of the Annual Conference on Environmental Toxicology (15th) Held in Dayton, Ohio on October 30, 31 and November 1, 1984.

TO ORDER THE COMPLETE COMPILATION REPORT, USE AD-A178 248.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDING, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#: AD-P005 139 - AD-P005 160 AD#:

AD#: AD#:

AD#: AD#:

The paragraph quoted below from the "Guidelines for Carcinogen Bioassay in Small Rodents" sets forth this purpose as stated in 1976 (Sontag et al., 1976):

> "VII.C.1. Need. The purpose of the subchronic study is to predict the toxic effects which may occur in the animals during the chronic administration of the test agent. Based on the subchronic observations, a prediction is made of the maximum tolerated dose (MTD) that can be given to the animals in the chronic study without producing unwanted side effects."

The relationship between the MTD and "unwanted side effects" was described as follows:

> "VII.C.7. MTD Determination. A MAXIMUM TOLERATED DOSE (MTD) should be selected for each sex of each strain to be used in the chronic study. The MTD is defined as the highest dose of the test agent given during the chronic study that can be predicted not to alter the animals' normal longevity from effects other than carcinogenicity. The MTD is estimated after a review of the subchronic data. Since these data may not always be easily interpretable, a degree of judgment is often necessary in estimating the MTD. The MTD should be the highest dose that causes no more than a 10% weight decrement[1], as compared to the appropriate control groups; and does not produce mortality, clinical signs of toxicity, or pathologic lesions (other than those that may be related to a neoplastic response) that would be predicted to shorten the animal's natural life span. Other measurements (see VII.C.6) also may be used to aid in predicting the MTD."

Thus, the information base from the prechronic studies initially designed only to select the doses to be used in the chronic study was limited, with considerable emphasis on weight gain decrement as the controlling factor. As the program matured, it became apparent that the dose selection process could be refined by use of a broader range of information obtained directly from the prechronic study as well as information from other research programs (chemical disposition, genotoxicity) being carried out by NTP and others (Schwetz, 1983).

---

[1] "Although a depressed weight gain is a clinical sign of toxicity, this particular effect is acceptable when estimating the MTD."

The Food Safety Council (1980) has presented in its Proposed System for Food Safety Assessment a detailed discussion of subchronic studies which addresses many of the issues of concern to this panel. The Department of Health and Social Security (1982) of the United Kingdom has recently published a set of "Guidelines for the Testing of Chemicals for Carcinogenicity" which also considers the factors associated with design and interpretation of carcinogenicity studies.

In addition to the dose selection issue, a number of additional operational factors were identified that had major impacts on the prechronic studies and chronic bioassay and were perceived to be appropriate for reevaluation and possible modification to improve the interpretability of the results and their relationship to assessments of toxic effects in humans.

As a consequence of discussions among members of the ad hoc panel and representatives of NTP who provided technical guidance, the following issues were reviewed:

1. Chemical selection process

2. Suitability of continued use of F344 rat and B6C3F1 mouse

3. Establish dose range, number of doses, administration route and vehicle

    a. Use of toxicological parameters (target organs, weight loss, organ function, histopathology, clinical chemistry, hematology, other toxic signs)
    b. Use of chemical disposition and metabolism data, and human exposure information
    c. Factors affecting dose route and vehicle
    d. Use of other supporting information (short term test data)

The process used was to develop a position paper on each item that would serve as a basis for the recommendations offered in each section. In addition to contributions by subcommittee members Richard Adamson, National Cancer Institute and Perry Gehring, Dow Chemical Company as well as the special panel assembled to review the issues of pharmacokinetics, the comments from outside reviewers in the government, academic, and commercial communities provided much valuable input. For several of the issues further amplification and discussion is provided in the paper dealing with design of the chronic study.

THE CHEMICAL SELECTION PROCESS

RECOMMENDATIONS

1. The development by NTP of a set of definitive criteria for each of the selection elements would provide a more uniform and justifiable process for chemical selection.

2. NTP should consider a methodology to make appropriate use of the exposure factor in the chemical selection process given its dominant position in assessment of risk.

3. The NTP Board of Scientific Counselors should ensure that the opportunity is maintained and enhanced for participation of interested parties in the review of decisions made between the subchronic test and the chronic bioassay.

BACKGROUND

Although the chemical selection process for NTP is not directly part of the prechronic studies area, its important early role in determination of which chemicals will be placed in the assay stream indicates that this process should be evaluated. We were aware of the National Academy of Sciences "Testing Needs Study" and one of their primary goals to "integrate the intensity of selected toxicity data elements . . . . . to set priorities." In light of public comments that were made with respect to the chemical selection process, we included its consideration in our evaluation.

The NTP chemical nomination and selection process shown in Figure 1 lays out the key elements in the process including (1) access to the nomination process by any interested party, (2) multiple review steps and (3) participation in this review by members of appropriate Federal agencies. The lineage of the process, though somewhat cumbersome, does appear to address the needs of the different constituencies involved. Several issues were raised by commentors from the public sector. The first of these dealt with the quality of the executive summaries prepared for the chemical evaluation committee. According to the commentors, the quality of these summaries was highly variable and in some cases incomplete and/or inaccurate. Concern was expressed that key decisions on setting priorities for entry of chemicals into the bioassay process may have resulted in the bioassay of chemicals of lesser urgency than those that were deferred or not tested at all.

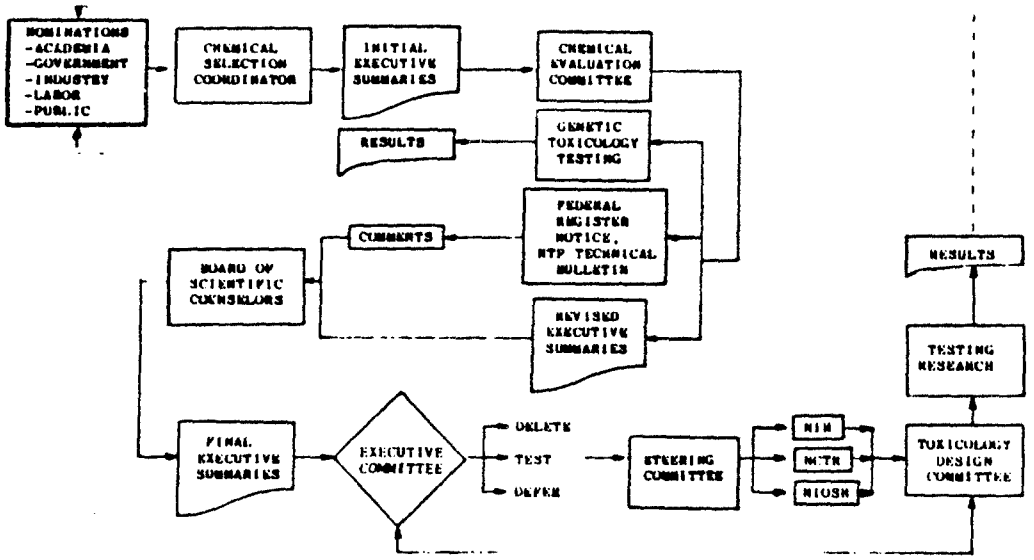


Figure 1. NTP Chemical Nomination and Selection Process.

A second related issue that was raised by a number of commentors related to external comment and peer review at the prechronic/chronic interface. In earlier times, the operation of the testing program was such that opportunities for comment at the prechronic/chronic interface were limited. As the program matured, information relating to testing plans and schedules for the components of the bioassay have been publicly announced in the Federal Register and the NTP Annual Plan and contain invitation for public input, and thus, there exist provisions for a regular and consistent opportunity for public input which have been utilized by interested parties. The balance between public participation in the decision process and the need to carry out an efficient program without undue delay appears to be the crux of the debate. The development and judicious application of a decision matrix based on the best data available (weight gain, organ specific toxicity, clinical chemistry, pharmacokinetics) would tend to reduce the concerns of those external to the program and obviate to some degree the need for multiple peer reviews. An important issue in the use of such a matrix is that it be used with flexibility recognizing the imperfect nature of the data that will be available and the basic scientific problems in experimental design of carcinogenesis studies.

One important segment of information that has not been publicly available in bioassay reports until relatively recently is a disclosure of the scientific rationale for selection of the

doses in the chronic study. Even in recent reports, the description of the manner in which the available data from the subchronic studies and other data are used to arrive at the chronic doses could be explained in more detail and made clearer and more definitive.

The third issue raised addressed the criteria used in application of the chemical nomination elements. Although the elements span the range of information suitable for setting priorities for chemical selection (Table 2), there are no defined criteria against which the data elements can be tested and weighted in the nomination and selection process. The development and application of a set of criteria would aid in providing a more rigorous and less subjective basis for chemical selection. Among the chemical nomination elements, there is no specific identification of how human exposure in either magnitude or frequency is used as a key factor. Since exposure is a dominant, if not controlling, factor in the assessment of risk, it would appear that its use in the nomination sequence should be of primary concern. While detailed exposure scenarios may not be readily available for a large number of chemicals, especially those now under consideration, an effort in this direction to obtain a more quantitative exposure analysis would add considerable power to the chemical selection process as well as providing guidance in setting a dose range for the chronic bioassay.

SUITABILITY OF CONTINUED USE OF FISCHER 344 RAT B6C3F1 MOUSE

RECOMMENDATIONS

1. For the present, it is recommended that NTP maintain the two species, (Fischer 344 rat and B6C3F1 mouse) presently used for carcinogenesis bioassays in the NTP program.

2. Based on data available both within the program and without, the NTP Board of Scientific Counselors should explore whether continued use of both a rat and mouse strain is needed for detection of carcinogens given the range of assay tools available to the oncology research community.

3. If a determination is made to maintain a two species bioassay protocol, give serious consideration to replacement of the B6C3F1 mouse with a strain having an established lower and less variable spontaneous incidence of important tumors that are induced by chemicals.

In addition, continued investigation of the use of other species as adjuncts or replacements for the ones now in use should be undertaken.

TABLE 2. NTP CHEMICAL NOMINATION ELEMENTS

I. Chemical Identification

- A. Chemical Abstracts Service (CAS) preferred name
- B. Common or generic name and synonyms
- C. CAS registry Number
- D. Chemical class and related compounds
- E. Physical and chemical properties
  - 1. Physical description
  - 2. Structural and molecular formula and molecular weight
  - 3. Melting and boiling points
  - 4. Solubility
  - 5. Stability and reactivity
  - 6. Other relevant information
- F. Commercial product(s) composition
- G. References

II. Production, Use, Occurrences, and Analysis

- A. Production
  - 1. Source and synthesis, year and pathway of first production
  - 2. Current production and pathway
- B. Uses
- C. Occurrence in the Environment
  - 1. Naturally occurring
  - 2. Air, water, and soil
  - 3. Occupational
- D. Analysis
- E. References

III. Toxicology

- A. Human data, case reports, and epidemiological studies
- B. Experimental animal information
- C. In vitro and other short-term tests
- D. Other relevant information
- E. References

IV. Disposition and structure-activity-relations

- A. Absorption, distribution, metabolism and excretion
- B. Structure-activity correlations and considerations
- C. References

V. Ongoing Toxicological and Environmental Studies in the Government, Industry, and Academia

VI. Rationale for Recommendation and Suggested Studies

HISTORICAL RATIONALE FOR SELECTION OF ANIMAL STRAINS FOR CARCINOGENICITY STUDIES

The recent description of the history of the carcinogenesis bioassay program at the National Cancer Institute by Elizabeth Weisburger (1983) provides a modest insight into the rationale for selection of the Fischer 344 rat and B6C3F1 mouse as the animals of choice to carry out chronic carcinogenesis studies.

"The search for better and more sensitive models for testing led to extensive tests on the development of mammary cancer in virgin female Sprague-Dawley rats. The use of infant mice and rats to increase sensitivity was explored thoroughly and several publications resulted therefrom. The Sprague-Dawley derived rat from Charles River Laboratories was also used in several other exploratory projects and in some bioassays. However, the large size and numerous spontaneous tumors of this strain of rat detracted from its value. In 1970 the F344 rat, which was easy to maintain, sensitive to various types of chemical carcinogens, and had a relatively low spontaneous tumor rate (except for the testicular Leydig cell tumors of aged nonbreeding males), was therefore adopted as the standard rat for the bioassay program. The CD-1 Ha/ICR mouse was also found to have an unacceptable high spontaneous tumor incidence and relatively short survival and was subsequently replaced in the bioassay program by the B6C3F1 mouse."

The issue of selection of appropriate experimental animals for biological tests, specifically cancer studies has also been commented upon by several members of the carcincgenesis research community.

At a symposium in 1967, Shimkin made several interesting observations that bear recalling. In particular, his views on the necessity for inbred strains of animals.

"The dictum that 'inbred mice are as essential to the biologist as pure chemicals are for the chemist' needs reexamination. The fact is that practically all key discoveries in chemical carcinogenesis were made, and continue to be made, on less exacting biological material which is easier and cheaper to procure and which is often more hardy and longer lived. We now have added pathogen-free and bacteria-free and newborn animals, which certainly have intriguing uses. But such animals, just as homozygous strains, are not necessarily more desirable than more prosaic ones for many biological and carcinogenic experiments. It is our hope that official standards will not be promulgated in regard to animal systems. More appropriate is good experimental design, with adequate controls, as selected by the individual investigator."

In contrast, Festing (1979) has, for a number of years, presented detailed arguments supporting the use of inbred strains of rodents for chronic toxicity and carcinogenicity studies. In

his view, the use of inbred strains will aid in achieving "uniform experiment material" which will lead to a reduction of variability in the responses. With proper controls of environment, nutrition, and infectious disease and allocation of animals to test control groups in an orderly and comparable fashion, the probability of false positive (Type I) and false negative (Type II) errors should be substantially reduced. A similar position supporting the use of inbred strains of test animals has been offered by Haseman and Hoel (1979) and also in contrast to his earlier position by Shimkin (1974) who proposed a general set of criteria that could be applied to the selection process for test animals in a carcinogenicity study.

They are:

1. Availability
2. Economy
3. Sensitivity to carcinogens
4. Stable as to response
5. Similarity to man in regard to metabolism
6. Similarity to man in regard to pathology responses

Most regulatory and government advisory bodies have not taken positions with respect to recommending specific strains of animals for use in carcinogenesis or other chronic exposure studies. For example, the British Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment (1982) in its "Guidelines for the Testing of Chemicals for Carcinogenicity" recognized the various competing factors in animal strain selection and concluded, "Extensive background or experience of a particular inbred or outbred strain in a particular laboratory may be a strong reason for choosing it. Where the test substance is chemically related to a known carcinogen, the choice of strains may be influenced by the strain used in the tests of the related substance. Otherwise the Committee felt unable to come down firmly in favor of specific inbred or outbred strains of rats or mice."

The Committee's report continues:

> "Many strains of rats and mice have high 'spontaneous' incidences of particular kinds of neoplasm. For instance, particular strains of mice, both inbred and outbred, may be exceptionally prone to the development of adenomatous tumours of the lungs, benign and malignant parenchymal cell neoplasms of the liver, malignant lymphoma or mammary tumours, and some strains of rats may be prone to the development of mammary

tumours, interstitial-cell tumours of the testis or pituitary adenomas. Exposure of such strains to test substances may be associated with an increased incidence of the same kind of tumours. Opinion is divided as to the extent to which such increases should be regarded as evidence of carcinogenic potential for man. Similarly, opinion is divided as to whether high-spontaneous tumour incidence strains should be used in preference to low-spontaneous tumour incidence strains for carcinogenicity testing."

The Food Safety Council (1982) in its "Proposed System for Food Safety Assessment" similarly sets out some general principles in the selection of test species.

It is only considerably after the fact that the xenobiotic metabolic characterization has been examined in the test strains. It should be emphasized that the NTP embarked on acquiring such metabolic data in an expeditious fashion upon acquiring responsibility for the bioassay program.

The overriding issue, by far, in the continued use of originally selected strains of test animals is the impact of the occurrence of spontaneous hepatic neoplastic tumors in B6C3F1 mice on the interpretation of the effects of chemicals in inducing these kinds of lesions.

THE FISCHER 344 RAT

At the time of the selection of the Fischer 344 rat for use in the NCI carcinogenesis bioassay studies, there was not a large data base on the sensitivity of this strain to carcinogens or on the metabolic capability of the strain for xenobiotics. Neither was there a detailed understanding of the profile of spontaneous tumors over the planned two year interval of the bioassay. In some of the earlier studies in the NCI Bioassay Program, Osborne-Mendel rats were used; however, this strain has not been used extensively by the NTP in its bioassay studies.

As the data base has grown, it has become evident that this strain of rat is vigorous with good survival characteristics (Solleveld et al., 1984). Moreover, the spontaneous tumor profile, which is now well documented, does not appear to have any substantial impact on the interpretation of chemically induced tumor response (Haseman, 1983).

Recent analyses of the responses of Fischer 344 rats in the NTP program have revealed some interesting data that bear further study as well as providing information that supports the

continued use of this strain of rat as a test animal in carcinogenesis bioassays. Haseman's (1983) evaluation of patterns of tumor incidence among 25 feeding studies in the NTP program revealed a negative correlation between incidences of liver tumors and leukemia/lymphoma. In addition, reduction in the incidence of breast fibroadenomas was associated with decreased weight gain. Evaluation of the spontaneous tumor spectrum reveals only one tumor type (testicular interstitial cell tumor) that reaches a very high incidence (92%) in the 104 week assay interval. Other tumors that exhibit greater than a 20% incidence in controls are mononuclear cell leukemia and adrenal medullar pheochromocytoma in males and breast fibroadenoma, mononuclear cell leukemia, and pituitary gland adenoma in females. The comparable survival (median lifespan of 28 months) of males and females and the observation that the variety of neoplastic lesions in animals carried for their lifespan was not greater than in animals killed between 110 and 116 weeks of age, although the overall frequency of these lesions increased markedly as the animals aged beyond the time they would usually be employed in a standard bioassay.

THE B6C3F1 MOUSE

The use of the B6C3F1 mouse as one of the strains in the carcinogenesis bioassay program has spawned a considerable debate that emanates from the occurrence of spontaneous liver tumors. This characteristic of substantial hepatic tumor incidence in untreated male mice has raised questions about the interpretation of experiments in which an increased liver tumor incidence is observed associated with chemical exposure. This issue has been considered by several individual investigators as well as by a number of workshops and symposia. Because of the attention focused on this issue and its apparent importance in the determination of the carcinogenicity of several classes of chemicals, especially the chlorinated hydrocarbons of both the aliphatic and aromatic types, it has been considered by several individual investigators as well as by a number of workshops and symposia (Squire and Levitt, 1975; Society of Toxicology Pathologists, 1982).

The basic source of the debate appears to center on the identification of proliferative lesions in the livers of mice of the B6C3F1 strain and the fate of those lesions. At one end of the spectrum, the position is espoused that proliferative lesions, whether hyperplastic nodules, benign adenomas or carcinomas, are all indicative of carcinogenic potential (Ward et al., 1979). In contrast other arguments have been made that evidence for malignancy is essential before designating a

chemical as a carcinogen (Newberne, 1982; Newberne et al., 1982; Vesselinovitch, 1982). In addition, the wide variation in spontaneous liver lesions in male B6C3F1 mice among the various test laboratories, and even within a single laboratory, among those participating in the carcinogenicity bioassay program raises further questions with respect to the interpretation and use of these lesions as markers for direct induction of carcinogenicity by a chemical agent (Nutrition Foundation, 1983). Indeed, the variation in occurrence of hepatic proliferative lesions among mice appears to be a general characteristic of these species (Grasso and Hardy, 1975).

Marenpot and Boorman (1982) summarize the NTP position with respect to liver tumor responses:

> "Examination of factors involved in the interpretation of liver tumor responses for the four bioassays discussed leads to some general conclusions regarding the current philosophy of assessing carcinogenicity on the basis of two-year rodent bioassays. First, there is agreement among most scientists that induction of hepatocellular carcinomas early in the course of a bioassay, as in the case of the pentachloroethane bioassay, is indicative of carcinogenicity under bioassay conditions. Furthermore, tumor frequency data must be appropriately analyzed when there is early mortality in the bioassay. The example of the pentachloroethane bioassay also exemplifies this point. Second, the occurrence of hepatocellular tumors in more than one species and/or more than one sex makes judgment regarding carcinogenicity more convincing. This was exemplified by the DEHP bioassay results. Third, neither the practice of combining benign and malignant liver tumors nor analyzing them separately for purposes of statistical determination of carcinogenicity is universally accepted. NTP scientists are in the process of establishing policy relative to combining benign and malignant tumors in the near future. Fourth, consideration of a chemical to be carcinogenic on the basis of an increase in hepatocellular adenomas (mice) or neoplastic nodules (rats) will remain controversial until such time as there are data to clearly indicate the biological behavior and significance of the benign liver tumors."

Because a substantial data base has been developed with the B6C3F1 mouse, there is a natural reluctance to change the test strain to one that might present fewer problems. Yet, if the continued use of the present strain does not provide a level of discrimination to yield reasonably definitive results, its continued use must be weighed seriously. Important factors that

should be considered in changing the test strain of mouse relate to evidence that the strain now used does not provide an adequate level of discrimination and that an alternative strain with superior characteristics in terms of longevity, responsiveness, and level of spontaneous tumors has been identified.

One final issue that merits mention is the genetic integrity of the B6C3F1 mouse. Since the occurrence of spontaneous mouse hepatomas is strongly genetically dependent (Grasso and Hardy, 1975), although the source of this genetic influence has not been elaborated, the occurrence of genetic impurity could have dramatic effects on both spontaneous and induced liver tumor incidence, especially related to the contribution of the C3H gene pool to the hybrid. The application of the NTP genetic screening program as described in the 1984 Annual Plan should provide the quality assurance needed to reduce genetic variability as a factor in the chronic bioassay.

Beyond the question of the appropriate mouse strain to use in carcinogenicity studies, the merits of using only a rat strain for carcinogenicity studies has been discussed (van Wittenau and Estes, 1983). With the wide range of short term tests that are applicable and a committment to develop pharmacokinetic data on bioassay compounds, employing the resources used in a long-term bioassay with a second species that may not be needed, and applying these resources to additional studies of mechanisms and/or additional substances is an attractive possibility that bears further study.

Finally, as the program has evolved, a variety of strains of rats have been employed for comparison of toxicologic and metabolic response for a select group of chemicals. The rat strains being used other than the Fischer 344 strain are Osborne-Mendel, ACI 9935, August 28807, Marshall, Long-Evans, Sherman, and Wistar. For the series of asbestos studies that have been carried out by NTP, the species of choice was Syrian golden hamster and for skin painting studies, CD-1 mice have been generally used.

TOXICOLOGICAL AND CHEMICAL DISPOSITION DATA FOR SELECTION OF DOSES FOR CHRONIC STUDIES

RECOMMENDATIONS

1. The use of the MTD as described in this report should be continued. The rationale for selection of the doses for a chronic study and the procedures relating to

utilization of the data from subchronic studies and other sources in this process should be included in all NTP bioassay reports.

2. NTP should continue to develop criteria and methodologies to employ toxicological and pharmacological data as well as human exposure estimates where appropriate to select doses for the chronic study.

3. Pharmacokinetic studies should continue to be conducted before or during the prechronic phase of a bioassay so that a data set as complete as possible will be available to aid in the design of the chronic protocol and the interpretation of the results of the chronic study.

BACKGROUND

As indicated earlier, the conduct of the prechronic studies and the procedures for design of the chronic study are undergoing a continuing evaluation within the NTP bioassay based on new information that is being developed as part of the research program.

The maximum tolerated dose (MTD) is that dose which, when given for the duration of the chronic study as the highest dose, will not shorten the treated animals longevity from any toxic effects other than the induction of neoplasms (Schwetz, 1983). The MTD should not cause morphologic evidence of toxicity of a severity that would interfere with the interpretation of the study. For example, necrosis of a degree to be associated with a significant amount of regeneration may complicate the interpretation of a neoplastic response in the organ. Thus, toxicity and pathology criteria from the subchronic study are the primary criteria for setting the MTD.

The spread of doses below the MTD in chronic studies is determined by one or more of the following factors:

1. The slope of the dose response curve for toxicity in the prechronic studies.
2. The need for a no-effect dose.
3. The need to spread each dose by more than a factor of 10.
4. Relevance to exposure guidelines such as a TLV.
5. The pharmacokinetics of the chemical such as saturation and altered metabolism.

The inclusion in the final bioassay report of a description of the process of dose selection for the chronic bioassay, including the data elements that were employed and the rationale for dose selection based on these data as well as any other input from NTP staff or public sources would aid in the evaluation of the results of the bioassay.

SUBCHRONIC TOXICITY CONSIDERATIONS AND GUIDELINES FOR USING SUBCHRONIC TEST RESULTS FOR SELECTING DOSES FOR CHRONIC TOXICITY TESTS

Before proceeding with a general discussion of toxicity data available for dose selection, it is appropriate to set down some definitions of what data are considered in the evaluation of prechronic and chronic toxicity.

Prechronic toxicity encompasses those effects observed in experiments in which the test animals are exposed to the chemical agent for a period of six months or less, usually 90 days. Common endpoints or markers of prechronic toxicity include organ specific and/or systemic pathology, body weight and organ weight alterations, mortality, clinical signs (including pharmacological or altered physiological parameters), and clinical laboratory measurements (hematology, urinalysis and clinical chemistry). Beyond these classical toxicity measures, the effects of test chemicals on the immune system or on hormonal balance, factors which could have a significant influence on the carcinogenic response, should be considered in the analysis of the subchronic studies for chronic study dose selection where such data are available.

Chronic toxicity studies elucidate effects in test species exposed for longer than six months, and generally focus on endpoints of organ-specific or systemic pathology, both neoplastic and non-neoplastic. In addition, pathophysiology (abnormal function) is usually observed either as a primary or secondary response to treatment.

It is very important to recognize that all observations made in such studies, irrespective of duration of exposure, do not necessarily constitute toxicity per se but commonly span the entire range of effects from altered physiology through pharmacology to toxicity.

The high dose for chronic toxicity testing is generally based upon endpoints achieved in prechronic tests ranging from 14 to 90 days in duration. The slope of the prechronic dose-response curve may influence the range and number of doses

selected for the chronic toxicity test. The high dose for the chronic toxicity test generally lies between the dose level producing a positive endpoint and the dose producing a minimal effect in the three-month subchronic test.

Organ-Specific and/or Systemic Pathology Endpoints

The use of organ specific and pathology endpoints in a subchronic study are important criteria to aid in setting doses for a chronic study. Although no formal documentation for these criteria are presently available, the experience of toxicologists has involved some operational guidelines. An example of how such guidelines may be used is presented below. It should be considered as a model and not a formal mandate for the described procedure.

Evidence of moderate to marked necrosis, (3+ to 4+), degeneration, irritation, inflammation, or atrophy is generally considered a positive effect (endpoint) in prechronic studies. Such positive effects should be found in at least 30% of the test animals at a given dose to judge that dose to yield positive toxicity. If the degree of necrosis, degeneration, irritation, inflammation, or atrophy is minimal to mild (1+ to 2+), then such an effect should be present in at least 60% of the animals to judge that dose to have positive toxicity. Additional judgments regarding positive effects must be made on the basis of the specific lesion and be influenced by the presence or absence of a dose-response for the endpoint in question and whether similar lesions are found in some of the vehicle or untreated controls. In rendering such judgments it is assumed that animals are free of complicating disease and that there is consensus agreement on the quality of the pathologic findings. In practice, the positive dose is often different for males and females of a given species; the same is true for a given sex of two species. Finally, the judgment must be made with respect to whether the pathologic lesion will affect the lifespan in a detrimental way.

An example of pathological findings obtained in a 90 day dermal study is depicted of Table 3. Although tempting, it is inappropriate to select the doses to be used in a chronic study from only the information set forth in this table. Prior to selection of the top dose it must be rationalized whether the purpose of the study is to reveal dermal or systemic toxicity, or both. In any case, the pharmacokinetics of dermally applied doses should be ascertained so data resulting from a chronic study can ultimately be related to human exposure, both dermally and systemically.

TABLE 3. PATHOLOGICAL ALTERATIONS OBSERVED IN A 90 DAY DERMAL STUDY

Lesion	Dose					
	0	1	2	4	8	16
Lung, congestion	2/10	6/10	4/10	4/10	10/10*	9/10*
Thymus, atrophy	0/10	0/10	0/10	0/10	0/10	7/07
Liver, cytomegaly (cellular hypertrophy)	0/10	0/10	4/10	6/10*	10/10*	10/10*
Testes, tubular atrophy	0/10	0/10	0/10	0/10	5/10*	10/10*
Subcutaneous tissue, inflammation	0/10	0/10	0/10	0/10	9/10*	7/10*
Epidermis, hyperplasia	0/10	0/10	0/10	0/10	8/10*	7/10

*P $\leq$ 0.05 (2 - tailed)

Comments on Example:

1. Lung congestion is not a sufficiently definitive lesion for determining an endpoint of toxicity.
2. Diagnosis of thymic atrophy should support thymic organ weight findings; modest reductions in organ weight in the absence of histological evidence of atrophy is too non-specific.
3. Liver cellular hypertrophy (cytomegaly) should support organ weight findings. The endpoint for this specific lesion should be determined in conjunction with pharmacokinetic data since hepatocellular cytomegaly may represent enzyme induction.
4. Although not life-threatening, testicular atrophy is a definitive endpoint of toxicity.
5. Subcutaneous inflammation and epidermal hyperplasia are target-organ-specific endpoints in dermal studies.

In Table 4, specific lesions from various subchronic toxicity studies are listed. These lesions provide rational endpoints upon which dose selection for chronic studies may be based.

In certain instances a constellation of pathologic effects may be present which are secondary to a generalized perturbation which is induced from the primary lesion. For example, the observation of splenic hemosiderosis, centrilobular fatty changes in the liver, gross presence of icterus, myocardial fiber degeneration and necrosis, and bone marrow hyperplasia are secondary effects of hemolytic anemia. Another example might be generalized metastatic calcification (heart, blood vessels in the brain and lung, gastric mucosa and submucosa), myelofibrosis of bone marrow, ulcers in the oral cavity, parathyroid hyperplasia, and renal tubular necrosis. In such a case, the majority of lesions are secondary to the uremia resulting from renal failure.

Judgments regarding selection of a toxic dose in the subchronic study should be influenced by the primary insult more than the secondary or systemic lesion responses.

TABLE 4. SELECTED EXAMPLES OF SPECIFIC TOXIC LESIONS THAT ARE ACCEPTABLE POSITIVE ENDPOINTS IN SUBCHRONIC (3-MONTH) STUDIES

Lesion	Chemical
Myocardial necrosis	isoproterenol adriamycin
Neuronal degeneration	neuroleptics
Brain necrosis	mercury
Status spongiosis	hexachlorophene
Peripheral neuropathy	acrylamide triethyl tin
[a]Thyroid hyperplasia	PBB
Forestomach ulceration with [a]epithelial hyperplasia	2,4-diamionphenol
Hepatocellular necrosis	bromobenzene cycasin
Alveolar edema	paraquat alpha-naphthylthiourea
[a]Nasal cavity epithelial necrosis	formaldehyde acrolein
Testicular atrophy	cadmium DES DDT
[a]Renal tubular karyomegaly	trichloroethylene TRIS
[a]Urinary bladder-transitional cell hyperplasia	saccharin t-butanol
Thymic lymphoid depletion	TCDD DES

[a] These specific lesions may represent pre-neoplastic changes. In selecting a high dose (MTD) for a two-year study, it is appropriate to select a dose known to produce a potential pre-neoplastic lesion to maximize the likelihood of eliciting a carcinogenic effect.

Body Weight and Organ Weight Endpoints

Body weight and organ weight alterations are usually sensitive, quantitative indicators of toxicity (Weil and McCollister, 1963). Decreases in the weight of thymus or testes can often be detected in the absence of clear-cut lesions in these organs. The same is true for increases in liver or kidney weights. Body weight gain data are even more sensitive indicators of toxicity since they can potentially reflect effects not otherwise apparent. Body weight evaluations of necessity must be made in

light of food or water consumption to help rule out decreased consumption due to poor palatability. Both absolute and relative organ weights should be evaluated in establishing dose effects.

Since body weights and organ weights are continuous variables, estimation of toxic endpoints is best achieved by the use of parametric statistical evaluation. Employing the fiducial limit of 0.05 (2-tailed tests), body weight or organ weight changes can be identified. Organ weight changes must be considered in light of other findings (e.g., pathologic alterations, clinical signs, reduced food consumption). Judgments regarding body weight in mice must be tempered by the profound effects that small changes have on measures of group average weights. In specific situations body weight and organ weight alterations can occur in 90-day studies at doses that do not produce pathologic alterations.

Clinical Laboratory Measurements

Hematologic, clinical chemistry, and/or urinary measurements are frequently performed in subchronic 90-day toxicity studies, usually only at the end. Routine hematologic measurements are justified on the basis of providing the most reasonable means for assessing toxic effects on the hematopoietic system. Routine clinical chemistry has generally been less sensitive than organ weights and histopathology in identifying metabolic and functional organ effects in rodents. Certain urinalyses (volume, specific gravity, and microscopic evaluation of urine sediment) are of benefit when renal toxicity is present while other urinalyses procedures (dip-stick qualitative measurements) are rarely more sensitive than histopathology.

Because present day technology permits high precision in clinical laboratory measurements, there are frequent instances where statistically significant effects are flagged in the absence of clinical significance. For example, animals treated with the high dose in a 90-day subchronic toxicity test quite frequently have RBC counts that are 0.5 to 0.8 million cells/mm^3 less than controls. While statistically significant at alpha ≤ 0.05, all results must be evaluated for clinical significance in assessments of treatment-related endpoints.

While the potential utility of hematologic parameters can be intuitively appreciated as useful measures of toxicity to the hematopoietic system, experience with routine clinical chemistry and urinalyses has been less satisfactory. These latter measurements rarely identify toxic endpoints that are not specifically reflected in organ weight and histomorphologic alterations. If

used at all, it is recommended that clinical laboratory tests be prudently selected and timed in 90-day as well as two-year study designs. Selective timing vs. routine performance at a predetermined time interval may enhance the utility of clinical laboratory studies, thereby permitting detection of initial organ insult with subsequent compensation. Such information would theoretically be useful in characterizing mechanisms of toxicity.

At the present time it is recommended that clinically significant changes in hematologic, clinical chemistry, or urinary measurements be used as supporting information to be correlated with more definitive positive toxicity endpoints. With the possible exception of hematologic results, clinical laboratory measurements as currently obtained are not generally sufficient as the sole reason for identifying toxic dose endpoints in prechronic studies, although with certain chemicals they may have specific value in identifying a toxic response.

Indeed, the need for routine collection of complete clinical laboratory data is highly questionable and it is recommended that plans to collect such data should be implemented only when the results are judged to be worthy of collection.

Statistics

Some general thoughts on the use of statistical analyses of data are appropriate here. Statistical evaluation is a tool to assist in the determination of a positive lesion endpoint. However, it is neither a substitute for making informed judgments nor should statistical tests be inappropriately applied in evaluating toxicological data. Many toxicology studies utilize an exploratory strategy which involves measurements of as many as 30 different parameters in a control group versus groups treated with several different dose levels of a chemical. Although some toxicologists may use the t-test, ANOVA with multiple comparisons are typically used for each parameter. The data for each parameter are, however, from the same subjects; thus, the measurements are correlated to an unknown degree and the ability to define the true level of type I or type II error is hopelessly confounded for the experiment as a whole and is difficult to assess.

One possible solution to this problem is to use statistical methods such as t-tests and ANOVA with multiple comparisons as screening devices to indicate those variables that show statistical evidence of biological effects. The methods section should acknowledge the multiple comparisons issue and give some indication of the number of significant effects that might be

expected by chance alone. The toxicologists must then examine the data, use all available information, and make a judgmental decision on a case-by-case basis about the toxicologic significance of the various statistically significant effects observed.

Mortality

Mortality is a highly definitive endpoint of toxicity in 90-day studies. Assuming that the cause of death is clearly associated with treatment, then high dose selection for the two-year study should be set lower than the lowest 90-day study dose with mortality. An example of mortality data which may be useful in selection of the top dose is given in Table 5. In the absence of attendant pathology, the top dose selected for a chronic study should be between the 4 and 8 dosage rates.

TABLE 5. EXAMPLE - MORTALITY DATA

Mortality	0	1	2	4	8	16
	0/10	0/10	0/10	0/10	1/10	10/10

Comments on example:

1. Dose group 8 with mortality constitutes a toxic endpoint. Consequently, the high dose for the two-year study should be lower than 8. Mortality = number dead/number treated.

Clinical Signs

Generally clinical signs will be correlated with and support other endpoints of toxicity. However, in special situations clinical signs seen in prechronic studies will preclude using high doses for the two-year study which cause definitive pathological effects (e.g., anesthesia produced by inhaling high doses of an aerosol; prolonged "intoxication" associated with ingestion of alcohols).

Example of a Composite Data Set

Table 6 presents a typical composite data set from a subchronic evaluation. Assuming that the indicated doses are useful for assessment of the human risk, the top dose selected for a chronic dose should not exceed "2"; in many cases selection of "1" will be appropriate. The latter dose will in most cases

significantly exceed allowable human exposure and still elucidate any insidious chronic toxicity affecting the liver and other tissues as well.

TABLE 6. EXAMPLE (90-DAY DERMAL STUDY) - COMPOSITE OF STUDY FINDINGS

Parameter	Doses					
	0	1	2	3	8	16
Liver/BW		alpha $\leq$ 0.01	alpha $\leq$ 0.01	alpha $\leq$ 0.01	p $\leq$ 0.01	died
Thymus/BW					p $\leq$ 0.01	died
Kidney/BW					p $\leq$ 0.01	died
Reticulocyte count					C.S.	died
Methemoglobin					C.S.	died
Lung, congestion	2/10	6/10	4/10	4/10	10/10	9/10
Thymus, atrophy	0/10	0/10	0/10	0/10	0/10	7/7
Liver, cytomegaly	0/10	0/10	4/10	6/10	10/10	10/10
Testes, atrophy	0/10	0/10	0/10	0/10	5/10	10/10
Subcutaneous inflammation	0/10	0/10	0/10	0/10	9/10	7/10
Epidermal hyperplasia	0/10	0/10	0/10	0/10	8/10	3/10
Mortality	0/10	0/10	0/10	0/10	1/10	10/10

C.S. = clinically significant

PHARMACOKINETIC STUDIES (CHEMICAL DISPOSITION)

In a very general sense, pharmacokinetics is the study of the dynamics of the fate of chemicals in the body. Perhaps the most significant contribution of pharmacokinetics in the design and/or interpretation of toxicology studies is the concept of dose-dependent or non-linear kinetics. Many of the physiological and biochemical processes which affect the disposition of chemicals in the body are capacity limited (i.e., saturable). When any of the rate-limiting processes involving absorption, distribution, metabolism, or excretion of a chemical become saturated, the internal concentration of chemical and/or activated metabolites may not be directly proportional to the administered dose. Instead, disproportionate increases or decreases may be observed, with concurrent effects upon toxicity. Levy (1968) has described some of the criteria which suggest that such saturation may be occurring:

1. The decline of levels of chemicals in the body does not follow an exponential time curve.

2. The biological half-life increases with increasing dose.

3. The area under the plasma concentration versus time curve (AUC) is not proportional to doses.

4. The composition of excretory products may be changed both qualitatively and quantitatively with increasing dose.

5. Competitive inhibition by other chemicals metabolized by the same enzymatic system(s) is likely.

6. Dose response curves show unusually large increases in response with increasing dose, starting with the dose level where saturation effects first become evident.

Another important area of pharmacokinetic studies is an ability to predict bioaccumulation. If a material is eliminated slowly, repeated dosing may result in the introduction of a new chemical into the body before the last dose is gone. Consequently, the eventual concentration of this material in the test organism during a chronic study may be many times higher than anticipated. Polybrominated biphenyls (PBB) are examples of such chemicals. Determination of pharmacokinetic parameters allows the prediction of the level in the body in toto or tissues subsequent to any number of repetitive doses as well as the ultimate steady state levels resulting from continuous or interrupted exposures. For a more detailed discussion of the application of pharmacokinetics, the investigator should consult the following references: Gehring et al. (1976), Tsuchiya and Levy (1972), O'Flaherty (1981), Gibaldi and Perrier (1982), Dayton and Sanders (1983).

Appendix I to this report describes a sample methodology that should be useful to aid in the dose setting process for the chronic bioassay. As indicated above, the establishment of the doses in a chronic study should be the result of an analysis of all the information available in the subchronic study such as (1) variations in weight gain, (2) physiological disturbances, (3) histopathology, (4) dose response of mortality and (5) pharmacokinetics.

Indeed, as indicated in Appendix I, additional methods are available to obtain more detailed information about pharmacokinetics, physiological disposition and metabolism of

xenobiotics. Some of these methods are the use of body compartment distribution models, target tissue binding, metabolic profile, and whole body autoradiography.

As a companion piece to this discussion of dose selection in chronic studies, another review prepared by a group of experienced toxicologists should be consulted for additional details relating to this critical issue (Grice et al., 1984).

FACTORS AFFECTING DOSE ROUTE AND VEHICLE

DOSE ROUTE

Recommendations

1. Exposure to compounds in the carcinogenesis bioassay should reflect tne predominant human exposure route where possible taking into account the need to achieve adequate dose levels for an appropriate study.

2. Should special circumstances indicate that an alternative exposure route may be necessary, an appropriate supporting rationale should be provided including the development of pharmacokinetic and toxicologic data associated with the surrogate exposure route.

3. Since inhalation exposure represents a significant route of human exposure to environmental chemicals, NTP should develop a data base for a selected group of substances tested by several routes to aid in determining whether a single exposure route is adequate to assess carcinogenicity.

Background

Although the general principle of using a dosing route for chronic studies that was comparable to the exposure route in humans has been known and accepted in the toxicology community, there was also general acceptance of use of gavage techniques. The rationales for the use of intubation procedures included a more exact measurement of dosage than available by diet or drinking water procedures and/or ease of delivery of materials whose properties of solubility, palatability, corrosiveness or volatility made them difficult to administer by the usual feeding or drinking modes of dosing (Weisburger and Weisburger, 1967).

In recent years, however, the practice of using gavage administration in corn oil has come under considerable criticism. Although the fundamental reasons for this criticism are not easy to discern, those that have been offered include 1) bolus administration results in highly atypical pharmacokinetic patterns compared to usual human exposures, except where the human exposure is, in fact, a bolus dose, 2) the test animals are subjected to stress on a daily basis with some likelihood of fatality resulting from faulty intubation techniques, 3) substantial increase in personnel resources to perform dosing as compared to diet or drinking water exposure, and 4) necessity to use vehicles (corn oil) that could have a substantial confounding influence on the final outcome of the bioassay. This issue of vehicle for gavage studies, especially as it relates to the use of corn oil, will be considered separately.

In connection with the use of diet and drinking water exposures to model oral intake in humans and the use of skin exposure procedures to model exposure by that route, Page (1977) has discussed extensively the methodology and problems associated with these procedures. The protocols as they are laid out appear to cover the necessary details. Adherence to the protocols requires a rigorous quality assurance program to ensure that the dosage consumed by or applied to the test animals is consistent over the interval of the study.

The inhalation route is the source of exposure of many materials in the environment. Yet, because of substantial resources, in terms of cost and trained personnel, associated with the thorough conduct of such bioassays, they have only had limited use in comparison to the numbers of compounds that would be desirable to test using this route of exposure.

Background

The gavage route of exposure appears to be the only procedure for which the vehicle has become an important issue. With appropriate assurances of test compound stability in food and drinking water and in atmospheres generated in inhalation chambers, there has been no major concern over the dosing vehicle.

Nonetheless, it has become clear that the use of corn oil as a vehicle in chronic gavage studies is not without some impact of its own on the test animals in a number of studies. However, this finding is not consistent in the bioassay program. Three primary issues have emerged. The first, and possibly most critical, of these is the finding in some gavage studies that pancreatic tumors are enhanced in rats receiving corn oil

alone. This observation is enigmatic in that the occurrence of pancreatic tumors in association with corn oil gavage has not been consistently observed over the interval of carcinogenesis bioassay program. Because the incidence of such tumors was further increased in test groups receiving test agents (methylene chloride, benzyl acetate), the problem is further complicated.

The second issue is the one of altered pharmacokinetics associated with administration of test chemicals as a bolus in a lipid solvent. There is substantial evidence that the pattern of metabolism varies considerably with respect to both rate and qualitative nature of the biochemical profile when dosing by gavage in corn oil and by inclusion in drinking water are compared (Newberne et al., 1979; Iritani and Ikeda, 1982; Wade and Norred, 1976; Williams et al., 1983).

The third issue is the nutritional and physiological impact of the corn oil dosing. In particular, the lipid nutritional profile has been shown to affect dramatically the outcome of carcinogenicity studies (Rogers and Newberne, 1980; Newberne et al., 1982; Reddy et al., 1980; Rogers, 1983).

USE OF OTHER SUPPORTING DATA (SHORT TERM TESTS)

Following a consideration of the applicabili y of short term test data for use in the design of a chronic bioassay protocol, it was determined that this information would provide only limited guidance. The primary uses for these tests at the present time appear to be in assisting the chemical selection process and in providing information to and in clarifying the responses observed in a carcinogenesis bioassay. For the future, the possibilities for their wider utility in the decision process to identify carcinogens has been considered by the Subpanel on Short Term Tests. In particular, the prechronic study can provide a resource of treated animals over a wide dose range that is not now employed to obtain valuable data on a variety of genetic endpoints. The Subgroup on Prechronic Studies concurs with the statements of the Short Term Tests Subgroup regarding the value of the prechronic and chronic studies for validation of the short term assay as a source of material for furtner study, as a means of developing parallel data, and as a possible lead to work in humans.

APPENDIX

1. Acquisition of the Minimum of Pharmacokinetic Data

No single procedure will be appropriate for all compounds because of the wide diversities in physical properties, routes of exposures, analytical sensitivities, and feasibility of radiochemical synthesis. In some cases it may not be feasible to gather pharmacokinetic data at all for technical reasons. Nevertheless, a general approach which will provide useful information is outlined below:

a. An analytical method for quantitating the parent material within a biological matrix should be developed. If sufficient sensitivity cannot be obtained by direct analysis, it may be desirable to use a radioactive form of the test material, coupled with one or more separation techniques to characterize the chemical nature of the radioactivity being measured (i.e., to see if it is still parent material).

b. The test material should be administered to animals of the appropriate species by the intravenous route if feasible. Initially, one sex should be studied. The time course of elimination of parent material from blood should be determined for a dose just beneath that needed to cause demonstrable toxicity and another which is the lowest feasible as determined by analytical sensitivity. It is desirable for the lowest dose for pharmacokinetic study to be in the range of anticipated human exposure for which risk assessment may be attempted. If comparison of the pharmacokinetic parameters for these two doses reveals nonlinearity, additional intermediate doses should be administered to allow estimation of the magnitude of doses which will accomplish 80 to 95% saturation.

c. Animals should be administered the test compound by the proposed route of administration for the chronic study. Again a series of single doses should be employed, spanning the range from near lethality to the level of analytical sensitivity. Generally, these studies may be done in one sex unless there is evidence from toxicity, other studies, or other reason to believe that significant sex differences are likely to occur. If appropriate, comparisons of different routes of administration should be conducted.

d. If radioactive forms of the chemical are available, the major routes of elimination, and the percentage of label eliminated by each route, may be determined for a series of doses in a single sex. Partial characterization of metabolites (e.g., urinary metabolites) might be useful. Elucidation of metabolic pathways is not the objective here; we are looking primarily for evidence of dose-dependency.

Once these data are obtained, the investigator can assess the rates of absorption and elimination of the test chemical. By plotting peak blood concentrations or, better, area under the blood concentration/time curves against administered dose the point at which non-linearity (if present) occurs can be estimated. Evaluation of elimination rates over a range of doses gives indications of saturable or capacity-limited elimination or metabolism of the chemical as well as an indication of potential accumulation with repeated exposure. Before proceeding with dose selection for the chronic study, some additional information is desirable:

a. Selected parameters should be monitored in the other sex of the species to be tested. It is not anticipated that every parameter studied above will be investigated in the other sex.

b. The effects of multiple dosing should be investigated by conducting limited pharmacokinetic studies in animals dosed for approximately two weeks. This time period is a compromise between the need to investigate the potential for changes in pharmacokinetic parameters in repetitively-dosed animals and practical considerations of the difficulty of obtaining such animals. It may be possible to coordinate subchronic toxicity evaluations with pharmacokinetic studies. This allows verification of any predictions about bioaccumulation, as well as providing time for induction/repression of enzyme systems and readjustment of metabolic pools (i.e., glutathione). Often the validity of using pharmacokinetic parameters determined from a single dose for chronically treated animals may be established by determining the steady state blood or tissue levels of the chemical in animals from the subchronic study. Using the pharmacokinetic parameters from the single dose study, the levels attained in tissues of animals used in the subchronic study should be reasonably predictable.

c. For some agents, differentials in pharmacokinetic parameters between trace and toxic doses may be caused by toxicity rather than saturable processes. Such information is important in dose selection; however, whether this differential is attributed to saturation or toxicity may be important in hypothesizing potential mechanisms of toxicity.

2. Acquisition of Additional Valuable Pharmacokinetic Data

Although pharmacokinetic studies may be very sophisticated, the pharmacokinetic studies proposed in the foregoing for use in dose-selection are easily performed and relatively straightforward. They require only minimal analysis of the curves, mostly by visual inspection or multiexponential analysis with standard feathering techniques. They attempt in a general way to answer only two broad questions: Are there non-linearities in the pharmacokinetic behavior, and at what doses or concentrations do these non-linearities occur? With the variety of analytical procedures now available for determining chemicals in blood, the approach should be readily implemented for most test chemicals, and interpretation of the pharmacokinetic studies would not require specialized pharmacokinetic expertise on the part of the chemical managers who will be responsible for defending the rationale for the dose levels selected.

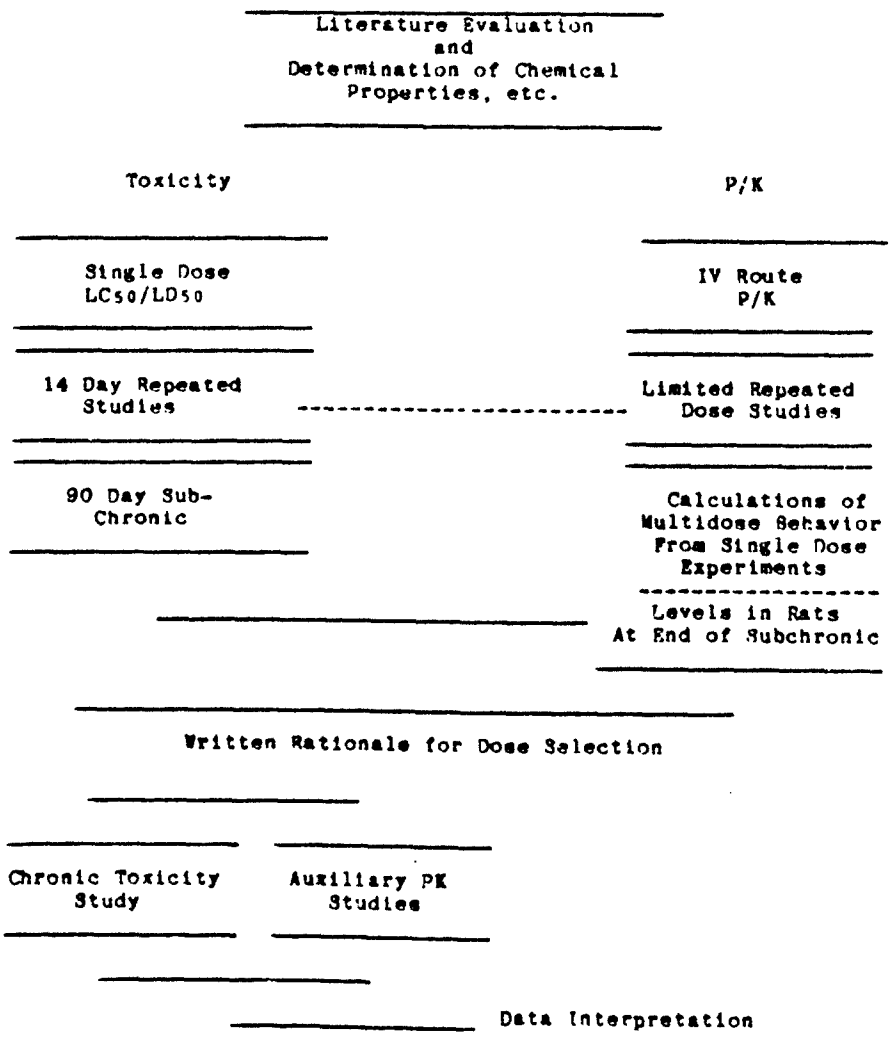
Potential ambiguities may still arise in establishing the non-linear dose region when comparing an experimental route of administration (single dose gavage) with an intended chronic route (chemical mixed with diet). For example, non-linear absorption may be seen for a chemical given via gavage. Similar doses delivered in the feed may result in an entirely linear system since the localized concentration in the gastrointestinal tract is lower even though the daily dose is equivalent. Problems of this type can be evaluated by more detailed pharmacokinetic work which might be conducted during the 90-day subchronic study or concurrently with the chronic toxicity study.

The idealized attainment of pharmacokinetic and toxicity data shown in Table 7 will allow the former to be utilized in selection of doses to be used in the 90-day study. Having basic single dose pharmacokinetic data in hand and using standard equations, the achieved body burdens after multiple doses can be predicted. As indicated previously, reasonable predictability constitutes evidence that pharmacokinetic parameters do not change with repetitive exposure. Other pharmacokinetic studies may also be conducted during the two-year chronic toxicity test to address the question of whether age influences pharmacokinetic

parameters. Other studies of value include (1) metabolite identification and evaluation of their pharmacokinetics, (2) analysis of glutathione depletion or covalent binding, and (3) determination of the effect of diet, disease, or altered physiology on basic pharmacokinetic parameters in test animals. These studies though valuable are incidental to the main objectives in the foregoing proposal. Such sophistication needs to be reserved for those materials warranting the best feasible estimation of risk; i.e., materials for which there is a small difference between human exposure and exposure which causes toxicity in animals.

TABLE 7. DOSE SELECTION/DATA INTERPRETATION

Integration of Toxicity Data and Pharmacokinetics (P/K) Data



Modeling efforts may be pursued by conventional compartmental analysis to understand the disposition of chemical in the test animal and the rate at which chemicals move between the various subcompartments. Physiological modeling with computer simulation is an area with exceptional promise for pharmacokinetic analysis (Bischoff et al., 1971). This approach may help explain the underlying basis of nonlinearities and in some instances the biochemical and physiological factors involved in the nonlinear processes. Ramsey and Andersen (1983) have described the nonlinearity in metabolism of inhaled styrene in the rat in relation to the maximum rate of metabolism and organ blood flow with a physiological model. This approach showed that the inhaled concentration at which nonlinear behavior would be observed should be the same from species to species. This extrapolation of physiological pharmacokinetic results from one mammalian species to another has been called animal scale-up (Dedrick, 1973). The styrene model was also readily adapted to evaluate the enzyme induction seen after repeated exposures of styrene, a phenomenon of potential consequence for long-term studies.

It appears likely that predictive pharmacokinetic models can be developed based on careful measurement of physical and biochemical constants for the test compound. This approach has already been used with anesthetic gases and inhaled gases and vapors (Fiserova-Bergerova and Holaday, 1979). Because physiological factors which control pulmonary absorption of vapors are fairly well understood, a general description of lung function can be obtained which together with physical constants for the vapor becomes predictive of pharmacokinetic behavior by the pulmonary route of administration. The processes of xenobiotic absorption from the gastrointestinal tract and skin uptake across the skin are not as well understood, so a priori predictive models are not yet available for these routes. One prospective use of a pharmacokinetic data base associated with a "national chronic toxicity" program would be to develop an improved understanding of the basic physiological processes which determine the kinetic parameters of absorption, distribution, and elimination. This would enhance the predictive powers of pharmacokinetic models in the future.

Implementation of predictive modeling would reduce the scope of experimentation necessary in pharmacokinetic data collection and reduce somewhat the numbers of animals required for toxicity testing. Limited, 'critical' results in rats, consistent with a general model could support simulation of expected behavior in other proposed test species, and in humans as well. Predicted behavior would be validated by appropriate but limited work in

other species. The ability to predict human kinetics would be extremely useful for risk assessment since it would tell how administered dose and internal target tissue concentrations are related in the species of interest - man. This would remove one of the uncertainties in species extrapolation - i.e., does the human handle the test chemical similarly to the test species? The technology to drive these developments now exists but a commitment by the NTP program to incorporate pharmacokinetic considerations would accelerate progress in this area.

CHOOSING THE DOSES

No matter what objectives are selected for a chronic toxicity evaluation, it is essential to document the dose selection process in detail. This documentation needs to address: What toxic effects were considered in establishing the highest dose? Why were the lower doses selected? Should three or four doses be given in addition to controls? Are there nonlinearities in pharmacokinetics and how did those relate to dose selection? How do the doses selected relate to anticipated human exposure? What results from the study may be anticipated to relate to human risk assessment? Are there any which may not?

To some degree two objectives of a chronic toxicity study may be in conflict - evaluation of data for human risk assessment versus evolution of data for elucidation of toxicity. Albeit infrequent, this conflict occurs when there is a large differential between doses needed to elicit toxicity and potential human exposure excepting abusive use. The first approach for dose selection set forth below assumes that the overwhelming objective is that of obtaining data for human risk assessment.

If human exposures are in the range where linear pharmacokinetics predominate, then the chronic toxicity study must contain at least two dose levels where linear pharmacokinetics predominate if extrapolation is to be possible; however, utilization of less than three doses is to be avoided except in rare, well documented instances. Conversely, if human exposures are near or above the area of nonlinear pharmacokinetics, then the chronic toxicity evaluation should place emphasis upon defining the dose response curve in this area. Dose selection thus is influenced by three variables: (1) pathological observations in subchronic studies; (2) pharmacokinetic studies; and (3) expected human exposure levels.

This approach may be illustrated by reference to Table 8. A test material is placed into one of the categories listed in the table in accordance with these three criteria and dose levels are

selected based on principles appropriate for that category. This approach is a model for how a dose setting regimen may be employed.

TABLE 8. DOSE(S) SELECTION BASED ON THREE CRITERIA

Category	Pathology Top Dose	Non-linear Pharmacokinetics	Human Pharmacokinetics	Human Exposure (Estimated)
A1	X	No	- -	H>.1X
A2	X	No	- -	.1X>H>.001X
A3	X	No	- -	.001X>H
B1	X	S > X	- -	H>.1X
B2	X	S > X	- -	.1X>H>.001X
B3	X	S > X	- -	.001X>H
C1	X	S = X	- -	H>.1X
C2.1	X	S = X	No	.1X>H>.001X
C2.2	X	S = X	Yes	.1X>H>.001X
C3	X	S = X	- -	.001X>H
D1	X	S < X	- -	H>.1X
D2.1	X	S < X	No	.1X>H>.001X
D2.2	X	S < X	Yes	.1X>H>.001X
D3	X	S < X	- -	.001X>H

In Table 8, "X" stands for the highest allowable dose on pathology criteria alone, "S" stands for the approximate dose where nonlinear pharmacokinetic behavior becomes apparent, and "H" stands for the estimated human dose.

When pharmacokinetic studies have not revealed any nonlinear behavior (Category A), or have shown such behavior only at doses well above the maximum judged allowable on pathological criteria (Category B), then pharmacokinetics will play no further role in the selection of doses. Three doses would normally be employed in such chronic toxicity evaluations. Geometric spacing will be employed for selecting these doses; i.e., the doses will be X, X/Y, and X/Y^2. A good rule of thumb in such cases would be to set Y equal to the square root of ten so as to span one order of magnitude in the selected doses.

Depending upon the pathological endpoint used in setting doses, it may be desirable to have an additional dose which is likely to cause a lesser effect than the selected maximum dose, while still retaining two doses anticipated to have no effect. For example, if 3X causes 20-30% increase in liver weight, and X causes approximately 10% increase in liver weight, then four doses might be employed: 3X, X, X/Y and X/Y^2.

The value of Y will be influenced to some degree by the steepness of the pathology dose response curve and the results of in vitro tests of genotoxicity. A shallow pathology dose response curve would suggest a larger value of Y than a steep dose response curve (increased dose spacing) while steep dose response curves would have the opposite effect. Similarly, indication of potential genotoxicity in short term tests would reduce the value of Y and increase the range of doses selected for the chronic toxicity test.

In Category C, pharmacokinetic nonlinearities occur in the same region of doses as those identified by pathological criteria. In such a case, categories C1 and C2.1 would be treated differently from C3 and C2.2.

Materials in category C1 have an expected human exposure close to that identified by pathological criteria (greater than or equal to 0.1X). Materials in category C2.1 have an estimated human exposure between 0.1X and 0.001X and lack human pharmacokinetic data to indicate that the test animals handle the material like man. In these cases, the top dose should be selected so as to produce an estimated 80 to 95% saturation, assuming the pathology associated with such a dose is not likely to be life shortening. Two other doses will be selected as follows: (1) Middle dose to produce an estimated 15 to 30% saturation as determined from AVC data (often an 0.3 log unit of the MTD will be suitable), (2) Lowest dose at least a half log unit below the middle dose. If it is anticipated that an even lower dose will be required to attain a no-effect level, additional logarithmically spaced doses should be included.

Materials in category C2.2 are those which have an expected human exposure level of between 0.1X and 0.001X with supporting human pharmacokinetic data to suggest that the test animals handle the material like man, and materials in category C3 are those where human exposures are less than 0.001X. For these, the highest dose is chosen to produce 15 to 30% saturation, with at least two geometrically spaced lower doses.

Materials are placed in category D when nonlinear pharmacokinetics can be demonstrated at doses well below those identified

by pathological criteria. In such a case, four dose levels are selected for categories D1 and D2.1. The highest dose level employed is X, and the next three dose levels are chosen so as to provide: (1) One dose level causing an estimated 80 to 95% of saturation, (2) one dose level causing an estimated 15 to 30% saturation, and (3) one dose at least one half log unit below that estimated to be 15 to 30% of the saturation level. This regimen will provide data which may be meaningfully extrapolated if chronic toxicity is incurred.

Materials in category D3 would be tested with three doses. The top dose would be selected so as to give 70-80% saturation with two additional doses selected as for C3.

Materials in category D2.2 may be tested as designated for materials in category D3 if human pharmacokinetic data are strong and consistent with the absence of saturation at anticipated human exposures and exposure is near the 0.001X value rather than 0.1X. Otherwise, chronic toxicity evaluations will normally require four doses as outlined for materials in category D1 and D2.1.

If the primary objective of a chronic toxicity evaluation is to elucidate toxicity regardless of its meaningfulness for assessment of human risk, the foregoing approach remains applicable except for category D. In this case, pathological criteria will be used solely for selection of the top dose, regardless of its relationship to known, allowed, or anticipated human exposure. Materials falling into category D will be evaluated with a minimum of four doses; the top dose based on pathological criteria, a dose equal to or greater than that needed to cause 80 to 95% saturation, a dose estimated to give 15 to 30% saturation and a dose at least one half log unit below the latter.

The foregoing guidelines for dose selection are not intended to address all of the possible combinations of pathological and pharmacokinetic criteria which may be useful in dose selection. Only with sound, considered, scientific judgment may doses be selected for chronic toxicity evaluations which will ultimately produce a data base which can be meaningfully extrapolated for human risk assessment. Historically, such judgment has all too often been attempted only after the chronic toxicity study is completed and at this point even sound judgment is too often precluded by results which have a poor relationship to human risk assessment. The most important scientific task in conducting chronic studies is that of designing the study and clearly documenting the rationale used in its design.

AUTHORS OF APPENDIX

Dr. Melvin E. Andersen
AFAMRL/THB
Wright-Patterson AFB, OH 45433

Dr. George D. Divincenzo
Health, Safety & Human Facility Lab
Eastman Kodak Company
Rochester, NY 14650

Dr. Perry J. Gehring
Vice President of Agricultural
Products R & D and Director of
Health & Environmental Science
Dow Chemical Company U.S.A.
Midland, MI 48640

Dr. Robert R. Maronpot
NIEHS
P. O. Box 27709
Research Triangle Park, NC 27709

Dr. John C. Ramsey
Health & Environmental Sciences
Dow Chemical U.S.A.
1803 Building
Midland, MI 48640

Dr. Richard H. Reitz
Health & Environmental Sciences
Dow Chemical U.S.A.
1803 Building
Midland, MI 48640

REFERENCES

Brischoff, K. B., R. L. Dedrick, D. S. Zaharko, and J. A. Longstreth (1971), Methotrexate pharmacokinetics, J. Pharm. Sci., 60:1128-1133.

Dayton, P. G. and J. E. Sanders (1983), Dose-dependent pharmacokinetics: emphasis on phase 1 metabolism, Drug Metab. Rev., pp. 347-405.

Dedrick, R. L. (1973), Animal scale-up, J. Pharmacokin. Biopharm., 1:435-461.

Department of Health and Social Security (1982), Guidelines for the testing of chemicals for carcinogenicity, Report on Health and Social Subjects by the Committee on Carcinogenicity of Chemicals in Food, Consumer Products and the Environment, London, United Kingdom.

Eustis, S. L. and G. Boorman (1983), Proliferative lesions of the exocrine pancreas: relationship to corn oil gavage in the NTP carcinogenesis testing program, In: 1983 NTP Working Paper.

Festing, M. F. W. (1979), Inbred strains and the factorial experimental design in toxicological screening, ICLA Symposium, Utrecht, Holland.

Fiserova-Bergerova, V. and D. A. Holaday (1979), Uptake and clearance of inhalation anesthetics in man, Drug Metab. Rev., 9:43-60.

Food Safety Council (1980), Proposed system for food safety assessment, Final Report of the Scientific Committee of the Food Safety Council, Washington, DC.

Gehring, P. J., P. G. Watanabe, and G. E. Blau (1976), Pharmacokinetic studies in evaluation of the toxicological and environmental hazard of chemicals, In: New Concepts in Safety Evaluation, Eds. M. A. Mehlman, R.E. Shapiro, H. Blumenthal, pp. 195-270, Wiley & Sons, New York.

Gibaldi, M. and D. Pervier (1982), Pharmacokinetics, 2nd ed., Marcel Dekker, New York.

Grice, H. C., D. L. Arnold, H. Blumenthal, J. L. Emmerson, and D. Krewski (1984), The selection of doses in chronic toxicity/carcinogenicity studies, In: Current Issues in Toxicology, Ed. H. C. Grice, pp. 6-49, Springer Verlag, New York, pp. 6-49.

Grasso, P. and J. Hardy (1975), Strain difference in natural incidence and response to carcinogenesis, In: Mouse Hepatic Neoplasia, Eds. W. H. Putler, P. M. Newberne, Elsevier Press, Amsterdam, pp. 111-132.

Haseman, J. K. and D. G. Hoel (1979), Statistical design of toxicity assays: role of genetic structure of test animal population, J. Tox. Environ. Hlth., 5:89-101.

Haseman, J. K. (1983), Patterns of tumor incidence in two-year cancer bioassay feeding studies in Fischer 344 rats, Fund. Appl. Toxicol., 3:1-9.

Iritani, N. and Y. Ikeda (1982), Activation of catalase and other enzymes by corn oil intake, J. Nutr., 112:2235-2239.

Laskin S. and A. Sellakumar (1974), Models in chemical respiratory carcinogenesis, In: Experimental Lung Cancer: Carcinogenesis and Bioassays, Eds. E. Karbe, J. F. Park, Springer-Verlag, New York.

Levy, G. (1968), Dose-dependent effects in pharmacokinetics, In: Importance of Fundamental Principles in Drug Evaluation, Eds. D. H. Tedeschi, R. E. Tedeschi, pp. 141-172, Raven Press, New York.

Maronpot, R. R. and G. A. Boorman (1982), Interpretation of rodent hepatocellular proliferative alterations and hepatocellular tumors in chemical safety assessment, Toxicolog. Pathol., 10:71-79.

Mattsson, J. L. and C. N. Park (1982), Fundam. Appl. Toxicol., 2:16a.

National Toxicology Program (1983), Annual Plan for Fiscal Year 1983, Research Triangle Park, NC.

Newberne, P. M. (1982), Assessment of the hepatocarcinogenic potential of chemicals: response of the liver, In: Toxicology of the Liver, Eds. G. Plaa, W. R. Hewitt, Raven Press, New York.

Newberne, P. M., J. Weigert, and N. Kula (1979), Effects of dietary fat on hepatic mixed-function oxidases and hepatocellular carcinoma induced by aflatoxin B_1 in rats, Cancer Res., 39:3986-3991.

Newberne, P. M., J. L. V. deCamargo, and A. J. Clark (1982), Choline deficiency, partial hepatectomy, and liver tumors in rats and mice, Toxicolog. Pathol., 10:95-109.

The Nutrition Foundation (1983), Report of the ad hoc working group on oil gavage in toxicology, July 14-15, 1983, Arlington, VA.

The Nutrition Foundation (1983), The relevance of mouse liver hepatoma to carcinogenic risk, Report of the international expert advisory committee, September, 1983.

O'Flaherty, E. J. (1981), Toxicants and drugs: kinetics and dynamics, Wiley & Sons, New York.

Organization for Economic Cooperation and Development (1981), "OECD Guidelines for Testing of Chemicals," Paris, France.

Page, N. P. (1977), Concepts of a bioassay program in environmental carcinogens," In: Advances in Modern Toxicology, Vol. 3, eds. H. F. Kraybill, M. A. Mehlman, Hemisphere Publ., Washington, D.C.

Ramsey, J. C. and M. E. Andersen (1983), A physiological based description of the inhalation pharmacokinetics of styrene in rats and humans, Toxicol. Appl. Pharmacol., (submitted).

Reddy, B. S., L. A. Cohen, G. D. McCoy, P. Hill, J. H. Weisburger, and E. L. Wynder (1980), Nutrition and its relationship to cancer, In: Advances in Cancer Research, Vol. 32, Academic Press, Inc.

Rogers, A. E. (1983), Influence of dietary content of lipids and lipotropic nutrients on chemical carcinogenesis in rats, Cancer Res., Suppl., 43:2477s-2484s.

Rogers, A. W. and P. M. Newberne (1980), Lipotrope deficiency in experimental carcinogenesis, Nutr. Cancer, 2:104.

Schwetz, B. A. (1983), Prechronic toxicological tests - scope and use, In: NTP working paper.

Shimkin, M. and V. A. Triolo (1969), History of chemical carcinogenesis: some prospective remarks, In: Progr. Exp. Tumor Res., Vol. 11, pp. 10-20, Karger, Basel/New York.

Shimkin, B. B. (1974), Section 2: Report of discussion group no. 1; species and strain selection, In: Carcinogenesis Testing of Chemicals, From the Proceedings of the Conference on Carcinogenesis Testing in the Development of New Drugs, May 23-25, 1973, Washington, D.C.

Society of Toxicologic Pathologists, Proceedings of the International Symposium of the Society of Toxicologic Pathologists: Rodent liver nodules; significance to human cancer risk, Toxicologic Pathol., Vol. 10, Number 2, 1982.

Solleveld, H. A., J. Haseman, and E. E. McConnell (1983), The natural history of body weight gain, survival and neoplasia in the Fischer 344 rat, National Toxicology Program, Research Triangle Park, NC (in press).

Sontag, J. M., N. P. Page, and U. Saffiotti (1976), Guidelines for carcinogen bioassay in small rodents, Report to the carcinogenesis program, division of cancer cause and prevention, National Cancer Institute, Bethesda, MD.

Squire, R. A. and M. H. Levitt (1975), Report of a workshop on classification of specific hepatocellular lesions in rats, Cancer Res., 35:3214-3223.

Tsuchiya, T. and G. Levy (1972), Relationship between dose and plateau level of drugs eliminated by parallel first-order and capacity-limited kinetics, J. Pharm. Sci., 61:541-544.

U. S. Environmental Protection Agency (1982), Health effects test guidelines and support documents, Washington, D.C.

U. S. Environmental Protection Agency (1982), Pesticide assessment guidelines, Sub-division F, Hazard evaluation: human and domestic animals, Washington, D.C.

U. S. Food and Drug Administration (1981), Staff Manual Guide, IND/NDA Pharmacology Review Guidelines, Washington, D.C.

U. S. Food and Drug Administration (1982), Toxicological principles for the safety assessment of direct food additives and color additives used in food, Washington, D.C.

Vesselinovitch, S. D. (1982), Liver tumor induction, Toxicol. Pathol., 10:110-118.

von Wittenau, M. S. and P. C. Estes (1983), The redundancy of mouse carcinogenicity bioassays, Fundam. Appl. Toxicol., (in press).

Wade, A. E. and W. P. Norred (1976), Effect of dietary lipid on drug-metabolizing enzymes, Fed. Proc., 35:2475-2479.

Ward, J. M., R. A. Griesemer, and E. K. Weisburger (1979), Toxicol. Appl. Pharmacol., 51:389-397.

Weil, C. S. and D. D. McCollistor (1963), Relationship between short- and long-term feeding studies in designing an effective toxicity test, J. Agr. Food Chem., 11:486-491.

Weisburger, E. K. (1983), History of the bioassay program of the national cancer institute, Prog. Exp. Tumor Res., 26:187-201.

Weisburger, J. H. and E. K. Weisburger (1967), Tests for chemical carcinogens, In: Methods in Cancer Research, Ed. H. Busch, Academic Press, New York.

Williams, G. M., K. Purukawa, Y. Maeura, and T. Croci (1983), Induction of hepatic enzymes by corn oil gavage, Science, (in press).