DRAFT

Background Document on the Development and Use of Reference Doses

Part I: Data Needs and Apportionment

Prepared for:

Office Of Solid Waste
U.S. Environmental Protection Agency
Washington, D.C.

Prepared by:

ENVIRON Corporation
1000 Potomac Street, N.W.
Washington, D.C. 20007

December 20, 1985

Table of Contents

1. Introduction --- Purpose and Organization

The Office of Solid Waste (OSW) of EPA is proposing certain restrictions on the land disposal of hazardous wastes. The principal concern of these restrictions is the problem of long-term, low-level release of hazardous chemicals from land disposal sites that may arise because of the deterioration of containment systems. To ensure protection of human health, OSW proposes to place limits on the extent of air or water contamination that may result from any such releases.

Limits are to be proposed for individual chemicals to protect humans from the possible adverse effects of repeated, low level exposure (chronic exposure). (The Agency has already promulgated regulations dealing with single or infrequent, high-level exposures that may arise because of accidents.) The two principal determinants of these limits are:

(1) for substances not known to display carcinogenic properties, the acceptable daily intake (ADI), hereafter to be referred to by EPA as the Reference Dose (RfD).

(2) for substances known to display carcinogenic properties, the lifetime average daily dose corresponding to a specific level of excess lifetime cancer risk, hereafter, the Risk-Specific Dose (RSD).

These two determinants are well-established and widely-accepted health protection criteria. They satisfy the goal of protecting humans from chronic exposures to chemicals that may be released from various sources, to the extent current scientific knowledge can allow (NRC, 1980; 1983). EPA is abandoning the use of the term "Acceptable Daily Intake", because it may be read to imply that doses in excess of it are necessarily "unacceptable." As will be seen, this is an incorrect interpretation, and the Agency believes use

of the more neutral term "Reference Dose" avoids this difficulty. The principles upon which RfDs are based, and the data used to derive them, are identical to those traditionally used to derive ADIs.

The term "Risk Specific Dose" has not previously been used; but it is simply a convenient way of identifying the dose of a carcinogen corresponding to a specified level of lifetime risk.

EPA is relying on a number of expert scientific reviews and agency documents to support use of these two health protection criteria. But there are several aspects of the proposal that require review and analysis not found in any existing documents. EPA thus asked ENVIRON to prepare such a review, focusing on the following principal issues:

(1) EPA proposes to develop toxicity data on chemicals for which limited or no data are currently available. It is thus necessary to assess available test methodologies and to identify those suitable for developing data from which RDs can be established. (No similar review is needed for RSDs, which are developed from carcinogenesis bioassay data.)

(2) Chemicals released from waste sites may enter both air and water, creating two possible routes of human exposure. In addition, chemicals found at waste sites may also be present in other media (e.g., a pesticide that is also present in the diet). It is thus necessary to decide whether and how to <u>apportion</u> RfDs or RSDs among the several possible human exposure media.

In addition to these two major issues, a number of ancillary points arise in the approach proposed by EPA. These include: 1) the scientific basis for the RfD as a protective device; 2) methodology for deriving RfDs from various types of toxicity data; 3) the accuracy and precision of RfDs; 4) the development of RfDs for certain metals that are also essential nutrients (e.g., copper, selenium, chromium); and 5) the problem of interactions among

chemicals. Although ENVIRON's report is organized around the two major issues, these additional points will be included in the discussion.

In the next section, a broad introduction to the concept of the RfD is provided. This is followed by a discussion of the types of toxicity data from which RfDs can be established, along with a presentation of the strengths and limitations of various types of data. The purpose of this section is to identify the types of data believed necessary to develop a reliable RfD.

The report then moves to a discussion of establishing RfDs from various types of data and of the several ancillary issues relating to RfDs described earlier. We then examine the apportionment issue, and describe the options available to EPA and the strengths and weaknesses of each. All of these issues are presented as Part I of this report.

Because all of the published scientific literature pertaining to RfDs refers to ADIs, we retain the latter term in the following discussion, even when we refer to EPA's own earlier literature. It should be noted that EPA has altered only the label attached to the term, and has not altered its underlying basis. It is for this reason that all of the information relating to ADIs is directly relevant to EPA's proposed development and use of RfDs.

A discussion of the various considerations influencing the design of protocols for toxicity testing, provided to guide identification of the most cost-effective means to collect toxicity data, is presented in Part II of this report.

2. Introduction to the Concept of Acceptable Daily Intake (ADI)

2.1 Origins

According to a Committee of the National Academy of Sciences

> "The acceptable daily intake (ADI) of a chemical is defined as the dose that is anticipated to be without lifetime risk to humans when taken daily. It is not assumed that this dose guarantees absolute safety (NRC, 1980)."

This definition is essentially the same as that given by the World Health Organization (FAO/WHO, 1958; 1965), the EPA (see, e.g., various Health Effects Documents), and the FDA (FDA, 1982).

Experimental data on toxicity is typically collected in small groups of experimental animals at doses sufficiently high to produce directly observable forms of toxicity. Such experimental studies can reveal the dose-effect relation for the chemical, as well as the maximum dose* at which toxicity is not observed (termed the no-observed-effect level, NOEL).

Faced with this type of data for several food additives, Lehman and Fitzhugh (1954) proposed that ADIs could be established by dividing the experimental NOEL by a "safety factor." These authors (who were FDA officials) cited acute toxicity data suggesting that, for some substances,

---

* Because of practical limitations on the number of dose levels used in an experiment, it is usually not possible to identify the true maximum NOEL. The measured NOEL is, in many cases, less than the true maximum NOEL.

small groups of relatively homogeneous experimental animals were ca. 10-fold less sensitive to their toxic effects than were members of the general human population, and then reasoned that the variability in response expected among members of the human population might make some members ca. 10-fold more sensitive than the "average". These notions, coupled with the long-standing idea that chemical toxicity did not become manifest until the dose exceeded a threshold value, led Lehman and Fitzhugh (and the FDA) to the conclusion that they could estimate a human population ADI by dividing the chronic, experimental NOEL by a "safety factor" of 100.

The FDA recognized that the ADI was not a guarantee of absolute safety. They also recognized that human exposures for many substances might well exceed the ADI by some (undefinable) amount for extended periods without resulting in human chronic toxicity. That is, it was recognized by the original developers of the ADI that the figure was only an estimate based on incomplete knowledge, and that it should not be considered a sharp dividing line between "safe" and "unsafe" chronic exposures (Lehman and Fitzhugh, 1954; FDA, 1982; Rodricks and Taylor, 1983). Instead, the "NOEL-safety factor" approach is a practical device for deriving acceptable exposure levels, for various regulatory and public health purposes, in the face of limited scientific information and knowledge.

FDA has also derived chronic ADIs for substances for which chronic (i.e, lifetime) toxicity data were not available. When, for example, the only data available for a substance revealed the effects of subchronic exposure (e.g., 90-day exposure studies in rodents), FDA incorporated an additional 10-fold safety factor to derive an ADI. Thus, the NOEL from subchronic studies was

divided by 1000 to establish the chronic ADI. The agency (and other investigators as well, see below) justified this practice on the principle that, with the exception of carcinogenicity, there is very high confidence that all of the major toxic effects of a chemical can be found in carefully designed subchronic studies, and that chronic studies would merely extend the dose-response curve (by ca. 10-fold) for the effects observed after subchronic exposure. (Additional, detailed discussion of this point is presented in Section 3, below). The additional 10-fold safety factor was thus used as a substitute for the dose-response data that would be obtainable at the lower doses used in chronic experiments.

Scientists associated with other national and international organizations have also adopted the concept of ADI as a health protection device. Scientists associated with the World Health Organization (WHO) and the Food and Agricultural Organization (FAO) have further justified a 100-fold safety factor for food additives based on differences among species in body size, food requirements, water balance exchange, and variations in susceptibility to the toxic effect. This rationale and approach were also accepted by the FAO/WHO Expert Committee for Pesticide Residues (FAO/WHO, 1965).

A committee of the National Academy of Sciences (NRC, 1977) estimated ADIs for contaminants in drinking water using an approach similar to that of FDA, but used "uncertainty" (rather than "safety") factors to account for the limitations in the data base and in our knowledge of inter- and intra-species variability in response.

2.2 Use at EPA

Several uncertainty factors have been used to estimate ADIs depending on the type and quality of available human or animal toxicity data. At EPA, the magnitude of the chosen uncertainty factors depends on the differences between the human exposure characteristics and the conditions of the experimental studies used to derive the ADIs (now, RfDs). Further, if the no-observed-effect level (NOEL) is sufficiently close to the ambient exposure level, and there is no evidence of adverse effects at these levels, then relatively small uncertainty factors have been used. Also, detailed knowledge of a chemical's mechanism of toxicity, critical effect, and pharmacokinetic behavior in humans and experimental animals may permit modification of the standard (generic) uncertainty factors for some substances. Such information is, however, seldom available to influence estimation of the ADI.

An uncertainty factor of 10 is used by EPA to estimate ADIs from appropriate human data; its purpose is to account for intraspecies variability in response to the adverse effects of a chemical. An uncertainty factor of 100 is used with relevant (with regard to duration and route of exposure) animal data from properly conducted chronic studies; this factor accounts for both intra- and inter-species variability. If only marginal data are available (e.g., data from subchronic studies in animals), an uncertainty factor of 1000 is used; this figure incorporates the uncertainty in extrapolating from one duration of exposure to another and also accounts for intra- and inter-species variability. This approach essentially matches that of FDA, WHO, and the NRC.

Additional uncertainty factors have been used to compensate for other short-comings in experimental information. These additional uncertainty factors were incorporated when the only data available revealed a lowest-observed-effect level (LOEL) rather than a NOEL, when subchronic data were used to project potential chronic effects for humans, or when there were other deficiencies in the data base upon which decisions had to be made. Recently, Dourson and Stara (1983) demonstrated that some of these additional factors (which typically range from two to ten) have some experimental support and are likely to be highly protective for many chemical substances.

ADI's have been developed by EPA's Environmental Criteria and Assessment Office (ECAO) (EPA, 1982; 1984) and EPA's Office of Pesticide Programs (OPP). The following guidelines for deriving ADIs from toxicity data have been adopted by some groups at EPA (EPA, 1980a).

- Doses associated with an increase in frank toxic effects, such as mortality or convulsions, are not suitable for derivation of an ADI.
- A free-standing NOEL is unsuitable for derivation of an ADI. If multiple NOELs of equal quality are available without additional data on LOELs, NOAELs, or LOAELs, the highest NOEL shoud be used to derive an ADI.*
- A NOAEL, LOEL or LOAEL can be suitable for an ADI derivation. A well-defined NOAEL from a chronic or subchronic study can be used directly, applying the appropriate uncertainty factor, and is preferred. For a LOEL, a judgment must be made as to whether it actually corresponds to a NOEL or a LOAEL. In the case of a LOAEL, an additional uncertainty factor is applied; the magnitude of the additional uncertainty factor is not to substitute levels for which severely adverse effects are seen. (For some groups at EPA, no differentiation is made between NOEL and NOAEL or between LOEL and LOAEL.)

---

* The NOAEL and LOAEL include the additional adjective "adverse"; in many cases an effect may be observed at a given dose, but may not be adverse to health. If this is the only effect observed, the dose may be labeled NOAEL.

- If -- for reasonably closely spaced doses - only a NOEL and a LOEL of equal quality are available, the appropriate uncertainty factor is applied to the NOEL.

As an additional general rule, EPA does not ordinarily consider it appropriate to use recommended occupational exposure levels, such as Threshold Limit Values (TLVs), for directly deriving ADIs; nevertheless, the data bases which were used to derive TLVs or other occupational exposure levels may be appropriate for use in deriving an ADI. In some instances the TLV may be directly useful, if its derivation was based on the same general principles used to devise ADIs.

2.3 Media-Specific Limits

For many chemicals, human exposure may occur through several media (air, water, food, direct soil contact). In such cases it is important that the total exposure from all media not exceed the ADI. Human intake or contact with various media must thus be taken into account when estimating the maximum level that can be tolerated in each medium without the ADI being exceeded for any individual. Several methods have been developed to deal with this issue, and they will be discussed in Section 5, on apportionment.

2.4 Utility and Limitations of the ADI

Since its introduction the ADI has been widely used as a practical, health-protection device. For this reason, it appears to be the appropriate criterion for establishing limits for substances migrating into air or water from hazardous waste sites. While no better means for accomplishing EPA's health protection objectives is available, the limitations in the ADI should be recognized. The following list of limitations and other characteristics of the

ADI has been assembled from all of the various sources cited in the foregoing discussion.

(i) The ADI does not represent a sharp dividing line between "safe" and "unsafe" exposures.

(ii) There is no readily definable way to estimate the magnitude of the uncertainty in any given ADI. Uncertainties arise from many sources, and are related to the quality and completeness of the toxicity data, the uncertainty in the dose-response information and the NOEL, and the lack of chemical-specific data on intra- and inter-species variability in response. There is no means available to quantify accurately the accuracy and precision of ADIs.

(iii) The safety (or uncertainty) factors in common use are probably overprotective for some substances, because they have been selected to protect against "worst-case" substances. It is possible, however, that in a few cases they are underprotective.

(iv) Brief excursions above the ADI can probably be tolerated without harmful effect by all members of the general population. There is no precise definition of "brief".

(v) Consistent protection at or near the ADI ensures that individuals will be protected from the acute effects of all chemicals. Protection against acute toxic effects usually requires safety factors no larger (and sometimes smaller) then those used for establishing the ADI. Because the chronic NOEL will always be a lower dose that the minimum effective acutely toxic dose, the ADI will clearly protect against acute toxicity.

(vi) ADIs have not been established or used for carcinogens, on the ground that threshold doses may not exist for this class of agent (i.e., carcinogens pose a finite risk at all finite doses, with the risk increasing with dose).

These characteristics and limitations of ADIs apply, of course, to RfDs, and should be kept in mind when using them or explaining their basis. Several will be discussed more fully in the Sections to follow.

3. Minimum Data Needs for Establishing ADIs

3.1 Introduction

Some toxicity data are available from controlled studies in exposed humans, but these are typically limited to short-term exposures, producing rapidly reversible effects. Epidemiological data are available for a relatively large number of important chemicals, but such data usually lack quantitative dose-response information and, in many cases, are ambiguous with respect to the issue of causation. In all but a few cases, toxicity data from human studies are either not available or inadequate to establish ADIs (NRC, 1980; EPA, 1984). Because of this it has become necessary to rely upon data from studies in experimental animals. In this section we describe the various types of experimental tests available for collecting toxicity data and the types of information provided by each of the various tests. We also describe the limitations in each of the types (i.e., what they cannot reveal). The discussion is limited to types of toxicity tests that have been sufficiently validated for use in regulatory standard-setting.

The purpose of this discussion is to identify the minimum amount and type of data necessary to establish an ADI for chronic human exposure. It should be noted that identifying minimum data requirements is not a strictly scientific undertaking, because it is possible, as a policy matter, to use safety or uncertainty factors to compensate for almost any kind of data gap (and some degree of scientific support can probably be found for such selections, see ENVIRON, 1985). While we adhere to these general concepts, we point out in the closing evaluation subsection what appears to be the current

consensus in the scientific and regulatory policy communities regarding minimally acceptable data for establishing ADIs.

3.2 Utility and Limitations of Various Types of Toxicity Tests

3.2.1 Basic Concepts

A fundamental assumption in the estimation of human health risks posed by chemicals is the ability to extrapolate animal test results to predict human response. This is the cornerstone of most regulatory decisions regarding the safety of substances in the environment, in the food supply, or in drugs. This assumption, however, is not based on complete certainty regarding the predictive power of animal models. Rather, it is based on the widely-accepted view that well designed animal studies provide an indication of potential human toxicity and that the strength of the indication is a function of the rigor, completeness, and reproducibility of the test animal studies. This section contains an examination of the various standard toxicity studies currently accepted by various regulatory and public health agencies and a summary of the information each provides (and does not provide) about potential human toxicity. This review will proceed from least to most complex test type, and emphasizes the confidence that can be placed in the results of each type of test. It will be made clear what types of effects might not be detected at each level of testing, and what uncertainty would remain if testing were to cease at a given level. A tabular summary of this information will be provided, with some estimate of the cost of moving from one level of testing to the next. The cost figures allow some judgment regarding the value of obtaining new information, or of the cost-effectiveness of each test type.

Required or widely accepted general toxicity tests are classified as acute, subchronic, and chronic. These tests are "general" because they are designed to identify the full range of toxicities associated with a chemical. Additional tests, however, have been found necessary to identify specific effects not readily observable in the general toxicity studies. Tests for reproductive injury, teratogenicity, and genetic effects are among those widely used for such purposes. Cancer bioassays are generally considered a subcategory of the chronic test. There are tests available for other specific endpoints (e.g., behavioral and neurological injury and adverse effects on the immune system) that have not yet been widely accepted as valid indicators of human toxic potential - i.e., they are still in the developmental state. Metabolism and pharmacokinetic studies are becoming increasingly important components of a toxicity profile.

The specific protocols for each of these types of tests vary somewhat among agencies, but they nevertheless provide the same basic information.

In the ideal, it would seem that determination of an ADI for a chemical to which humans would be chronically exposed would require chronic animal test data in several species. These studies are used to identify the range of chronically toxic doses, and to establish the NOEL or NOAEL. Fiscal, manpower, and legal constraints often require that the "ideal" level of testing be adjusted to a more realistic level. Most regulatory agencies have, therefore, adjusted the level of testing required for a compound according to the magnitude of expected human exposure and the outcome of previous, less-than-lifetime toxicity studies. Current toxicity testing strategies are, therefore, hierarchical sequences of tests designed to develop a profile of a

substance's toxicity. The hierarchy generally consists of several levels, or tiers, of tests. The lowest or initial tier consists of relatively rapid, inexpensive tests intended to identify the acute toxicity of the compound. Often included in the initial tier are short-term and genetic toxicity tests that rapidly provide information about potential carcinogenic effects. This first tier information, although not directly useful in predicting chronic adverse effects in humans, can be used to guide decisions about the need for and type of more extensive testing. The second tier of testing may include subchronic, whole-animal types of tests that require 1-3 months to conduct. Later tiers are intended to yield direct information on the chronic toxicity of the substance, and on its effects on reproduction or development (FDA, 1982).

3.2.2 Acute Toxicity Studies

Acute toxicity studies are used to provide an estimate of the adverse effects that would be associated with a single exposure to a chemical. In addition, they provide an estimate of the relative susceptibilities of various species and sexes, identify target organs, suggest mechanisms of action, and assist in selection of dose levels to be used in longer-term studies. The most common measure of acute toxicity is the median lethal dose (LD_{50} or LC_{50}). It should be emphasized that estimation of an LD_{50} is not equivalent to describing the acute toxicity of a compound. A well designed acute toxicity study will also include consideration of non-lethal parameters of morbidity or pathogenesis.

In general, a battery of acute exposure studies is usually used to describe the acute toxicity of a compound by several routes of administration.

These include tests by the oral, dermal, and inhalation routes; skin and eye irritation studies are also considered at this phase of testing. These data are necessary to protect workers and others who may be exposed during production, transport, use, and disposal of a chemical. For the toxicologist, acute studies can provide information on the possible mechanism of action of the compound, its target organs, the reversibility of the effects, and on structure-activity relationships. Such information also assists, indeed is necessary, in the design of longer-term studies.

It should be emphasized that acute studies do not provide any information about the cumulative effects from subchronic or chronic exposure to a compound, reproductive effects, teratologic effects, or carcinogenic effects. In other words, they reveal nothing about the nature of toxicity that will arise after repeated exposures.

McNamara (1976) examined a series of 122 non-carcinogenic compounds and concluded that the LD_{50} could be divided by 100 to estimate a subchronic NOEL or by 1,000 to estimate a chronic NOEL. Furthermore, in a report prepared by ENVIRON (1985), it was shown that for 85 non-carcinogenic compounds, the LD_{50} could be divided by 119 to estimate a subchronic NOEL and 3,120 to estimate a chronic NOEL. Use of such safety factors to develop a chronic or subchronic NOEL from LD_{50} data, based on these type of empirical analyses, has not been accepted by the general toxicology community and is thus not yet considered a reliable method for estimating ADIs, at least for deriving health-protection limits. An ADI derived on the basis of these types of empirical observations, which concern only quantitative factors and which are not based on any real knowledge of toxicity, would seem to be useful only

in cases where decisions clearly have to be made before more appropriate data could be collected.

3.2.3 Subchronic Toxicity Studies

Subchronic toxicity studies are used to determine the toxic effects that occur from repeated exposure for various fractions of an animal's life span, and to identify the NOEL for these effects. Such studies provide information about target organs, physiologic and metabolic capacity of the animal to tolerate prolonged exposure, and cumulative toxicity.

An important component of subchronic studies is the use of a broad screen of measures which will detect most forms of toxicity. These include daily behavioral observations, periodic physical examinations, body weight and food consumption monitoring, analysis of hematologic parameters, and clinical screening of blood and urine. Of most importance is the conduct of gross and histopathologic examinations of animals, and collection of organ weight data at sacrifice.

The period of exposure for a subchronic study is dependent on the species of animal used and how the study will be used. In general, rodents are maintained on test for 3 months while longer lived animals, such as dogs and monkeys, for one year or more. If the subchronic study is being used as a range finding study for selection of doses that will be administered in reproductive, chronic, or carcinogenesis studies, then a one-month exposure period is probably adequate for most compounds.

Well designed and conducted subchronic studies have been found to be reliable predictors of most forms of toxicity except for carcinogenic, teratogenic, or reproductive effects. The FDA (1982), in its toxicological principles for safety assessment, suggests that if a compound tested in a subchronic study is found to cause focal hyperplasia, metaplasia, proliferative lesions, or necrosis, then a carcinogenicity study in two rodent species is indicated. Finally, if a subchronic study indicates reproductive organ toxicity, then a two-generation reproduction study with a tetratology phase may be appropriated. This type of approach (which may be most appropriate for substances requiring premarket approval) implies that only under limited circumstances is chronic or reproductive toxicity data necessary.

Because of the enormous cost of conducting chronic studies, several authors have examined the question of what additional information is gained by extending the subchronic study. Weill and McCollister (1963) compared the results of 90-day studies to those obtained in 2-year studies for 33 chemicals tested in rats. Only body weight gain, relative weight changes of the liver and kidney, and liver and kidney pathology were monitored. They found that, for 95% of the studies, the 90-day maximum effect level was only 6 times larger than the 2 year maximum effect level. Peck (1968) examined eleven drugs tested for periods of up to 2 years and found that only one study showed additional new forms of toxicity after 3 months and only four showed additional new toxicity after 2 months. The author supported the use of 3-6 month studies for detecting long term effects.

McNamara (1976) examined data on 122 compounds for which subchronic and chronic studies were available. Of these, only 2.5% produced previously

unnoted toxic effects after 3 months of exposure. For another 6.5% of the compounds, effects were found in less than 3 months at the highest dose, but effects at a lower dose were then seen after 3 months. In almost all cases, new toxicities, not found before 3 months of exposure, did not appear after longer periods of exposure.

McNamara also estimated the relationship between the chronic, and subchronic NOELs. He concluded that, for 95% of the chemicals, the subchronic NOEL will be no more than ten times larger than the chronic NOEL. McNamara concluded that the 90 day subchronic NOEL could reliably predict the chronic toxicity and the NOEL. It should be noted that this finding may more simply reflect the relative design characteristics (specifically, dose-spacing) of subchronic and chronic studies. Of course, the findings remain useful as long as the two types of studies continue to be designed as they have been and now are.

In a review by EPA (1980), the work of several authors (Barnes and Denz, 1954; Boyd, 1968; Davey, 1964; Peck, 1968; WHO Technical Report, 1966) was reviewed, and was found to support the hypothesis that tests of 3-6 months can predict chronic toxicity and NOELs. EPA (1980) also reviewed several primary studies and found that 90-day studies were reliable predictors of chronic effects.

3.2.4 Chronic Toxicity Studies

The chronic toxicity study is used to determine the effects of a substance after repeated exposure for the major portion of an animal's lifetime. There are two forms of the chronic toxicity study. One is

concerned with establishing NOELs for toxic endpoints which have a long latent period or are cumulative in nature. Such a protocol will monitor the animal throughout its lifetime for general toxicity, including neurological, physiological, biochemical, and morphological measures. The second type of chronic study is designed to determine whether a compound can induce cancer after near-lifetime exposure. This cancer bioassay does not involve monitoring the animals for general toxicity, except as it relates to shortening the lifespan of the animal. The chronic cancer bioassay is thus more limited in the information it provides about general toxicity.

It is possible to combine a chronic toxicity study with a carcinogenesis study but this requires adding more animals and including interim sacrifice groups and additional measurements. In most cases a carcinogenesis study cannot be used to replace a chronic toxicity study. However, a positive finding of carcinogenicity in a chronic toxicity study can be used as evidence of carcinogenesis.

As was true for subchronic studies, chronic studies do not provide information on potential reproductive or teratogenic hazards. Any suggestions of reproductive organ toxicity would suggest the need for a reproductive study.

3.2.5 Reproductive Toxicity Studies

Multigeneration reproduction studies are designed to assess reproductive function of an animal by evaluating effects on gonadal function, estrous cycle, mating behavior, conception, parturition, lactation, weaning, and postnatal growth and development of the offspring. Most guidelines recommend continuous exposure in two or three generations of animals, with careful

monitoring of the reproductive performance of the parents. The design of these studies provides a qualitative indication that reproduction is being adversely altered, but usually cannot provide information on the specific mechanisms causing these effects. Further tests will usually be required to ascribe the reproductive effects specifically to male or female influences.

The traditional three-generation reproduction study required two litters per generation. This was used because the first litter of a new generation was usually considered highly variable in response. Furthermore, it was felt that three generations were needed to detect transmitted genetic damage and cummulative effects that occur due to this damage. Today, both the EPA (1978, 1979) and FDA have decreased the number of required generations to two and decreased the number of litters needed per generation to one. These changes have considerably decreased the cost of a reproduction study and are now thought to provide data as reliable as that provided by the former protocols (Dixon and Hall, 1982).

Although multigeneration reproduction studies can give an indication of the presence of a potent teratogen, they are not well-suited to measure teratogenicity. For this reason, separate teratology studies are often conducted. However, a multigeneration reproduction study can be expanded to include a full teratology screen at a cost savings to the testor. FDA (1982) suggests the use of such a protocol. Furthermore, FDA (1982) suggests that, if reproductive toxicity is found in the two generation study, a teratology study should be conducted.

3.2.6 Teratology Studies

The purpose of teratology studies is to determine the effects of exposure of the embryo and fetus to the substance. Such studies are conducted for a period of time that includes the stages of organogenesis for the particular species being used. For rats and mice the period of exposure is usually 6-15 days past conception and for the rabbit 6-18 days. One day prior to birth the dams are sacrificed and the fetuses are removed for examination of gross, visceral, and skeletal abnormalities.

Teratology studies are also performed in conjunction with multigeneration reproduction studies. In these cases exposure to the substance is continuous and would occur before, during, and after conception and would continue until one day before the dam was to deliver, at which time she would be sacrificed and the fetuses removed for full examination.

Teratology studies can provide some information about reproductive function as it relates to preimplantation loss. This measure is, however, only one of many causes of infertility and cannot be used as a replacement for a full reproduction study.

3.3 Summary and Conclusions

Five major types of animal toxicity studies are routinely used by regulatory agencies to establish acceptable daily intakes of a substance. Starting from acute studies, each successive level of testing provides more reliable information with which to determine an ADI. Table 1 presents these five major toxicity tests and indicates the primary information provided by the tests, the debatable information provided, the critical information

lacking for determination of an ADI, and the agencies with standard protocols published. Through the use of this table a decision can be made as to the degree of confidence one wishes to achieve and at what cost.

In general, it appears subchronic testing provides the most information about toxicity and can, in most cases, reliably be used to predict chronic (non-carcinogenic) toxicity. Moreover, the chronic NOEL can be estimated from the subchronic NOEL with relatively high reliability. When this is considered in relation to the relative costs of subchronic and chronic tests (Table 1), it appears that subchronic tests are a considerably more cost-effective means of collecting data suitable for chronic ADI estimation than are chronic tests.

There appears to be substantial empirical support for the proposition that subchronic toxicity data can be reliably used to establish ADIs. At the same time, it needs to be recognized that for some substances, certain findings from such studies, or from other studies reported in the scientific literature, may suggest the possibility of effects not detectable in subchronic studies. Whenever such findings are reported, it is probably prudent to consider an ADI based on subchronic studies to be tentative, and to seek additional toxicity data.

For example, subchronic studies do not provide information about reproductive and teratogenic effects, but certain results from them may suggest that a substance may cause the former. If adverse reproductive effects are suggested at relatively low doses, it may be appropriate to consider a two-generation reproduction study. If reproductive damage is seen

Table 1
Toxicity Information Provided by Various
Types of Toxicity Studies

Test Type	Primary Information Provided	Debatable Information Provided	Information Lacking	Cost Per Study	Agencies with Published Protocols
Acute Exposure	1) LD_{50}, LC_{50} 2) Irritation Potential 3) Target Organ Toxicity 4) Dose Response Information for selection of doses to be used in longer term studies	1) Estimate subchronic NOEL by dividing LD_{50} by 100 2) Estimate chronic NOEL by dividing LD_{50} by 1,000-3,000	1) Cumulative toxicity 2) Subchronic toxicity 3) Chronic toxicity 4) Reproductive effects 5) Teratogenic effects 6) Carcinogenic effects	$5,000 for two species	EPA FDA OECD
Subchronic Exposure	1) Determines subchronic NOEL for estimating ADI 2) Target organ toxicity 3) Cumulative toxicity 4) Reversibility of effects 5) Physiologic and metabolic tolerance to dosing	1) Estimate chronic NOEL by dividing subchronic NOEL by 10 2) Predict all chronic effects except cancers 3) Preneoplastic changes suggest need for carcinogenic testing 4) Gonadal changes suggest need for reproductive testing	1) Carcinogenic effects 2) Reproductive effects 3) Teratogenic effects	Mouse, $35-40,000 Rat, $50-90,000	EPA FDA OECD
Chronic Exposure	1) Determine chronic NOEL for estimating ADI 2) Detect toxicities with long latent periods 3) Carcinogenic potential of the compound	1) Chronic studies do not detect any new non-carcinogenic toxicities than subchronic studies 2) A chronic carcinogenesis study cannot be used to replace a chronic toxicity study unless properly modified.	1) Reproductive effects 2) Teratogenic effects	Mouse, $200-250,000 Rat, $300-325,000	EPA FDA OECD

Table 1 (continued)

Test Type	Primary Information Provided	Debatable Information Provided	Information Lacking	Cost Per Study	Agencies with Published Protocols
Reproduction Studies	1) Determine NOEL for reproductive impairment and and use for estimating ADI 2) Reproductive impairment measured 3) Postnatal growth and development evaluated 4) Effects on lactation measured	1) Estimate teratogenic potential of a compound 2) Ascribe effects to the male or female	1) Teratogenic effects 2) Subchronic toxicity 3) Chronic toxicity 4) Carcinogenic potential	Rodent, $40-50,000	EPA FDA OECD
Teratology Studies	1) Determine NOEL for teratogenic effects and use for estimating ADI 2) Measure effects on organogenesis 3) Observe gross fetal abnormalities 4) Measure skeletal abnormalities in the fetus 5) Observe visceral abnormalities in the fetus	2) Provide some estimate of reproductive impairment due to pre-implantation loss	1) Reproductive effects 2) Subchronic toxicity 3) Chronic toxicity 4) Carcinogenic potential	Rodent, $25-30,000	EPA FDA OECD

in a two-generation reproduction study, then a teratology study might be needed to measure definitively the teratogenic potential of a compound.

Similarly, carcinogenicity can be detected and measured only through the conduct of a chronic bioassay. Before proceeding with or recommending carcinogenicity studies, however, the results of previous tests should be used to decide on the advisability and priority with which scarce monetary resources will be committed to their conduct. For example, the FDA (1982) has suggested that if a subchronic study demonstrates a substance causes focal hyperplasia, metaplasia, proliferative lesions, or necrosis, then priority should be given to conducting a carcinogenicity study. Likewise, the results of short-term mutagenicity studies have been suggested as a screen to select compounds for cancer bioassay (Food Safety Council, 1980; FDA, 1982).

Because acute toxicity data do not provide information about the effects of repeated exposure, and because predicting subchronic or chronic NOELs from LD_{50} values is not considered a validated methodology in the scientific community, it would appear that subchronic toxicity studies constitute the minimally necessary data for establishing reliable ADIs. Subchronic studies reveal a great deal about the toxic properties of chemicals and at relatively modest cost. Thus, in most cases, subchronic studies are not only minimally necessary, but are also entirely adequate to establish reliable ADIs. In addition, for the cost and in the time necessary to develop chronic toxicity data on a single chemical, ADIs can be developed on the basis of subchronic data for several (perhaps 4-6) substances.

4. Special Issues in the Use of Toxicity Data to Derive ADIs

4.1 ADIs for Essential Nutrients

Several metals that exhibit toxicity at high doses are known to be essential nutrients for humans. Among these are zinc, copper, selenium, and chromium. Selenium is not only toxic to the liver, but also induces hepatocellular tumors in experimental animals. Chromium (VI) is carcinogenic in humans, at least by inhalation (NRC, 1980).

For nutrients that are not known to be carcinogenic, the application of the standard uncertainty factors may well lead to ADIs below the recommended dietary intake level. For carcinogens, application of standard extrapolation models will reveal a finite risk of cancer at the recommended intake level. Should this be the case, it should not be inferred that a toxic risk exists at and below the recommended nutrient intake level. Rather, it suggests that the standard uncertainty factors are unnecessarily large, and, for the carcinogens, that a non-threshold model may not be appropriate for these categories of elements, probably because mammalian systems have developed homeostatic mechanisms for dealing with the toxic properties of these elements when exposures are at or near the nutrionally necessary levels (Stults, 1981; Roberts, 1981). In the case of, at least, chromium, it may also mean that the route of exposure is critical -- i.e., that carcinogenicity is not expected for ingested, rather than inhaled chromium.

ADIs have to be established for essential elements on the basis of case-by-case analysis. Judgments have to be made by first examining the toxicity data and NOELs and comparing the NOEL with the recommended daily

intake. If a very wide margin exists, it may be possible to apply the standard extrapolation factors and derive an ADI that is greater than the recommended intake level (one that allows the recommended intake to be exceeded by a significant degree). In other cases, where the margin is relatively small, it will be necessary to decide what intake in addition to that recommended for nutritional well-being can be tolerated before toxicity will almost certainly arise (i.e., it will be necessary to allow some additional intake beyond that which is essential if the ADI is to be a figure other than zero). The Safe Drinking Water Committee of the National Academy of Sciences has undertaken these types of analyses for several substances (copper, chromium, selenium, iodide, fluoride*, phosphorous, etc.) and their work can be consulted for guidance (NRC, 1977; 1980; 1983).

4.2 Mixtures and Toxicological Interactions

At most hazardous waste sites and in many other situations, there are many chemicals that may enter air or water simultaneously. It is thus of interest to examine the questions of possible biological interactions among these substances to determine whether ADIs should be adjusted to account for them.

4.2.1 General Types and Mechanisms of Interaction

Some chemicals may interact in ways such that the risk to health from exposure to a combination of chemicals differs, either qualitatively or quantitatively, from the estimated risk from exposure to each chemical by

---

* Fluoride is not essential, but is added to water supplies at a level sufficient to reduce the incidence of dental cavities.

itself. Such interactions can be synergistic or antagonistic, and can occur during absorption, distribution, metabolism, or excretion, or at the site of biological action (target site).

One striking and well-studied example of interactions among hazardous agents is the marked synergism between cigarette smoking and asbestos in the induction of lung cancer. Epidemiologic evidence indicates that an asbestos worker who smokes cigarettes has 8 times the risk of smokers of the same age who do not work with asbestos, and 92 times the risk of men who neither work with asbestos nor smoke cigarettes (NRC, 1980a).

Although the ways in which chemicals interact are complicated and incompletely understood, three general mechanism by which chemicals can interact have been identified:

- Chemical-Chemical Reactions

  A chemical may react with another in such a way that: (1) the potentially injurious chemicals(s) never reach the target site(s) in an active form; or (2) the chemical products reach the target site(s) and cause enhanced injury or an altered form of injury (NRC, 1980a).

- Chemical Competition at Macromolecules

  This mechanism of chemical interaction involves the competition for binding or reaction at a limited number of reaction sites or cellular macromolecules. These sites may control absorption, activation, detoxification, injurious action, or excretion, with competition for these sites resulting in either enhanced or reduced toxicity. This mechanism generally requires that the interacting chemicals or derivatives be present in the organism at the same time.

- Altered Cellular Responsiveness or Reactivity

  A cell or tissue may be altered by one chemical in such a way that the cell's or tissue's response to a second chemical is altered, even if the first chemical is no longer present. This mechanism is demonstrated by the initiation-promotion theory of carcinogenesis. Administration of a promoting agent weeks or even months after administration of an

initiating agent generally enhances tumor formation, while administration of the promotor before the initiator has little or no effect.

4.2.2 Interactions in Contaminated Air or Water

In most cases people exposed to contaminated air or water will be exposed to a mixture of chemicals rather than to a single substance. Because of the possibility of interactions of the types described above, a question arises about the total risk due to all the contaminants present. At least four possibilities arise:

1. The total risk is equal to the sum of the risks of each of the chemicals (more specifically -- the risks of agents showing similar effects would be strictly additive, so that one could calculate, by addition, total carcinogenic risk, or total risk of liver damage, etc.).

2. The total risk is greater than the risks obtained by addition (this represents the phenomenon of synergism).

3. The total risk is less than the risks obtained by addition (this is the phenomenon of antagonism).

4. Within a given combination of chemicals, various combinations of synergism, antagonism, or strict additivity may occur for different toxic effects.

It might be possible to subject these four possibilities to empirical tests in properly designed animal experiments. Any such experiments would, however, be extraordinarily costly and results from them are likely to be of limited generality. Thus, experiments may be conducted on certain commonly occurring combinations of chemicals (the effects of which would have to be compared with the effects of the individual constituents of the combination), but it is unlikely that any such results would be clearly applicable to other combinations of chemicals, or to the same chemicals occurring in different proportions. This view may be unduly pessimistic, however, and it may be useful to consider this testing question further.

Even if major research efforts were initiated on the question of interactions, no useful data would likely be available for several years. No generally applicable protocols for such tests appear to be available, and problems of study design, conduct and interpretation have only been discussed from a theoretical viewpoint (NRC, 1980a; 1982). If risk assessments or ADI derivations are to take into account possible interactions, they must for the present be based on consideration other than empirical evidence (EPA, 1984).

Unless the use of uncertainty factors in the derivation of ADIs is considered in part to compensate for the possibility of interactions, then there is no area of risk or safety evaluation that has, as a matter of course, taken interactions into account. The major reason for this is that there is very little data available to demonstrate toxic interactions, especially of chemicals found at waste sites (most data come from studies of drugs, Calabrese, 1983).

It thus appears that, unless specific data become available that reveal the mode of interaction among groups of chemicals that will be known always to co-occur in the same relative proportions, and that will reveal the quantitative effect of one upon the others, it is probably not possible to take interactions generically into account in deriving ADIs. Under such circumstances, it would appear that treating each substance as an independently-acting toxicant would be the course most in keeping with current scientific understanding.

5. Apportionment of RfDs and RSDs

5.1 Introduction

The final issue concerns apportionment of RfDs and RSDs (we shall now use the preferred EPA terms) among sources and between air and water. To ensure that potentially exposed individuals do not experience intakes significantly in excess of the RfD, it is necessary to establish limits on the concentrations of individual substances in air and water.

In deriving such limits, two major issues must be considered:

1. Apportionment of RfDs among media and sources; and
2. Relationships between concentrations in air and water, and human doses.

5.2 Apportionment Among Media and Sources

The RfD for a chemical represents the maximum allowable daily dose of the chemical that is anticipated to have no adverse effect in humans following chronic exposure. For most chemicals, it is the systemic dose that is important in defining the RfD; the route of exposure, whether it be by inhalation, ingestion, or dermal contact, is not of major concern.

Since exposure to a chemical may arise from different sources (that is, not solely as a result of waste disposal) and via different routes of exposure (for example, inhalation from the air, ingestion in food or drinking water, or dermal contact), the RfD must represent the limit on total dose received from all sources and via all routes. Derivation of limits on concentrations

of substances in a particular medium (for example, air or water) must allow for the possibility that other sources and routes of exposure to the substance exist.

In the present case, we are interested in deriving limits on concentrations of chemicals from waste sites in air and water. Chemicals migrating from waste sites may enter either or both of these media to varying degrees depending on chemical-specific physical characteristics. For example, a chemical which is highly volatile but poorly soluble in water, such as dichlorodifluoromethane, will tend to migrate into the air, while a water soluble chemical of low volatility, such as many inorganic salts, will tend to migrate into water. However, contamination of water and air from a chemical waste disposal site is not the only way in which people may be exposed to a chemical. Some compounds of wastes (for example, certain pesticides and organic solvents) are widespread environmental contaminants to which human exposure may occur via air, food, or consumer products. Others, for example, many inorganic substances such arsenic and cadmium, occur naturally in the environment, and again may reach people through food and several other media.

In setting maximum allowable concentrations of chemicals in water and air, allowance must be made for potential exposure from other sources and by other routes. Since exposure may occur by several routes, and the RfD represents, by definition, the total allowable exposure, the RfD must be apportioned over the various possible routes of exposure.

The concept of apportionment of a chemical by medium and by route of exposure is not new. The NRC Safe Drinking Water Committee (NRC, 1980)

calculated suggested no-adverse-response levels (SNARLs) for chronic exposure to non-carcinogens in drinking water while incorporating an "arbitrary assumption" that 20% of the intake of the chemical was from drinking water. The EPA in setting maximum contaminant levels (MCLs) for chemicals in tap water, also selects a fraction of the ADI (usually 20% if there are no data to suggest some other fraction).

Another use of apportionment was a risk evaluation procedure developed for EPA's Office of Emergency and Remedial Response to evaluate and manage the risks for specific remedial action sites. This procedure apportioned concentrations equally in environmental media (e.g., air and water) as an initial basis for calculating an allowable rate of release to the environment; at times, unequal apportionment was selected, if there were significant cost and feasibility differences in controlling exposures via the different pathways (ENVIRON, 1983).

The Food, Drug and Cosmetic Act [Section 409(c)(5)] specifies that in deciding whether a proposed use of a food additive is safe the FDA must consider certain relevant factors. Included as one of these factors is a consideration of "the cumulative effect of such additive in the diet of man or animals, taking into account any chemically or pharmacologically related substances or substances in such diet." This language has been interpreted by the FDA as requiring that all sources of exposure must be combined in order to estimate the total exposure to an additive. This total exposure level is then compared to the ADI to decide if the proposed uses are safe.

In Section 706(b)(5)(A) of the Food, Drug, and Cosmetic Act it is also specified that for a color additive the FDA will determine if a color additive is safe after specific factors are considered. The factors include "the probable consumption of, or other relevant exposure from, the additive and of any substance formed in or on food, drugs, devices, or cosmetics because of use of the additive. In addition the Secretary must consider, "the cumulative effect, if any, of such additives in the diet of man or animals, taking into account the same or any chemically or pharmacologically related substance or substances in such diet." As for food additives, colors must be reviewed for safety after all uses of the color have been combined and compared to the ADI.

In its review of lead in the human environment, the NRC (1980) stressed that one step in a comprehensive risk assessment requires that the level of exposure be estimated quantitatively for each pathway that affects the target population. Quantitative assessment of exposure for lead required estimation of population exposure from dust, air, soil, water, paint, food, and cosmetics. Although uncertainties are associated with exposure estimates for these various sources, it was agreed that all contributions to the "total exposure" were needed to establish the safety of lead.

NRC (1980) suggested that although it is difficult to measure specific lead source contributions directly, various modeling techniques can assess the relationship between lead in the environment and human exposure levels of lead. The EPA's Office of Drinking Water had developed a detailed model for estimating the relative contributions of air, water, and other sources of lead to the total exposure (Drill et al., 1979 as cited in NRC, 1980) for various populations.

It should also be noted that if the total RfD is used for hazardous waste sources of a chemical, then the only scientifically supportable decisions regarding other sources of the chemical is zero exposure. While the issue of apportioning allowable exposure levels among various sources and media has strong scientific justification, and considerable precedent, the choice of fractions of total exposure to allot to the various media is less clear. In the context of the present work, two issues are important in making this choice. The first relates to how much of the total allowable exposure may come from sources other than water and air, and the second relates to how exposure should be apportioned between water and air.

Many of the chemicals of concern are common environmental contaminants. The most rigorous procedure, scientifically, would be to analyze on a case-by-case basis each potential exposure situation to determine background levels of exposure to the substance in question and allot an appropriate fraction of the RfD on that basis, retaining the remaining fraction for air and water exposure resulting from escape of the substance from the waste disposal site. This would entail a level of effort that is out of proportion to its importance in protecting public health, and, because of data gaps for many chemicals, is not likely to be productive. As an alternative, alloting 50% of the RfD to background exposures and 50% to waste site-related air and water exposure seems a useful first approximation. A more rigorous apportionment is not called for in light of the facts, noted previously, that the RfD is itself subject to considerable uncertainty, and that occasional excursions above the RfD, for a relatively small number of substances, will not likely produce excess toxic risk (see Section 2). It would be prudent, however, to reconsider the 50% figure whenever readily available data on background levels

of specific substances reveal it to be too generous, and to use an alternative value if sufficient data on environmental distribution of the substance are available to justify an alternative.

This 50% allotment of the RfD is probably not necessary for carcinogens. For such substances, the equivalent dose (the RSD) is estimated by a procedure which introduces unavoidable uncertainties. The procedure used is deliberately selected to be conservative; so that a twofold difference in dose is well within the margin of uncertainty of the estimated RSD.

Moreover, for carcinogens, the determinate of risk is the daily dose averaged over a full lifetime. Small variations around the daily dose have little effect on the lifetime risk, as long as the average is not affected. For this reason, a two fold reduction in the RSD is relatively insignificant. For non-carcinogens, it is possible that not applying the 50% reduction (the indirect effect of which is to permit an approximate doubling of the ADI) may cause the threshold to be exceeded on some or even many days of the human exposure period. Exceeding the threshold of effect may have significant health consequences for some individuals. Thus, there appears to be justification for treating non-carcinogens differently from carcinogens with respect to this apportionment issue.

Before turning to the question of apportionment between air and water, we first discuss the interrelationship between the concentration of a substance in air or water and the human dose of the substance resulting from drinking the water or breathing the air.

5.3 Relationships Between Air or Water Concentration and Human Dose

The RfD or RSD is defined in terms of a daily dose or daily intake, generally measured in mg/kg/day. To define maximum allowable concentrations of a chemical in environmental media (in this case, water and air) it is necessary to know the relationship between the concentration in each medium and the daily human dose resulting from normal intake of that medium. For many substances, the daily intake (in mg/kg/day) may be calculated simply by multiplying the concentration of the chemical in the medium by the daily human intake of that medium and dividing by the human body weight:

Intake of chemical (mg/kg/day) = concentration in medium (mg/l) x daily intake of medium (l/day)/body weight (kg)

Adjustments to this simple equation may be necessary to account for incomplete absorption of the ingested or inhaled chemical. However, if the RfD is based on intake by the same route as that by which the human intake is being calculated, and if it can be assumed that the degree of absorption occurring in the experimental situation is the same as that in the humans of concern, such an adjustment will not be necessary (EPA, 1984; NRC, 1983).

It is generally assumed that the greatest contribution to the exposure from tap water results from direct ingestion and that inhalation of vapors and aerosols of water contaminants while showering, or dermal absorption of those substances during bathing, are relatively trivial contributors to exposure (NRC, 1983). It is advisable, however, that empirical verification or refutation of this premise be sought, because it is not clear it would hold for all chemicals (especially highly volatile ones).

If the RfD is based on exposure through one medium, say air, and we are attempting to derive a safe concentration in another medium, say water, adjustment for absorption will be necessary, unless the degree of absorption of the chemical when inhaled is the same as the degree of absorption of the chemical when ingested in water.

Specific absorption information is known in experimental animals for numerous compounds, but in humans for far fewer substances (Calabrese, 1983). However, the data from experimental animals represent reasonable approximations of those parameters in humans (Calabrese, 1983). In general, absorption of retained foreign compounds is greatest (i.e., both rate and efficiency) via the lungs, less by water in the gastrointestinal tract, still less by food in the gastrointestinal tract, and least by the skin (Klaassen, 1980).

Absorption data are not available on all substances of interest by all routes of exposure. The conversion from RfD to media concentrations, therefore, necessarily relies on knowledge from similar compounds for which such information is available. For example, among the chlorinated alkanes, it is possible to approximate crudely the degree of absorption of members of the class whose absorption is not known from absorption data on those members of the class for which absorption is known in experimental species considered predictive of chemical behavior in humans. Similar estimations are possible among inorganic metals (Calabrese, 1983; NAS, 1975).

For airborne particles, it is important to distinguish between alternative sites of deposition in order to apportion systemic doses since the

site of deposition influences how and to what extent a substance may be absorbed. Although deposition is determined by a number of variables including particle charge and shape, aerodynamic diameter can be used to estimate the most likely site of deposition. For airborne materials, those that are in particulate form are often not retained, and hence not absorbed, if smaller than 2-5 microns in diameter. Those between 5 and 20 microns in diameter are efficiently retained for absorption. Those greater than 20 microns in diameter are generally deposited in the upper respiratory tract from which they may be cleared by the mucocilliary escalator and swallowed. For such substances, therefore, it is gastrointestinal absorption rather than absorption through the lung that is critical in defining absorption.

As noted earlier, the conversion of the RfD to concentrations in environmental media requires knowledge of the extent of human exposure to the media themselves in addition to knowledge of the extent of absorption. Where specific data of this type exist, they are incorporated into the analysis of daily exposure. However, in the absence of such data, assumptions must be made. In the present case a variety of assumptions have been made to convert between daily human dose levels and media concentrations (see Table 4).

The previous discussion dealt with agents that produce injury through systemic distribution and selective affinity and injury to specific tissues. Some compounds, such as acids and alkalies, when present in adequate concentrations will damage the tissues with which they come into direct contact. Such effects are unusual for substances that have migrated from waste sites because of dilution, buffering, and other physical and biological influences, and are not considered further in this document.

The data in Table 4 (or other specific data if available) may be used to convert a daily human dose to a corresponding concentration in air or water using the general equations below.

Conversion of Daily Human Dose to Equivalent Air Concentration

$$\text{Air Concentration } (mg/m^3) = \frac{\text{Daily dose } (mg/kg/day) \times \text{body weight } (kg) \times \text{correction factor}}{m^3 \text{ air breathed/day}}$$

Conversion of Daily Human Dose to Equivalent Water Concentration

$$\text{Water Concentrations } (mg/l) = \frac{\text{Daily dose } (mg/kg/day) \times \text{b/w } (kg) \times \text{correction factor}}{\text{liters water consumed/day}}$$

The procedures described above for interconverting between media concentrations and human doses are typical of those commonly used by regulatory agencies and other bodies for such purposes. For example, EPA currently used identical procedures to set ambient water quality criteria (USEPA 1980) and maximum contaminant levels.

The National Research Council (1977) describes calculations directly analogous to those above for deriving acceptable water concentrations from acceptable daily doses. Likewise, the American Conference of Governmental Industrial Hygienists (1980) uses similar procedures to derive maximum permissible air concentrations (TLVs) for substances for which the only relevant data are derived from studies in which the substance in question is administered in a different route (for example, orally).

These procedures are also virtually identical to those originally published in 1958 by Stokinger and Woodward (1958). The validity of this procedure was discussed in a recent EPA conference (USEPA 1984), which

Table 4. Assumptions for Converting Daily Human Doses to Media Concentrations

Adult male body weight	70 kg (ICRP, 1975)
Adult male body surface area	1.8 m^2 (ICRP, 1975)
Volume of air breathed by adult male per day	23 m^3 (ICRP, 1975)
Efficiency of pulmonary absorption	100% (unless data to the contrary) (Calabrese, 1984)
Amount of water consumed by adult male per day	2 liters (ICRP, 1975)
Efficiency of gastrointestinal absorption	100% (unless data to the contrary) (Calabrese, 1984)
Correction factor	1.0*

* Adjusts for different extents of absorption if the RfD is based on a route of exposure other than that for which the RfD is being derived. If differences in the extent of absorption have not been reported, they are assumed to be identical, and therefore, the correction factor is 1.0.

concluded that it provides a reasonable first approximation, though estimates obtained are likely to be somewhat inaccurate if factors such as absorption and pharmacokinetics are not taken into consideration.

In the present context, the "daily dose" which is being converted to an air or water concentration would be the portion of the RfD or RSD that is alloted to the medium. What proportion of the total RfD or RSD is alloted to each medium is the subject of the next section.

5.4 Apportionment Between Air and Water

We have already noted how it is appropriate to allot just a portion (50%) of the total RfD to air and water contamination resulting from escape of chemicals from hazardous waste disposal sites to ensure that the total RfD is not exceeded in the likely event that some exposure to the chemicals of concern occurs via other media, particularly food. In this section we describe how the portion of the RfD that is alloted to air and water can be partitioned between these two media. In the end it must be ensured that the maximum concentrations permitted in air (in mg/m^3) and in water (in mg/l) yield a total exposure no greater than 0.5 RfD.

Many volatile and semivolatile chemicals may be present in both air and water, which present dual pathways of exposure. Atmospheric dispersion may substantially reduce the concentration of chemicals in air to a much greater extent normally than would be expected in surface water or ground water systems. However, airborne chemicals may accumulate in poorly ventilated

areas such as enclosed buildings where significant airborne exposures may occur.

There are analytical methodologies that could represent the individual, complex processes that affect the partitioning of chemicals between air and water, and the transportation of these chemicals from the source area to the receptor. However, no analytical approach would be suitable to all such processes. To make such predictions would require site and chemical specific data and reliable, verified models of nonconservative atmospheric and water borne transport processes. The application of such models for the purpose of partitioning the RfD is not justified in most instances, given the lack of data and verified models at most sites. Moreover, use of such refined models would seem incompatible with the relatively crude approximations used to derive the RfD.

There are two physical characteristics of chemicals that describe their behavior in air and water. The octananol-water partition coefficient (K_{ow}) is a measure of a chemical's partitioning between water and an organic phase (approximately represented by octanol). The K_{ow} for a chemical provides an indication of its solubility in water, and may describe its behavior in an environment likely to be present at a hazardous waste site. Specifically, chemicals with large values of K_{ow} are preferentially retained in an organic phase and only poorly soluble in water. Conversely, substances with small values of K_{ow} are more readily soluble in water than in an organic phase. It would thus appear that chemicals with small values of K_{ow} (e.g., high solubility chemicals such as phenols and halogenated phenols) are more likely to escape a landfill in an aqueous phase than those with large values of K_{ow}.

Once present in water, a chemical may partition between air and water. For a dilute solution the ideal gas vapor pressure (an alternative expression of concentration) of a volatile solute is proportional to its concentration in the solution. The gas vapor pressure, or air:water partition, is described by Henry's law (Tinsley, 1979), which can be expresssed as:

$$C_a = H_c\ C_w$$

where C_a is the chemical concentration in the gas (air) phase (mg/m^3), C_w is the chemical concentration in the liquid (water) phase (mg/ℓ), and, H_c is Henry's Law constant ($mg/m^3/mg/\ell$). It is important to note that Henry's Law applies rigorously only to dilute solutions where solute-solute interactions are negligible. H_c is customarily expressed as atm-m^3/mole when C_a is expressed in terms of partial pressure (atm). Using Boyle's Law for ideal gases, PV = NRT, the conversion of partial pressure (P) to moles/m^3 can be achieved by dividing P by RT. The gas constant, R, is 8.205 x 10^{-5} atm m^3/moles° K, and the ambient temperature T is normally assumed to be 20°C (293°K). Therefore, to convert H_c from units of atm-m^3/mole to $mg/m^3/mg/\ell$ (i.e., ℓ/m^3), one must multiply by 1/RT (i.e. 4.16 x 10^4).

The relative air and water concentrations of a chemical, at equilibrium will be indicated by the value of Henry's Law constant H_c, (units of $mg/m^3/mg/\ell$ or atm m^3/mol). Chemicals with large values of H_c will have a tendency to exist predominately in air, whereas those with low values will partition preferentially to water.

It is recognized the values of K_{ow} and H_c for specific chemicals are obtained in the laboratory under idealized conditions, and that their application in a setting as complex as a hazardous waste site can not be expected to yield accurate predictions of the behavior of a chemical. It is nevertheless true that these two physical constants will predict trends in behavior. That is, chemicals with low values of K_{ow} and low values of H_c are highly soluble in water and poorly volatile, so that their concentrations are likely to be high in water and low in air, relative to chemicals with high values of the two constants, assuming that the source is not limiting. One approach to deciding how to partition a chemical between air and water would depend on these trends in behavior, as described by K_{ow} and H_c. There are, of course, no standard definitions of "high" or "low" with respect to the two physical parameters K_{ow} and H_c. Values of K_{ow} and H_c for selected chemicals that have been reported in the literature are presented in Table 5. The chemicals are those considered as especially important by EPA. A simple scheme to partition the RfD using K_{ow} and H_c is presented in Table 6. The partition chosen is meant only to reflect the general direction of expected migration of the chemical from a source to water and air.

A model more refined than that shown in Table 6, i.e, depending only on K_{ow} and H_c, and which results in a partitioning of the RfD into more than three broad groups (partitioning mainly into air; partitioning mainly into water; and approximately equal partitioning between the two media) would not appear to be warranted, given the uncertainties inherent in the RfD and the

Table 5. Henry's Law Constants and Octanol-Water Coefficients for Selected Hazardous Chemicals

Chemical	Henry's Law Constant atm m^3/mole	Henry's Law Constant ℓ/m^3	Octanol-Water Coefficient
Carbon Disulfide	1.68E-02	6.99E+02	1.45E+02
Chlorobenzene	3.46E-03	1.44E+02	7.41E+02
Cresols	5.05E-06	2.10E-01	1.41E+02
1,2 Dichlorobenzene	1.88E-03	7.82E+01	3.80E+03
Methylene chloride	3.19E-03	1.33E+02	1.80E+02
Trichloromonofloromethane	8.02E-01	3.34E+04	3.31E+02
Isobutyl Alcohol	1.23E-05	5.12E-01	5.50E+00
Methyl Ethyl Ketone	2.61E-05	1.09E+00	2.00E+00
Nitrobenzene	2.40E-05	9.98E-01	7.94E+01
Pyridine	1.95E-07	8.11E-03	4.79E+00
Tetrachloroethylene	2.87E-02	1.19E+03	5.80E+02
2,3,4,6 Tetrachlorophenol	4.53E-06	1.88E-01	2.14E+04
Toluene	5.93E-03	2.47E+02	6.61E+02
Methylchloroform	2.76E-02	1.15E+03	3.16E+02
Trichloroethylene	1.17E-02	4.87E+02	2.29E+02
2,4,5 Trichlorophenol	2.84E-05	1.18E-01	7.24E+03
2,4,6 Trichlorophenol	1.77E-05	7.36E-01	2.93E+03
Pentachlorophenol	4.62E-06	1.92E-02	1.15E+05
1,2,2 Trichloro-1,2,2, Trifluoroethane	9.00E+00	3.74E+05	1.26E+03
Ethylbenzene	6.44E-03	2.68E+02	1.41E+03

Table 6: RfD Partition
Between Water and Air using K_{ow} and H_c

K_{ow}		H_c Low	H_c High
	Low	Air:Water 50:50	Air:Water 80:20
	High	Air:Water 20:80	Air:Water 50:50

low predictive power of H_c and K_{ow} in the context of a hazardous waste environment.

A further simplification of the partitioning scheme can be envisioned. The ultimate question to be addressed in partitioning the RfD is the relative rather than absolute concentrations in air and water that a chemical might achieve at equilibrium. The relative concentrations can be indicated by the use only of H_c.

The RfD is comparable to the total dose (or intake) of a specific chemical by the receptor. The dose is a function not only of the chemical concentration in air and water, but also of the breathing and ingestion rates, and absorption through the respiratory and gastrointestinal systems. Lacking specific data, inhaled or ingested chemicals are assumed to be totally absorbed. However, breathing and ingestion rates are well established and can be considered in the partitioning of the RfD.

The total dose of a specific chemical by the combined air-water pathways is given by:

Total Dose = dose inhaled + dose ingested by consumption of water

$$= BR * C_a + IR * C_w$$

where BR is the breathing rate (m^3/day), IR is the water ingestion rate (ℓ/day), and C_a and C_w are as defined previously.

It is assumed that at equilibrium the chemical concentration in air, C_a (mg/m^3) is related to the chemical concentration in water, C_w (mg/ℓ), by the Henry's law constant, H_c (ℓ/m^3). The dose model can be rewritten as:

$$\text{Total Dose} = BR * H_c * C_w + IR * C_w$$
$$= (BR * H_c + IR) * C_w$$

where the air and water based doses can be represented by ($BR * H_c * C_w$) and ($IR * C_w$), respectively.

Using the dose model, the partitioning of the RfD for each chemical between the air and water pathways can be calculated by the proportion that each contributes to the total dose.

(Air)		(Water)
$\frac{BR * H_c}{BR*H_c+IR}$	:	$\frac{IR}{BR*H_c+IR}$

When calculating the air:water partition, breathing and ingestion rates of 23 m^3/day and 2 ℓ/day, respectively, can be assumed (see previous section). The partition model can then be rewritten as:

(Air)		(Water)
$\frac{23 * H_c}{23\ H_c+2}$	:	$\frac{2}{23\ H_c+2}$

where H_c must be expressed in units of $mg/m^3/mg/\ell$, (ℓ/m^3).

An even more simplified approach might involve using the dose model described above with ranges of H_c, in which the air:water partition could be represented by high, moderate and low ranges. Such an approach is illustrated in Table 7. Although this approach may be less precise than a chemical specific calculation, it may more reasonably reflect the inherent uncertainties in these simplistic models of complex natural processes.

Table 7
RfD Partition Between Air and Water Using
Henry's Law Constant (H_c)

H_c Range (ℓ/m^3)	Air:Water Partition
>0.35	80:20
0.35 - 0.02	50:50
<0.02	20:80

* The values of H_c shown in this Table were derived by assuming the air:water partition values (e.g., 80:20, air to water) and calculating H_c from:

$$\frac{23 * H_c}{23 H_c + 2} \quad (\text{Air}) \quad : \quad \frac{2}{23 H_c + 2} \quad (\text{Water})$$

The representation of the air:water partition by the equilibrium relationships described herein is admittedly a somewhat simplistic approximation to the partitioning of the RfD. The approach may accurately represent the relative contribution of air and waterborne chemicals only in close proximity of the chemical source. At greater distances the predicted air:water partition would be less precise. However, it is intended that the partitioning of the RfD will be established within reasonable limits that are chemical specific, but not to be reevaluated on a site by site basis. The proposed approaches accomplish this purpose.

6. Conclusions and Recommendations

1) Since its introduction in the early 1950s, the ADI has been widely used by regulatory and public health agencies as a practical, health protection tool. It appears to be the appropriate criterion for establishing limits for substances that may migrate into environmental media to which humans may be chronically exposed. EPA now proposes to adopt the term "Reference Dose" (RfD) to replace ADI. All references to ADI in the following also describe the RfD.

2) The ADI should not be considered a sharp dividing line between "safe" and "unsafe" exposures. It is a practical tool, subject to considerable scientific uncertainty, and occasional excursions above the ADI should not be considered cause for concern. It is not possible, however, to quantify the magnitude of uncertainty associated with any given ADI.

3) Consistent protection at or near the ADI ensures that individuals will be protected from the acute effects of all chemicals. Protection against acute toxic effects usually requires safety factors no larger (and sometimes smaller) then those used for establishing the ADI. Because the chronic NOEL will always be a lower dose that the minimum effective acutely toxic dose, the ADI will clearly protect against acute toxicity.

4) In general, it appears subchronic testing can be used to predict chronic (non-carcinogenic) toxicity. When this is considered in relation to the relative costs of subchronic and chronic tests (Table 1), it appears that

subchronic tests are a considerably more cost-effective means of collecting data suitable for ADI estimation than are chronic tests.

5) Although the data base supporting the conclusion (4) is relatively extensive, it is not without uncertainty. Evaluations of the need to move beyond subchronic should be made on a case-by-case basis, giving due consideration to the results of acute and subchronic studies, chemical structure, and metabolic information.

6) Subchronic and chronic tests do not provide information about reproductive injury and teratogenic effects. Conduct of a two-generation reproduction study can provide most of this information as well as an indication of frank teratogenic effects. If reproductive damage is seen in a two-generation reproduction study, then a teratology study would be needed to measure definitively the teratogenic potential of a compound.

7) Carcinogenicity can be measured only through the conduct of a chronic bioassay. Before proceeding with cancer bioassays, however, the results of previous tests should be used to decide on the advisability and priority with which scarce monetary resources will be committed to their conduct. For example, the FDA (1982) has suggested that if a subchronic study demonstrates focal hyperplasia, metaplasia, proliferative lesions, or necrosis, then priority should be given to conducting a carcinogenicity study. Likewise, the results of short-term mutagenicity studies have been suggested as a reliable screen to select compounds for cancer bioassay (Food Safety Council, 1980; FDA 1982).

8) Because acute toxicity data do not provide information about the effects of repeated exposure, and because predicting subchronic or chronic NOELs from LD_{50} values is not considered a validated methodology in the scientific community, it would appear that subchronic toxicity studies constitute the minimally necessary data for establishing reliable ADIs. Subchronic studies reveal a great deal about the toxic properties of chemicals and at relatively modest cost. In most cases, data from such studies should be fully adequate to establish an ADI. Decisions about the need for additional toxicity data, and their value relative to the costs involved, should be made on a case-by-case basis, under the general criteria described above (items 5, 6, 7).

9) ADIs should be established for essential elements on the basis of case-by-case analysis. Judgments will have to be made by first examining the toxicity data and NOELs and comparing the NOEL with the recommended daily intake. If a very wide margin exists, it may be possible to apply the standard extrapolation factors and derive an ADI that is greater than the recommended intake level. In other cases, where the margin is relatively small, it will be necessary to decide what intake in addition to that recommended for nutritional well-being can be tolerated before toxicity will almost certainly arise (i.e., it will be necessary to allow some additional intake beyond that which is essential if the ADI is to be a figure other than zero).

10) Since most people will be exposed to mixtures of chemicals rather than single substances, the possibility of interactions among chemical raises questions about the total risk due to all the contaminants present.

It appears that, unless there are specific data available that reveal the mode of interaction among groups of chemicals that will be known always to co-occur, and that will reveal the quantitative effect of one upon the others, it is probably not possible to take interactions generically into account in deriving ADIs.

11) The final issue concerns apportionment of ADIs between air and water. To ensure that potentially exposed individuals do not experience intakes in excess of the ADI, it is necessary to establish limits on the concentrations of individual substances in air and water. Because indivuduals may be exposed through other media, only a portion of the ADI can be used for allocation to air and water contamination. There are several precedents for making such allocations for non-carcinogenic chemicals.

12) While the issue of apportioning chemicals between air and water is susceptible to some degree of analytic examination, there appears to be no readily definable means to select prospectively the portion of the ADI that is allotted to "all other exposures." In the general case, allocation of 50% of the ADI for air and water at hazardous waste sites and the remaining 50% to other exposures would, in the absence of data to the contrary, seem to represent a reasonable first approximation. It would probably be prudent, however, to reconsider the 50% figure whenever readily available data on background levels of specific substances reveal it to be incorrect.

13) In many (perhaps most) cases, substances will likely enter both air and water. The 50% of the ADI alloted to waste site exposures thus must be partitioned between the two media, and maximum allowable concentrations (in mg/l water and mg/m^3 air) must be established to ensure total exposure does not exceed the acceptable limits. The octanol-water partition coefficient (K_{ow}) together with Henry's Law Constant (H_c), can be used to indicate trends in partitioning. Alternatively, it is also possible to use a simpler scheme, involving only H_c, to decide on the approximate apportioning specific chemicals will assume.

References

American Conference of Governmental Industrial Hygienists 1980. Documentation of the TLVs. ACGIH, Cincinnati.

Barnes, J.M. and Denz, F.A. 1985. Experimental Methods used in Determining Chronic Toxicity. Pharmcol. Rev. 6:191-242.

Boyd, E.M. and Lloyd, G.R. 1965. Minimal Oral Doses of Benzylpenicillin Producing Signs of Toxicity in Albino Rats. Arch. Int. Pharmacodyn. 156:172-179.

Calabrese, E.J. 1983. Principles of Animal Extrapolation. John Wiley and Sons, New York. Pp 55-69.

Carcinogen Assessment Group. 1985. Relative Carcinogenic Potencies Among 53 Chemicals Evaluated by the Carcinogenic Assessment Group as Suspect Human Carcinogens. in EPA's Health Assessment Document for Dichloromethane (Methylene Chloride). February, 1985.

Davey, D.G. 1964. The Study of the Toxicity of a Potential Drug: Basic Principles. Eur. Soc. Study Drug Toxic. Suppl. 6:1-13.

Dixon, R.L. and Hall, J.L. 1982. Reproductive toxicology. In Hayes, A.W., ed. Principles and Methods of Toxicology. Raven Press, New York.

Doull, J., Klassen, C.D., and Amdur, M.O., eds. 1980. Casarett and Doull's Toxicology. Second Edition. Macmillan Publishing Co., Inc., New York.

Dourson, M.J., and Stara, J.F. 1983. Regulatory History and Experimental Support of Uncertainty (Safety) Factors. Regulatory Toxicology and Pharmacology 3:224-238.

Environ Corporation. 1985. Documentation of the Development of Toxicity and Volume Scores for the Purpose of Scheduling Hazardous Wastes. Prepared for EPA, Office of Solid Waste, Contract No. 68-01-6861. March 28, 1985.

Environmental Protection Agency (EPA). 1978. Proposed guidelines for registering pesticides in the United States; hazatd evaluation: Humans and domestic animals. Fed. Reg. 43(163):3733-37403.

Environmental Protection Agency (EPA). 1979. Support Document Test Data Development Standards: Chronic Health Effects. EPA-560/11-001.

Environmental Protection Agency (EPA). 1980. Water Quality Criteria Documents; Availability. Fed. Reg. 45:79318-79379.

Environmental Protection Agency (EPA). 1980a. Appendix C - Guidelines and Methodology Used in the Preparation of Health Effect Assessment Chapters of the Consent Decree Water Criteria Documents. 45 FR 79347-79357.

Environmental Protection Agency (EPA). 1980a. Proceedings of the Workshop on Subchronic Toxicity Testing; Denver, Colorado, May 20-24, 1979. EPA-560/11-780-028.

Environmental Protection Agency (EPA). 1980b. Scientific Rationale for the Selection of Toxicity Testing Methods: Human Health Assessment. EPA-560/1-80-001.

Environmental Protection Agency (EPA). 1982. Health Effects Assessment Summary for 300 Hazardous Organic Constituents in Support of Regulatory Impact Analysis of the Land Disposal Branch (LDB) and Interim Final Incinerator Regulations of the Technical Branch (TB) of the Office of Solid Wastes, OSWER/EPA, Draft, September 20, 1982.

Environmental Protection Agency (EPA). 1984. Approaches to risk assessment for multiple chemical exposures. Proceedings of a workshop. EPA-600/9-84-008.

Environmental Protection Agency (EPA). 1984. Summary of Current Acceptable Daily Intakes (ADIs) for Oral Exposure. Internal Review Draft. Environmental Criteria and Assessment Office, Cincinnati, Ohio. February, 1984.

FAO/WHO Expert Committee on Food Additives. 1958. Procedures for the Testing of Intentional Food Additives to Establish Their Safety for Use. World Health Organization Technical Report Series. Rep. 229. Geneva. pp 104.

FAO/WHO Expert Committee on Pesticides Residues. 1965. Report of the Second Joint Meeting of the FAO Committee on Pesticides in Agriculture and the WHO Expert Committee on Pesticide Residues.

Fears, T.R., and Douglas, J.F. 1977. Suggested Procedures for Reducing the Pathology Workload in a Carcinogen Bioassay Program, Part I: J. Environ. Toxicol. 1:125-137.

Fears, T.R., and Douglas, J.F. 1978. Suggested Procedures for Reducing the Pathology Workload in a Carcinogen Bioassay Program, Part II: Incorporating Blind Pathology Techniques and Analysis for Animals with Tumors. J. Environ. Toxicol. 1:211-222.

Food, Drug, and Cosmetic Act, as Amended January 1980. Published by DHHS. Publication No. (FDA 80-1051).

Food and Drug Administration (FDA). 1982. Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food. FDA, Bureau of Foods.

Food Safety Council. 1980. Proposed System for Food Safety Assessment. Washington, D.C. pp 160.

International Commission on Radiological Protection (ICRP). 1975. Report of the Task Group on Reference Man. Pergamon Press, New York.

Klaassen, C.D. 1980. Absorption, Distribution, and Excretion of Toxicants. In Doull, J., Klaassen, C.D., and Amdur, M.O., eds. Casarett and Doull's Toxicology. Second Edition, Macmillan, New York. Pp 28-55.

Kulwich, B.A. Hardisty, J.F., Gilmore, C.E., and Ward, J.M. 1980. Correlation Between Gross Observations of Tumors and Neoplasms Diagnosed Microscopically in Carcinogenesis Bioassays in Rats. J. Environ. Pathol. Toxicol. 3:281-287.

Lehman, A.J. and Fitzhugh, O.J. 1954. 100-Fold Margin of Safety. Assoc. Food Drug. Off. U.S.Q. Bull. 18:33-35.

Loomis, T. 1978. Essentials of Toxicology, 3rd edition. Lea and Febiger, Philadelphia, PA. pp 245.

McNamara, B.P. 1976. Concepts in Health Evaluation of Commerical and Industrial Chemicals. In New Concepts in Safety Evaluation (Mehlman, M.A., Shapiro, R.E. and Blumenthal, H., eds.) Adv. Mod. Toxicol. 1:61-140.

National Academy of Sciences (NAS). 1975. Principles for Evaluating Chemicals in the Environment. National Academy of Sciences, Washington, D.C.

National Research Council (NRC). 1977. Drinking Water and Health - Volume 1. National Academy of Sciences, Washington, D.C.

National Research Council (NRC). 1977a. Principles and Procedures for Evaluating the Toxicity of Household Substances. National Academy of Sciences, Washington, DC.

National Research Council (NRC). 1980. Drinking Water and Health, Volume 3. National Academy Press, Washington, D.C.

National Research Council (NRC). 1980a. Principle of Toxicological Interactions Associated with Multiple Chemical Exposures. Panel on Evaluations of Hazards Associated with Maritime Personnel Exposed to Multiple Cargo Vapors, Board on Toxicology and Environmental Health Hazards, and Committee on Maritime Hazardous Materials, National Materials Advisory Board, National Research Council. National Academy Press, Washington, D.C.

National Research Council (NRC). 1980b. Lead in the Human Environment. National Academy Press, Washington, D.C.

National Research Council (NRC). 1982. Strategies to Determine Needs and Priorities for Toxicity Testing. Volume 2: Development. National Academy Press, Washington, DC.

National Research Council (NRC). 1983. Drinking Water and Health. Volume 5. National Academy Press, Washington, D.C.

National Toxicology Program (NTP). 1984. Report of the NTP Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation. pp 280.

Organization for Economic Co-operation and Development (OECD). 1981. OECD Guidelines for Testing of Chemicals. Paris, France.

Peck, H.M. 1968. An Appraisal of Drug Safety Evaluation in Animals and The Extrapolation of Results to Man. In Importance of Fundamental Principles in Drug Evaluation (Tedeschi, D.H. and Tedeschi, R.E., eds.) pp. 449-471, Raven Press, New York.

Roberts, H.R. 1981. Food Safety in Perspective. In Food Safety edited by Howard Roberts. John Wiley and Sons, Inc., New York.

Rodricks, J. and Taylor, M.R. 1983. Application of Risk Assessment to Food Safety Decision Making. Regulatory Toxicology and Pharmacology 3:273-307.

Sontag, J.M., Page, N.P., and Saffiotti. 1976. Guidelines for Carcinogen Bioassay in Small Rodents. U.S. Department of Health, Education, and Welfare Publ. No. (NIH)76-801. pp 65.

Stokinger, H.E. and Woodward, R.L. 1958. Toxicologic methods for establishing drinking water standards. J. Am. Water Works Assoc. 50:515-529.

Street, A.E. 1970. Biochemical Tests in Toxicology. In Methods in Toxicology (Paget, G.E., ed.) pp. 313-337, F.A. Davis Company, Philadelphia PA.

Stults, V.J. 1981. Nutritional Hazards. In Food Safety edited by Howard Roberts. John Wiley and Sons, Inc., New York, NY.

Tinsley, I.J. 1979. Chemical Concepts in Pollutant Behavior. John Wiley and Sons, Inc., New York.

Weill, C.S. and McCollister, D.D. 1963. Relationship Between Short and Long-term Feeding Studies in Designing an Effective Toxicity Test. Agric. Food Chem. 11:486-491.

World Health Organization (WHO). 1966. Principles for Pre-Clinical Testing of Drug Safety. WHO Tech. Rep. Ser. 341, Geneva. pp. 3-12.

World Health Organization (WHO). 1978. Principles and Methods for Evaluating the Toxicity of Chemicals: Part I. Environmental Health Criteria 6. Geneva. 273 pp.

Zbinden, G. 1963. Experimental and Clinical Aspects of Drug Toxicity. Adv. Pharmacol. 2:1-112.

Background Document on the Development and Use of Reference Doses Part II: Considerations Related to the Development of Protocols for Toxicity Studies

Prepared for:

Office Of Solid Wastes
U.S. Environmental Protection Agency
Washington, D.C.

Prepared by:

ENVIRON Corporation
1000 Potomac Street, N.W.
Washington, D.C. 20007

December 20, 1985

Table of Contents

1. Introduction

In Part I of this report, we identified the types of toxicity studies minimally acceptable for establishing ADIs. In this Part we provide a detailed examination of the protocols to be used for such tests. Specifically, we examine the options available for selecting the species and sex, species number, number of dose levels, sample size, and extent of animal examination, toward the end of identifying the minimally acceptable protocol for conducting a required toxicity test. As in the first report we set forth the options available and identify the strengths and weaknesses of each. Based on this information, and on the judgments made by various regulatory and public health bodies in the past, we recommend how decisions about experimental design should be made to achieve the maximum amount of relevant information at minimal cost.

2. Species Selection

A basic premise in toxicity testing is that laboratory animals can be used to predict toxic responses in humans (NRC, 1977). It is also widely recognized, however, that animal models are not infallible predictors of human toxic response. Qualitatively different responses are sometimes seen in different species. More often, however, different species have similar responses to a given substance and differ only in the doses which elicit the toxic response (WHO, 1978).

The reason for the differences in toxic response is most often attributed to differences in metabolism (e.g., detoxification and activation) among species. Absorption, distribution, and elimination of a given compound can also vary among species and can produce differences in toxic responses. And for reasons that are not always understood, a given target organ in different species will exhibit differences in sensitivity or response to a given substance.

One possible error in toxicity testing is choosing a test species that is less sensitive than humans to the test substances. Knowledge about the pharmacokinetics of the test substance and the availability of previous test data can reduce the chances of making this error. For example, if it is known that a particular substance is metabolized much differently in a rat than in a human, then a different species would probably be chosen. Similarly, if the rat is known to exhibit a sensitivity substantially different from humans to the anticipated effects of a substance, then another species would probably be chosen. Another way of reducing the chances of making this error is to test the substance in more than one species and to use the result from the most sensitive species, unless it is known to be substantially more sensitive than humans. The use of data from the most sensitive species is particularly popular when an NOEL is being identified. If information on the test substance itself does not exist, knowledge of the pharmacokinetics or toxic effects of structurally similar compounds may help in the choice of a test species and in the decision to test in more than one species.

Toxicity tests are usually conducted using one or two species. Occasionally, tests are conducted in three species. As a general rule, most

test protocols recommend the use of at least two species. For some tests (e.g., acute toxicity, carcinogenicity), this may be two rodent species. In others (e.g., subchronic or chronic toxicity), this may be a rodent and a non-rodent (most commonly the dog). As previously discussed, the primary reason for testing in more than one species is to identify quantitative and qualitative differences in response. Presumably, results can be extrapolated to humans with more certainty if responses in different animal species are similar. For at least some tests, however, testing in an additional species has little marginal benefit with regard to identifying either NOELs or potential toxic responses. For example, Weil and McCollister (1963) evaluated 21 chemicals that had been studied in rats, for 2 years, and also in dogs, for at least 1 year. They found that in none of the 21 cases was the dog more sensitive than the rat. In a similar study, Aviado (1978) evaluated 110 chronic studies performed on both the rat and dog and concluded that the use of both species was unnecessary.

Thus, if the purpose of toxicity testing is to estimate NOELs and to establish an ADI, the primary reason to test in more than one species of test animal is to reduce the uncertainties associated with extrapolating the results to humans. However, as mentioned above, the additional information gained from testing in a second animal species may not be very substantial, particularly if the first species is the rat and the second is the dog. Other information concerning the toxic effects or metabolism of the test substance, or structurally similar substances, can also be effective in reducing the uncertainty of extrapolating results to humans. Thus, the decision to incur the additional expense of testing in more than one species needs to be evaluated in light of what additional information the test is likely to reveal

and what is already known about the substance. For example, if the test substance produces a toxic effect for which the test species is known to have a substantially different sensitivity than humans, then testing in a second species may be warranted.

3. Selections on the Basis of Sex

There are many examples of chemicals that produce substantially different toxicities in males and females. This includes both carcinogenic and non-carcinogenic effects (Barnes and Denz, 1954; EPA, 1980b). It is the opinion of at least some toxicologists, however, that males and females of the same strain and species usually exhibit only slight differences in toxic response (Doull, 1980).

The mechanisms of all sex-related toxicity differences are not known (EPA, 1980b). Sex hormones are thought to play an important role, either by being the target or by modifying the toxic response (Chan et al., 1982). Sex-related differences in the biotransformation of foreign substances appear to be the most common reason for the development of different toxic responses (Doull, 1980). Such differences appear at puberty in some mammals (Dauterman, 1980). Differences between males and females in fat-free bodyweight and food consumption may also influence toxic responses.

Because there are many examples of sex-related differences in toxic responses, most protocols suggest or require testing in both sexes (EPA, 1978; EPA, 1979; EPA, 1980a; FDA, 1982; NAS, 1975; NRC, 1977a; OECD, 1981; WHO, 1978). Barnes and Denz (1954) suggested that it may not, however, be

necessary to test equal numbers of each sex. Until a quantitative evaluation can convincingly demonstrate that the fraction of chemicals with significantly different toxicities in males and females is very small, testing in only one sex is likely to evoke criticism of the tests' reliability.

4. Number of Dose Levels

The ability to characterize a dose-response function and to estimate a NOEL from toxicity test results are strongly affected by the number of dose levels used in the test. The dose levels tested usually range from a high dose level which produces toxic effects but minimal mortality, to a low dose level which produces no signs of toxicity (EPA, 1978; OECD, 1981; WHO, 1978). The number of intermediate dose levels tested is variable, and the testing of more intermediate levels provides a better characterization of the dose-response relationship. The clear demonstration of a dose-response relationship allows increased confidence that the relationship is not spurious (WHO, 1978). Consequently, when more intermediate dose levels are used, the estimate of the threshold dose for the test species can be made with more precision and confidence; and the human health risks extrapolated from such results can also be made with more precision (EPA, 1979).

For the reasons discussed above, virtually all protocols for subchronic and chronic toxicity tests either require (EPA, 1978) or recommend (EPA, 1979; EPA, 1980b; NAS, 1975; NRC, 1977a; FDA, 1982; FSC, 1980; OECD, 1981) that at least three dose-levels (i.e., high, low, and one intermediate) be tested. While tests for effects other than cancer have been conducted using only one or two dose levels (Barnes and Denz, 1954), there are two serious drawbacks to

using less than three dose levels. The first is that two dose levels are not sufficient for characterizing a dose-response function. The second consideration is that the use of only two doses allows little margin of error if the dose range is incorrectly chosen. For example, if the tested dose range is too high, the NOEL may be missed at the low dose, or the high dose may have such high mortality that too few survivors remain for meaningful statistical evaluation. Thus, the use of fewer than three dose levels increases the chances that a study will have to be repeated.

If only two dose levels are to be tested then it would be important to conduct a prior range-finding study. Pharmacokinetics data allowing a determination of the test substance's accumulation potential would also be useful. Even with this information, an element of luck remains (Barnes and Denz, 1954). Testing at less than three dose levels is discouraged by the fact that the additional effort and expense of testing a third dose level is small in comparison to the effect and expense of having to repeat an entire study.

Thus, the use of more dose levels can improve the characterization of the dose-response relationship and can reduce the chances of having to repeat a study. The use of more dose levels is, however, limited by practical considerations; it can either require the use of unmanageable numbers of animals or the use of small groups of animals, which may be unsatisfactory for purposes of statistical evaluation (Barnes and Denz, 1954). Unless there is special need for a particularly well characterized dose-response function, there appears to be little reason to test at other than three dose levels.

4.1 Sample Size

The number of animals used at each dose level is a compromise between the need to have a sufficient number of animals to allow adequate statistical analysis of results and practical considerations of needing to limit the cost and the workload (EPA, 1980b; EPA, 1979; WHO, 1978; Barnes and Denz, 1954). Benitz (1970) commented that the use of a large number of animals may diminish the thoroughness and care needed for a good study. It is his opinion that more useful information is obtained from a thorough study using relatively few animals than is obtained from an incomplete experiment using larger numbers of animals. An EPA-sponsored conference on the subchronic toxicity test came to a similar conclusion with its statment that the use of animals in excess of the recommended numbers would substantially increase the study cost and diminish the efficient use of facilities and personnel (EPA, 1980a).

Several current protocols and recommendations call for the use of at least 10 rodents of each sex per dose level (EPA, 1980a; OECD, 1981; Chan et al., 1982; Loomis, 1978; WHO, 1978; NTP, 1984). A few recommendations state that 20 rodents of each sex should be used (EPA, 1978; FDA, 1982; FSC, 1980). When non-rodents are used, recommendations for the number of animals of each sex to be used at each dose level fall to 3 to 8 (EPA, 1978; FDA, 1982; EPA, 1980a; FSC, 1980; Chan et al., 1982). Chronic toxicity protocols may recommend the use of more animals, particularly in carcinogenicity studies (EPA, 1980b; FDA, 1982; OECD, 1981; EPA, 1978; FSC, 1980).

The number of animals used at each dose level is a choice based primarily on practical compromise rather than on theoretical principle. In spite of this, there is a general consensus, within a fairly narrow range, on the

number of animals to be used. Using more animals than current standard practice would entail a substantial cost increase and only a marginal improvement in the ability to detect low-incidence effect (Barnes and Denz, 1954). Using fewer animals would produce a study result with less statistical confidence than has become the accepted norm.

5. Examination of Test Animals

The discovery and understanding of toxic effects depend, in large part, on the methods used to examine the test animals. Because it is not feasible to apply all known tests for toxicity or to examine every cell for signs of pathology, choices must be made concerning the methods and extent of examination of the test animals. The choices made play an important role in determining the reliability of a study. The following section discusses some of the factors to consider when making these choices. The methods discussed include hematology, blood chemistry, urinalysis, and pathology.

5.1 Hematology

Hematology tests are essential for the detection of toxic effects to the hematopoietic system. Minimum testing should provide information on cell damage and hemorrhagic effects (EPA, 1980b; NTP, 1984). This usually includes looking for signs of anemia, changes in leukocytes, and some indicator of clotting ability (Bushby, 1970).

Some of the most commonly performed hematology tests are listed in Table 2. As is shown in the table, some variation exists in what is considered to be a minimum set of hematology tests. For purposes of pesticide registration,

Table 2. Hematology Measurements Often Performed in Toxicity Investigations

Hematocrit [1,2,4,5,6]
Hemoglobin [1,2,3,4,5,6]
Erythrocyte count [1,2,4,5,6]
Total and differential leukocyte counts [1,2,3,4,5,6]
Platelet count [1,3,6]
Reticulocyte count [1]
Prothrombin time
Packed cell volume [3]
Mean corpuscle hemoglobin
Mean corpuscle volume
Methemoglobin
Thrombocyte count
Sedimentation rate [3]
Sulfhemoglobin
Examination of stained film for polychromasias and abnormal leukocytes and platelets [3]

[1] recommended by EPA (1980b)
[2] recommended by McNamara (1976)
[3] primary tests recommended by Bushby (1970)
[4] primary tests recommended by Zbinden (1963)
[5] recommended by NRC (1977a)
[6] required by EPA (1978)

the EPA (1978) requires determinations of hematocrit, hemoglobin, erythrocyte count, total and differential leukocyte counts, platelet count, and, if signs of anemia are present, reticulocyte count. In addition, the EPA requirements specify the timing of the test. For example, in a 90-day rodent study, they require hematology determinations to be made before the dosing begins, at 30-day intervals, and upon termination of dosing (EPA, 1978). A microscopic examination of the bone marrow is also included in some recommendations for minimum hematology examination because it may reveal hematotoxic effects (e.g., anemia) which often only slowly appear in the circulating blood cells (Egan et al., 1980; WHO, 1978).

As indicators of toxicity, the hematology tests are insensitive; they rarely are the effect seen at the lowest toxic dose (Weil and McCollister, 1963). Performing less than the minimum hematology determinations would leave the blood cells and clotting mechanism unexamined and would, therefore, increase the chances of missing toxic effects in them. In some instances more than the minimum tests may be indicated by effects that appear during the course of the toxicity test or by previous indications of hematotoxicity.

5.2 Blood Chemistry

Many chemical analyses can be performed on the blood as indicators of toxicity to organs, especially the liver and kidneys. Opinions vary as to the sensitivity, specificity and overall value of the individual test (Tyson et al., 1985; WHO, 1978). Fluctuations in the chemical indicators may result from transient changes in organ homeostasis rather than toxic lesions (WHO, 1978). Many toxicologists prefer that blood chemistry indicators of toxicity

be confirmed by histopathology or other evidence of toxicity (EPA, 1980b; Benitz, 1970).

Among the most commonly applied tests are serum glutamate-pyruvate transaminase (SGPT), serum glutamate-oxaloacetate transaminase (SGOT), sorbitol dehydrogenase, alkaline phosphatase, blood-urea nitrogren (BUN), and creatinine (EPA, 1980b; WHO, 1978; Environ, 1985a). The EPA Pesticide Registration protocol for subacute tests states that the following tests shall be used; calcium, potassium, serum lactate dehydrogenase (SLDH), SGPT, SGOT, glucose, BUN, direct and total bilirubin, serum alkaline phosphatase, total cholesterol, albumin, globulin, and total protein (EPA, 1978). Additional tests many also be chosen based on the institution and experience of the researcher.

While various individuals and institutions have their preferred tests, there does not appear to be a strong consensus as to what constitutes a minimum battery of blood chemistry tests. There is little advantage to eliminating very many of the tests since automated equipment allows for quick, inexpensive test results (Tyson et al., 1985); and the results can be helpful in deciding which organs to look at in the histopathology examination. It should be mentioned, however, that time spent for interpretation of the test results can be expensive (Tyson et al., 1985). Unfortunately, currently available data do not allow determining the cost-effectiveness of the individual tests (Tyson et al., 1985).

In conclusion, the blood chemistry tests can provide some indication of organ toxicity and can help guide the expensive histopathology exam. Because they are relatively inexpensive, there is little incentive to perform fewer

tests than are mentioned in the minimum test recommendations described above. The choice of individual tests will depend on the judgment of the researchers responsible for the toxicity testing.

5.3 Urinalysis

The purpose of urinalysis is to detect toxic effects in the kidneys. As with the previously discussed clinical tests, there are many specific analyses which can be performed (Environ, 1985a). Some of the most commonly used and recommended tests are listed in Table 3. There is very little consensus in the literature as to which tests are the most useful or as to the value of urinalysis in general. Individuals and institutions disagree. For example, Berndt (1976) believes that urinalysis is superior to histology and anatomical techniques for detecting nephrotoxicity, whereas Grice (1972) believes that histopathology is more sensitive than urinalysis. Similarly, the EPA (1979) states that routine urinalysis is a significant early indicator of renal damage whereas the National Academy of Science (NRC, 1977a) sees little value in routine urinalysis, particularly for subchronic tests. The objections of the National Academy of Science are primarily due to problems associated with sample collection and interpretation of results (NRC, 1977a).

It appears that urinalysis can be useful in determining the nature or specific location of nephrotoxic lesions. Its usefulness as a sensitive screen for detecting nephrotoxicity is, however, questionable. It would appear to be of most value when nephrotoxicity is suspected and the individual test would be chosen on the basis of the researchers' judgment regarding the possible lesions.

Table 3. Urinalysis Measurements
Often Performed in Toxicity Investigations

Color[3]
pH[3,10]
Volume[3,6,7]
Concentration/specific gravity[1,2,3,6,7,8,10]
Enzyme activities
- urinary glutamic oxaloacetic transaminase (UGOT)[1]
- lactate dehydrogenase[4,5]
- alkaline phosphatase[4,5]
- acid phosphatase
- glutamate dehydrogenase
- leucine aminopeptidase[5]

Protein[2,3,6,8,9,10]
Creatinine
Ketones[10]
Glucose[6,9,10]
Urobilinogen[10]
Bilirubin[9,10]
Addis count/formed elements including casts[1,2,3,9,10]

[1] recommended as most sensitive by Balazs et al. (1963, cited in EPA 1980b)
[2] recommended by Hoe and O'Shea (1965, cited in EPA 1980b)
[3] recommended by Street (1970)
[4] recommended by Wright and Plummer (1974, cited in EPA 1980b) for acute renal damage
[5] recommended by Cottrell et al. (1976, cited in EPA 1980b) for acute renal damage
[6] recommended by Berndt (1976)
[7] recommended by the National Toxicology Program (1984)
[8] described by the Food Safety Council (1980) as beneficial when renal toxicity is present
[9] recommended by World Health Organization (1978)
[10] required by EPA Pesticide Regulation Guidelines (1978)

5.4 Pathology

The pathology exam looks for changes in the structure and function of the organs and tissues of the test animals. This generally includes a visual examination of the intact organs (i.e., gross examination) as well as the examination of tissue under a microscope (i.e., histopathologic examination). The pathology exam is usually the most valuable method for detecting toxicity, a consideration which favors a more extensive examination. However, it can also be the most expensive element in a toxicology study, a consideration which favors limiting the pathology examination.

5.4.1 Gross Pathology

The gross pathology examination includes a visual examination of each animal and organ for signs of toxicity, and usually includes the weighing of individual organs. Visual examination of the organs can identify some toxic effects, supplement information from clinical tests, and help determine which tissues to examine microscopically. There is wide consensus on the need for gross examination of all organs from all animals at all dose levels (EPA, 1980b; EPA, 1978; EPA, 1979; NRC, 1977a; WHO, 1978; FSC, 1980; Barnes & Denz, 1954).

Changes in the absolute organ weight and ratio of organ weight to body weight are commonly used as indicators of possible toxicity. Weil and McCollister (1963) found changes in liver and kidney weight, for example, to be sensitive indicators of toxic effects. However, it must also be recognized that changes in organ weight may be the result of functional hypertrophy, metabolic overloading or changes in body weight, rather than the result of a specific toxic effect (EPA, 1980b). The significance of absolute and relative organ weight changes varies form organ to organ and is argued in the literature

(EPA, 1980b; Benitz, 1970). At least, organ weight change is useful as a guide when choosing organs for histopathology examination.

A common recommendation is for the weighing of all major organs, with the definition of "major" left to the judgment of the researcher. Other organs should also be include if toxic effects are suspected (EPA, 1980b). In its subchronic and chronic test protocols for pesticide registration, the EPA specifies which organs shall be weighed (EPA, 1978).

The gross examination of organs is a relatively fast, effective screen for toxic effects; reductions of this step would not result in much cost or time savings and would substantially reduce the reliability of the study.

5.4.1 Histopathology

Microscopic examination of tissue can identify toxic effects not seen in other examinations, confirm toxicity indicated by other tests or exams, and provide an indication of any dose-effect relationship. Because it is a time-consuming and expensive process, much thought has gone into ways to conduct an efficient histopathology examination. The two variables that are most often considered in discussions on efficient histopathology examinations are: 1) the number of dose levels at which all animals are to be examined, and; 2) the number of tissues to be examined per animal.

Recommendations from the National Cancer Institute for cancer bioassays called for histopathologic examination of all test and control animals, although positive controls may be exempted (Sontag et al., 1976).

Recommendations by the EPA for chronic (EPA, 1978; EPA, 1979) and subchronic (EPA, 1978) procedures repeat the recommendation that all animals be examined. Other scientific panels recommend the routine examination of all animals only in the high-dose group and in the control group (NRC, 1977a; WHO, 1978; FDA, 1982; OECD, 1981; FSC, 1980). Most of these recommendations also specify that any organs of animals at intermediate dose levels should be examined under a microscope if gross examination indicates that toxic effects may be present. Histopathological examination of the tissues of animals that die prior to the end a of study is also a common recommendation. Zbinden (1963) suggested that examining only 25-50% of the control animals would be a reasonable way to reduce the histopathology workload.

The number of tissues recommended for histopathologic examination in various necropsy protocols varies substantially. Examination of all (up to 41) tissues is called for in some protocols (EPA, 1979; EPA, 1978; Benitz, 1970) while others call for examining only selected tissues (as few as 16-18) (NTP, 1984; EPA, 1980b). The rationale for examining all tissues is to maximize the sensitivity of the test (EPA, 1979; Benitz, 1970). The rationale for examining only selected tissues is that some tissues are sufficiently poor indicators of toxicity that their examination is not justified by the cost (EPA, 1980b; Barnes & Denz, 1954).

The recommendations that only a selected number of organs be examined are based primarily on the likelihood of a positive finding, (i.e., sensitivity) (EPA, 1980b; Zbinden, 1963; NTP, 1984). Tissues that have rarely exhibited signs of toxicity in previous studies in a wide range of compounds are eliminated from the list of tissues recommended for routine microscopic

examination. Other considerations are the desire to examine tissues representing all organ systems and the frequency with which a particular tissue has been examined in past studies (EPA, 1980b). The exact number of tissues to be examined will also depend on the researcher's judgment as to any other target organs the test substance may affect. In addition to the examination of the pre-selected tissues, any tissues showing indications of toxicity (e.g. grossly observed lesions, weight chantes, etc.) are also usually recommended for histopathologic examination (EPA, 1980b). Other methods for reducing the pathology workload, such as the use of random sampling techniques have also been proposed (Fears and Douglas, 1977 and 1978).

A third approach to selecting tissues is to perform no routine examination of any particular tissue and to only look at tissues if there are other signs of toxicity (Barnes and Denz, 1954). This approach could include the evaluation of a tissue at all dose levels once an effect is seen at any dose level. For example, as a result of the gross observation of liver lesions, in the high-dose group, the livers of all animals at all dose groups would be subjected to microscopic examination. Prepared slides from all tissues and all animals could also be prepared and saved in case reason to evaluate the slides arises at a later time. Kulwich et al., (1980) and Frith et al., (1979) compared gross and microscopic examination results and found that reliance on gross examination to identify neoplastic lesions would have missed many (50% or more in some tissues) of the lesions identified when at least one histological section was examined from each organ. The nature of the lesion and the size of the organ affected the correlation between gross detection and microscopic detection of lesions (Kulwich et al., 1980; Frith et al., 1979), but the two studies illustrate the point that some toxic effects

are likely to be missed if gross examination alone is relied on to select tissues for microscopic examination.

The extensiveness of the histopathology examination represents a balance between the desire for sensitive detection of toxic effects and the high cost of histopathology examinations. Evaluation of all tissues in all animals maximizes the sensitivity of the study but is the most expensive option. Complete reliance on gross pathology to identify tissues for examination under a microscope will diminish sensitivity to a point some would consider unacceptable.

The design of an intermediate approach depends, at least in part, on the information about the test chemical that exists before testing begins. If very little is known about the possible toxic effect of the chemical, routine examination of a wider range of tissues may be warranted. If there is reason to believe the chemical affects a particular organ, a more focused histopathology design may be possible. Before the pathology component of the study can be designed, the investigator needs to know if the microscopic examination of all tissues in all animals is going to be required for the establishment of an Acceptable Daily Intake or whether a more selective approach will be acceptable. If a more selective approach will be acceptable, then the design will largely depend on the availability of pre-existing information about the chemical and the investigators' judgment.

6. Conclusions

The importance and expense of toxicity testing has stimulated much thought and discussion over the last forty years on how to make these tests more cost-effective. As a result of these efforts, each test variable in the subchronic toxicity test, for example, has a fairly well defined range within which an individual investigator must exercise judgment in designing and conducting a study for a given chemical. Even though each variable has a circumscribed range of adjustment, the cumulative effect of adjustments of all the variables can substantially influence the sensitivity and overall cost of the study.

The most important single decision in the study design is probably the extent of histopathology to be required. Pathology can account for as much as 40% of the overall cost of a toxicity study (EPA, 1979). Thus, the decision to require microscopic examination of all tissues from all animals at all dose levels or to allow a more selective approach could substantially affect the study cost.

The other factors which recommended protocols leave to the investigator's judgement can also cumulatively have a significant effect on the cost and sensitivity of the study. The previous discussion of test variables described how the information available to the investigator could affect such decisions. This information would include previous toxicity and metabolism test data when available. However, in many cases very little of this information will be available, and for these chemicals, knowledge about the toxic effect and metabolism of structurally related compounds may be valuable.

References

Aviado, D.M. 1978. Overwhelming Similarities and Minimal Differences in Toxic Responses of Rats and Dogs (Part I: 1974-1978 and Part II: 1966-1974 (Unpublished; cited in EPA, 1980b).

Barnes, J.M. and Denz, F.A. 1954. Experimental Methods used in Determining Chronic Toxicity. Pharmcol. Rev. 6:191-242.

Benitz, K.-F. 1970. Measurement of Chronic Toxicity. In Methods in Toxicology (Paget, G.E., ed.) pp. 82-131, F.A. Davis Company, Philadelphia, PA.

Berndt, W.O. 1976. Renal Function Tests: What Do They Mean? A Review of Renal Anatomy, Biochemistry, and Physiology. Environ. Hlth. Pers. 15:55-71.

Bushby, S.R.M. 1970. Haematological Studies During Toxicity Tests. In Methods in Toxicology (Paget, G.E. ed.) pp. 338-371, F.A. Davis Company, Philadelphia, PA.

Chan, D.K., O'Hara, G.P., Hayes, A.W. 1982. Principles and Methods for Acute and Subchronic Toxicity. In Principles and Methods of Toxicology (Hayes, A.W., ed.) pp. 1-51, Raven Press, New York, NY.

Dauterman, W.C. 1980. Physiological Factors Affecting Metabolism of Xenobiotics. In Introduction to Biochemical Toxicology (Hodgson, E. and Guthrie, F.W., eds.) pp. 133-142, Elsevier Science Publishing Co., Inc. New York, NY.

Doull, J. 1980. Factors Influencing Toxicology. In Casarett and Doull's Toxicology. (Doull, J., Klaassen, C.D., and Amdur, M.O., eds.) pp. 70-83. Second Edition. Macmillan Publishing Co., Inc., New York.

Egan, G.F., Lewis, S.C. and Scala, R.A. 1980. Experimental Design for Animal Toxicity Studies. In Experimental and Clinical Neurotoxicology (Spencer, P.S. and Schaumburg, H.H., eds.) Williams and Wilkins, Co., Baltimore, MD.

Environmental Protection Agency (EPA). 1978. Proposed guidelines for registering pesticides in the United States; hazard evaluation: Humans and domestic animals. Fed. Reg. 43(163):3733-37403.

Environmental Protection Agency (EPA). 1979. Support Document Test Data Development Standards: Chronic Health Effects. EPA-560/11-001.

Environmental Protection Agency (EPA). 1980a. Proceedings of the Workshop on Subchronic Toxicity Testing; Denver, Colorado, May 20-24, 1979. EPA-560/11-80-028.

Environmental Protection Agency (EPA). 1980b. Scientific Rationale for the Selection of Toxicity Testing Methods: Human Health Assessment. EPA-560/1-80-001.

Environ Corporation. 1985. Issues Relating to the Development and Use of Reference Doses. Draft Report. Prepared for EPA, Office of Solid Waste. November 12, 1985.

Environ Corporation. 1985a. Severity of Toxic Effects Ranking Scheme. Draft Report. Prepared for EPA, Environmental Criteria and Assessment Office. August 5, 1985.

Fears, T.R., and Douglas, J.F. 1977. Suggested Procedures for Reducing the Pathology Workload in a Carcinogen Bioassay Program, Part I: J. Environ. Toxicol. 1:125-137.

Fears, T.R., and Douglas, J.F. 1978. Suggested Procedures for Reducing the Pathology Workload in a Carcinogen Bioassay Program, Part II: Incorporating Blind Pathology Techniques and Analysis for Animals with Tumors. J. Environ. Toxicol. 1:211-222.

Food and Drug Administration (FDA). 1982. Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food. FDA, Bureau of Foods.

Food Safety Council. 1980. Proposed System for Food Safety Assessment. Washington, D.C. pp 160.

Frith, C.H., Booth, A.D., Greenman, D.L., and Farmer, J.H. 1979. Correlations Between Gross and Microscopic Lesions in Carcinogenic Studies in Mice. J. Environ. Pathol. Toxicol. 3:139-153.

Grice, H.C. 1972. The Changing Role of Pathology in Modern Safety Evaluation. CRC Crit. Rev. Toxicol. 1:119-152.

Kulwich, B.A. Hardisty, J.F., Gilmore, C.E., and Ward, J.M. 1980. Correlation Beteewn Gross Observations of Tumors and Neoplasms Diagnosed Microscopically in Carcinogenesis Bioassays in Rats. J. Environ. Pathol. Toxicol. 3:281-287.

Loomis, T. 1978. Essentials of Toxicology, 3rd edition. Lea and Febiger, Philadelphia, PA. pp 245.

McNamara, B.P. 1976. Concepts in Health Evaluation of Commerical and Industrial Chemicals. In New Concepts in Safety Evaluation (Mehlman, M.A., Shapiro, R.E. and Blumenthal, H., eds.) Adv. Mod. Toxicol. 1:61-140.

National Academy of Sciences. 1975. Principles for Evaluating Chemicals in the Environment. National Academy of Sciences, Washington, D.C.

National Research Council (NRC). 1977. Drinking Water and Health - Volume 1. National Academy of Sciences, Washington, D.C.

National Research Council (NRC). 1977a. Principles and Procedures for Evaluating the Toxicity of Household Substances. National Academy of Sciences, Washington, DC.

National Toxicology Program (NTP). 1984. Report of the NTP Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation. pp 280.

Organization for Economic Co-operation and Development (OECD). 1981. OECD Guidelines for Testing of Chemicals. Paris, France.

Sontag, J.M., Page, N.P., and Saffiotti. 1976. Guidelines for Carcinogen Bioassay in Small Rodents. U.S. Department of Health, Education, and Welfare Publ. No. (NIH)76-801. pp 65.

Street, A.E. 1970. Biochemical Tests in Toxicology. In Methods in Toxicology (Paget, G.E., ed.) pp. 313-337, F.A. Davis Company, Philadelphia PA.

Tyson, C.A., Story, D.L., Green, C.E., and Meierhenry, E.F. 1985 Traditionally used Indicators in Relation to Pathologic Lesions. In Organ Function Tests in Toxicity Evaluation (Tyson, C.A., and Sawhney, D.S., eds.) pp. 23-89, Noyes Publications, New Jersey.

Weill, C.S. and McCollister, D.D. 1963. Relationship Between Short and Long-term Feeding Studies in Designing an Effective Toxicity Test. Agric. Food Chem. 11:486-491.

World Health Organization. 1978. Principles and Methods for Evaluating the Toxicity of Chemicals: Part I. Environmental Health Criteria 6. Geneva. 273 pp.

Zbinden, G. 1963. Experimental and Clinical Aspects of Drug Toxicity. Adv. Pharmacol. 2:1-112.