John R. McKernan, Jr.
Governor

Rollin Ives
Commissioner

STATE OF MAINE
DEPARTMENT OF HUMAN SERVICES
AUGUSTA, MAINE 04333

POLICY FOR IDENTIFYING AND ASSESSING THE HEALTH RISKS OF TOXIC SUBSTANCES

Norman T. Anderson
Environmental Toxicology Program
Division of Disease Control
Bureau of Health
Appropriation No. 1310.1012

February, 1988

RAM TRAC Corporation

Robert A. Michaels, Ph.D., President
Toxicology & Risk Assessment Consulting

22 February 1988

Dr. Irwin Greenberg
Director
Maine Bureau of Health
State House, Station 11
Augusta, ME 04333

Dear Dr. Greenberg:

I am writing to transmit to you the Bureau's *Risk Assessment Policy* document, prepared by Norman T. Anderson and peer-reviewed by Maine's Scientific Advisory Panel (SAP). This document has evolved to its present form as an internal working document over a long period of time. It has attained a level of technical excellence, comprehensiveness, and consistency that SAP now deems appropriate for its intended use as an official guidance document for risk assessments to be conducted or judged acceptable by the Bureau of Health, including assessments conducted as part of Maine's Hazardous Air Pollutant Program. Accordingly, it is being published at this time to also serve as a resource for members of State agencies and the public seeking guidance in preparing or evaluating risk assessments in Maine.

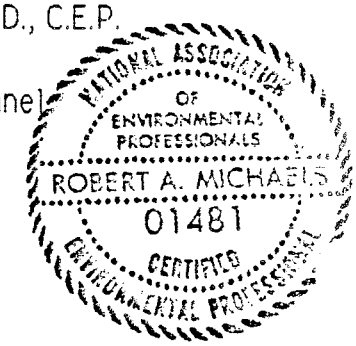
In this context, it is important for users to understand the evolutionary nature of this policy document. Its ongoing evolution is anticipated because the risk assessment process frequently poses issues which have not been addressed in the Bureau's prior assessments, and because the state of the art in risk assessment is currently advancing at a rapid pace. Neither of these sources tends to introduce changes at regular intervals. To release this document while also accommodating policy evolution, the Bureau has suggested and SAP has endorsed use of a loose-leaf format which can be updated as appropriate by occasional replacement of outdated pages. Users are, therefore, counseled to consult the Bureau to determine current dates of all pages, and acquire replacements as needed from the Bureau, prior to use of the document.

21-41 34th Avenue, Suite 11C
Long Island City, NY 11106 718-626-4947

Although no formal procedure has been established by the Bureau for soliciting public input about this policy document, SAP believes that any policy-related fact(s) brought to the Bureau's attention should be accorded due consideration as a possible basis for augmenting or altering the present document. Such responsiveness can exert only a positive effect upon the document's overall quality and its ability to serve the citizens of Maine during its continued evolution.

Very truly yours,

Robert A. Michaels, Ph.D., C.E.P.
Chairperson
Scientific Advisory Panel



cc: Norman T. Anderson
Scientific Advisory Panel members:

Barbara D. Beck, Ph.D.
Toxicologist
Gradient Corporation
44 Brattle Street
Cambridge, MA 02138

Tina Goldfine, Ph.D.
Research Biologist
Foundation for Blood Research
Route 1
Scarborough, ME 04074

Edgar Caldwell, M.D.
Pulmonary Physician
Maine Medical Center
22 Bramhall Street
Portland, ME 04102

Larry E. Mobraaten, Ph.D.
Staff Scientist
The Jackson Laboratory
Otter Creek Road
Bar Harbor, ME 04609

Ron Deprez, Ph.D.
Consultant in Epidemiology
17 West Street
Portland, ME 04102

William Sherman, Ph.D.
Environmental Engineer
Scott Paper Company
Winslow, ME 04910

Preface

As the Maine Bureau of Health has become increasingly involved with conducting and evaluating risk assessments, the need has arisen for the establishment of a risk assessment policy. The intent of this document is to identify relevant risk assessment policy issues, and to provide science policy guidance for addressing these issues. The scope of this document is broad enough to be applicable to any situation requiring risk assessments for toxic substances.

The scope of each risk assessment should be defined by its purpose. For example, the word "environment" is used extensively in this document. The term refers not just to the ambient environment, but to the total environment surrounding an organism or group of organisms. This distinction is important, as it does not limit the scope of the policy to any single exposure medium (such as air, water, or food), or to any particular category of exposure (for example, indoor air pollution in residential environments versus air pollution in occupational environments). It is possible that some assessments may require the evaluation of all exposure media and exposure categories. Conversely, other risk assessments may require only a limited evaluation of exposure, or perhaps none at all.

Sufficient latitude is also required in defining the scope of the health assessment. The health effects identified in the policy are not restricted to any disease endpoint or group of endpoints. Unless specifically stated in the assessment's purpose (for example, that the assessment will only consider carcinogenic effects), this policy allows the assessments sufficient flexibility to consider all potentially relevant parameters of adverse biological effects. Consideration of all relevant parameters may provide a better scientific insight into both the biological basis for any particular effect, or for the mechanisms of toxicity, than would be achieved by consideration of a limited number of health endpoints.

It should be continually emphasized, however, that risk assessment is a tool, not a substitute, for the determination of responsible social policy. It is an attempt to combine what is known and what is not known about a chemical into an abstract concept loosely defined as an "action level." As such, it generates a dynamic process which inevitably improves the methods by which risk assessments are conducted. It does not, unfortunately, provide much comfort when our scientific understanding fails to meet the challenges we increasingly impose upon it.

CONTENTS

SECTION III: EXPOSURE ASSESSMENT

SECTION IV: HAZARD IDENTIFICATION

SECTION V: HAZARD ASSESSMENT

SECTION VI: RISK CHARACTERIZATION

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Since the inception of its Environmental Health Unit in 1981 (M.R.S.A., 1981), the Maine Bureau of Health has become increasingly involved with conducting and evaluating risk assessments. With this responsibility has come the growing need for the Maine Bureau of Health to develop a risk assessment policy. This document describes the issues of potential concern in risk assessment, and the Bureau's policies regarding how these issues should be addressed.

The policy considerations contained in this document are guided by three basic principles: 1) that uncertainties associated with the assessment of health risk should be reflected by a conservative approach towards the protection of public health, 2) to the extent feasible, all relevant data should be evaluated in the assessment process, and 3) that risk assessments should reflect the best scientific understanding of chemically related health effects. Conformity to these principles results in a dynamic approach to risk assessment while ensuring that the public is given adequate health protection.

Risk assessments consist of four basic steps: exposure assessment, hazard identification, hazard assessment, and risk characterization. Within the first three risk assessment steps, specific parameters need to be evaluated. These parameters provide the basis for an overall characterization of risk. It is in the evaluation of these parameters that specific risk assessment policy choices must be made. Criteria are either followed or established to provide guidance in making these choices. Inadequacies identified in this process are incorporated into the recommendations for further study.

Risk assessment policy issues begin with defining the scope of the assessment and, subsequently, with the procedures by which studies are identified and selected for evaluation. The goal of identifying and reviewing all relevant data must be weighed against the limitations of data availability and the extent to which resources can be devoted to the assessment. The procedure should indicate how priority is given to studies critical to key risk assessment issues and specify those studies which were selected for review but could not be retrieved.

Studies selected for risk assessment are used to identify and evaluate key exposure and health parameters. Because this process generally requires information from several lines of investigation, it is important that the selection process identify different types of studies as well as different risk assessment parameters. Exposure parameters may need to be evaluated for different environmental media and for different exposure routes. Relevant exposure information may come from either monitoring or modelling studies. Health parameters should describe effects as a function of exposure duration and exposure dose. Relevant health information may come from epidemiological studies, controlled human exposure studies, animal studies, and cell culture studies.

Studies reviewed in the exposure assessment are used to determine how much is known about the extent and magnitude of population exposure. Studies reviewed in the hazard identification section are used to determine what is known and what needs to be known about the hazard potential of a chemical. Estimates of the health risks, which are derived in the hazard assessment section, need to consider both what is known and what is not known. Specific areas of uncertainty considered in the exposure assessment and hazard assessment sections should be identified. In addition, manner in which these uncertainties are quantitatively or qualitatively addressed should be specified. The distinction between the contributions of the empirical findings and those of the uncertainty factors to the risk estimates should be clearly made in the assessment.

The findings of the hazard assessment section are commonly expressed quantitatively in terms of action levels. Exposures greater than the action levels indicate a basis for health concern; exposures less than the action levels indicate an insignificant health risk. For threshold effects, action levels are estimates of no adverse effect levels for the general population. For non-threshold effects, such as carcinogenesis, any level of exposure is associated with some degree of risk. Action levels for non-threshold effects thus depend on the level of risk which society is willing to assume. In the absence of specific risk management policy guidance, exposure doses corresponding to lifetime cancer risks of 10^{-5}, 10^{-6}, and 10^{-7} should be presented in the risk assessment. The action level is the exposure dose corresponding to a lifetime cancer risk of 10^{-5}. In the risk characterization section, the action levels for threshold and non-threshold effects are compared with the exposure estimates to determine whether a current or projected exposure warrants a significant health concern.

The derivation of action levels relies on the use of reasonable worst case assumptions for estimating the health risks associated with chemical exposure. By using worst case assumptions, a plausible upper bound can be set on the estimation of uncertainty. Uncertainty exists in both the exposure assessment and the hazard assessment. Worst case assumptions thus need to be developed in both of these steps, as well as in the final, risk characterization step. Given the uncertainties associated with risk estimation, risk assessments may also describe approaches using less conservative assumptions. Unless these alternative assumptions reflect a greater certainty in the estimation of the actual exposure and toxicity, however, estimates based on these less extreme assumptions lack the scientific confidence necessary to ensure that the public's health is adequately protected.

After the risk assessment has been completed and sufficiently reviewed, its findings should be communicated to the appropriate agencies or individuals. While a procedure exists for chemicals evaluated in the Hazardous Air Pollutant Program, risk assessments done for other purposes have no defined risk communication procedures. The development of such procedures would enhance the effectiveness of risk assessment as a tool in public policymaking.

SECTION I: INTRODUCTION

1. Purpose of Risk Assessment.

The National Research Council (1983a) defines risk assessment as "the qualitative or quantitative characterization of potential health effects of particular substances on individuals or populations." Often in the risk assessment process, choices must be made from an array of scientifically plausible alternatives. These choices which are made in risk assessment comprise the risk assessment policy (NRC, 1983a).

The general goal of a risk assessment is to identify and evaluate the contribution of chemical exposure to the adverse human health effects. Risk assessments are conducted for a variety of reasons. They may be done to assess the health impacts of a current or anticipated exposure to a toxic substance. They may be conducted to assess chemical impacts from specific environmental media, or on specific population groups. Risk assessments may also be conducted to provide the basis for regulatory guidelines or standards. The purpose will, therefore, define the scope and focus of the risk assessment, the parameters which need to be evaluated, and the specific criteria which the assessment should address.

Once a risk assessment is completed, its findings need to be presented to interested, or potentially interested, parties. This process is called risk communication. Often, a risk assessment provide health criteria to be used in establishing public policies for particular chemicals or mixtures of chemicals. The process by which public policies are established for regulating population exposures to chemicals is called risk management (NRC, 1983a). Risk management decisions reflect a number of criteria in addition to health criteria. They include economic and political factors, statutory requirements, analytical limitations, technological feasibility, and impacts of alternative actions.

Since the inception of its Environmental Health Unit in 1981 (M.R.S.A., 1981), the Maine Bureau of Health has become increasingly involved with conducting and evaluating risk assessments. In particular, the Hazardous Air Pollutant Program (M.R.S.A., 1984) mandated the Bureau of Health to conduct risk assessments on potentially hazardous air pollutants emitted in the state. The legislation which established the program also established a peer review committee of scientists, the Scientific Advisory Panel, to provide critical review of these assessments. While carrying out their responsibilities, both the Bureau of Health staff and the Scientific Advisory Panel have realized the need for a consistent and scientifically defensible approach to risk assessment. This need has also been expressed by representatives from state regulatory agencies who request the Bureau of Health's advice on various environmental health issues. This risk assessment policy document was developed in response to these concerns.

This document has three principal uses. Structurally, it specifies the general format which the Maine Bureau of Health will use when conducting formal risk assessments. In terms of content, it identifies the specific parameters which assessments may need to consider when evaluating the health risks of toxic chemicals. Thirdly, it states the general science policy issues associated with the evaluation of these risk assessment parameters, the criteria used in their evaluation, and the general approaches recommended by the Bureau of Health when a policy judgment needs to be made. In many cases, establishment of all parameters may not be necessary. Decisions to evaluate or not evaluate parameters depend on the purpose and scope of the assessment.

The policy considerations contained in this document are guided by three basic principles. The first principle is that the uncertainties associated with the assessment of health risk should be reflected by a conservative approach towards the protection of public health. In this context, a conservative approach means that when confronted with scientific uncertainty, errors associated with the policy choices should be in the direction of increased public health protection. The second principle is that, to the extent feasible, all relevant data should be evaluated in the assessment process. The third principle is that the risk assessments should reflect the best scientific understanding of chemically related health effects.

Conformity to these choices implies a complex and dynamic approach to risk assessment policy. Specific risk assessment policy issues have been evaluated by the International Agency for Research on Cancer (IARC, 1980), National Research Council (1977a-b, 1986), the U.S. Environmental Protection Agency (USEPA, 1986a-e), the Occupational Safety and Health Administration (OSHA, 1985), and the Office on Science and Technology Policy (OSTP, 1985), among others. The criteria developed by such organizations as these provide the science policy guidance for most of the issues examined by the Maine Bureau of Health. Policy choices associated with these specific issues are modified as new scientific information becomes available.

2. Substance Identification and General Properties.

Once the purpose of the assessment is defined, the investigation of the substance begins with its chemical identification. This chapter contains information of the chemical formula, synonyms, and important identification codes for the substance being assessed. This chapter also contains a short, qualitative description of the substance's general characteristics, its major physical and chemical properties, and appropriate conversion factors. This information is useful to several areas of the assessment. In the exposure assessment, for example, such data may be helpful in determining likely routes of exposure and the substance's ability to be transferred across different environmental media. The data may also be useful in the pharmacokinetics chapter (Chapter 10), when specific data concerning partition coefficients within the body are not available. Finally, information on sustance's properties may provide assistance when describing the mechanism of action (Chapter 14.4).

SECTION II: METHODS

The procedure by which information is systematically identified, retrieved and evaluated is described in the Methods section. Although the degree of information gathering and analysis depends on the purpose of the assessment, there is a general protocol which is applicable to all assessments. First, the scope of the assessment is defined. Definition of the scope is followed by the procedure for 1) selecting the studies to be assessed, 2) how these studies are to be evaluated, and 3) how these findings of the assessment are to be communicated.

3. Scope of the Risk Assessment.

The scope of the assessment broadly determines its nature and the level of effort to be applied to it. The failure of the assessment to consider certain key areas may lead to erroneous conclusions. The assessment may also suffer from the failure to include information not directly related to its primary purpose, but which may provide insight into the overall nature of the problem. On the other hand, consideration of information not directly relevant to the assessment's purpose may create a drain on available resources and time which could better directed elsewhere. At a minimum, the scope should include those areas critical to satisfying the primary purpose of the assessment. Decisions to include additional areas should consider how such additional effort would divert resources from other environmental health assessments.

3.1 Risk Assessment Steps.

Risk assessment steps are the general areas of investigation and analysis which risk assessments should consider. According to the National Research Council, risk assessment contains one or more of the following steps: exposure assessment, hazard identification, hazard (or dose response) assessment, and risk characterization (NRC, 1983a). The extent to which these steps are investigated depends on the assessment's scope. The National Research Council's recommendations concerning how risk assessment steps should be combined and integrated into the risk management process is presented in Figure 3.1. The Maine Bureau of Health follows this general guidance when conducting risk assessments.

3.2 Risk Assessment Parameters and Criteria.

Assessment of each risk assessment step requires consideration of specific risk assessment policy issues. Resolution of these issues is important if the health and exposure parameters critical to the assessment of human health risks are to be appropriately determined. In many cases, these parameters are determined by using specific risk assessment criteria. Health parameters identified in the Hazard Identification section (Section IV) are evaluated in the Hazard Assessment section (Section V). The assessment of these parameters in combination with the exposure parameters (Section III) provides a characterization of the types and severity of potential health effects at exposure levels of concern (Section VI).

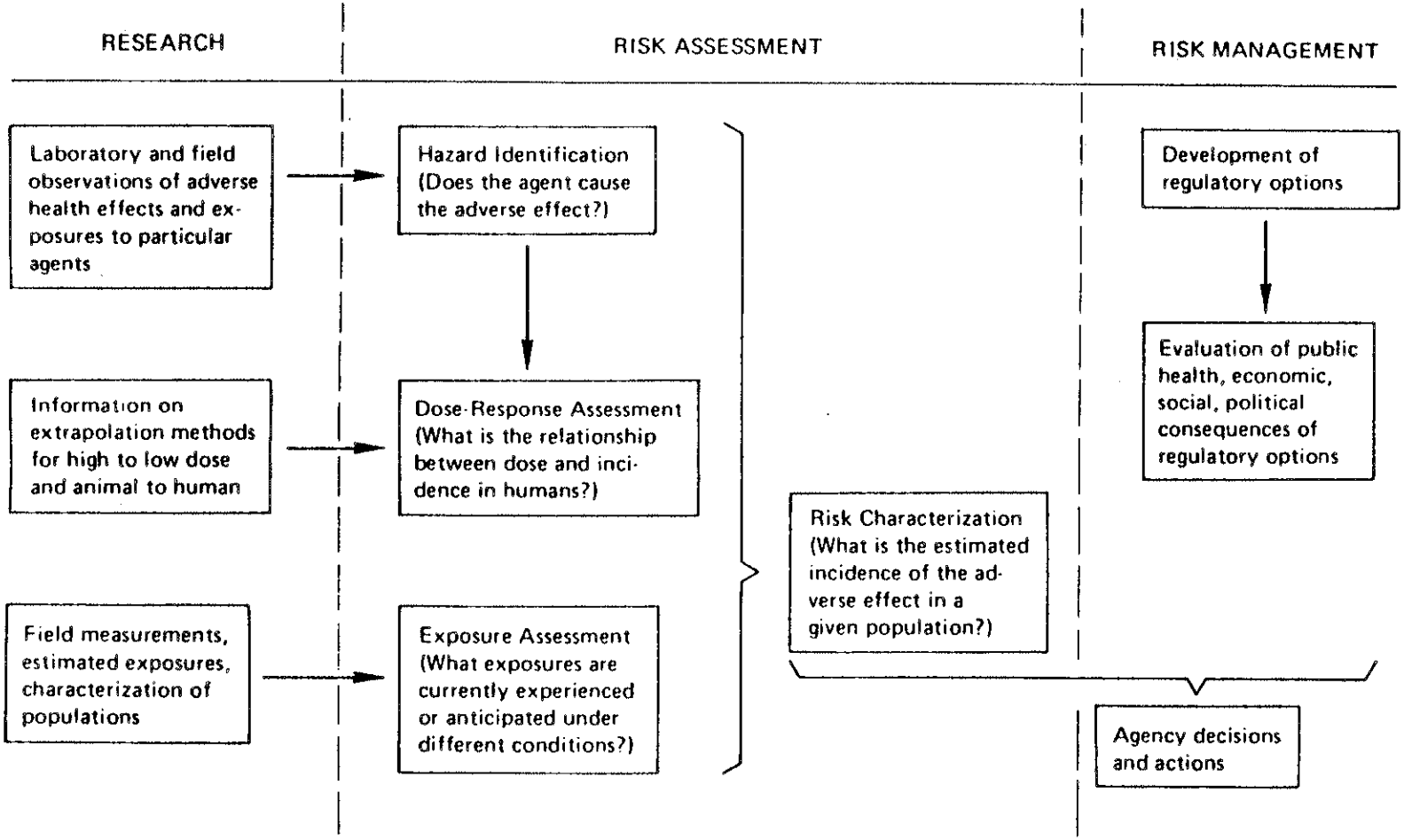


Figure 3.1. Elements of risk assessment and risk management.

Source: NRC (National Research Council), 1983a, Risk Assessment in the Federal Government: Managing the Process, National Academy Press, Washington, D.C.

The parameters identified and evaluated in the Exposure Assessment address the potential for human exposure from various emission sources and through different exposure routes, and the fate of the substance in the environment. These parameters are used in the derivation of quantitative exposure estimates and estimates of the body burdens. Criteria need to be followed or developed in order to judge the validity and precision of these exposure values.

The Hazard Assessment contains both a qualitative and a quantitative evaluation of health risk. The qualitative assessment analyzes all findings from the Hazard Identification which relate to a qualitative basis for toxicological concern. Hazard identification parameters for a particular substance are identified from studies on pharmacokinetics, health effects, environmental effects, and comparisons with other substances which have similar biological or chemical characteristics. From this information, an overall weight of evidence determination is made which relates substance exposure to potential health effects. As with the exposure assessment section, criteria should be followed or developed in this section from which to judge weight of evidence and other qualititative measures of toxicological concern.

Once the potential for a health effect to occur is established, parameters for estimating the quantitative relationship between dose and response should be determined. Included in these determinations should be an evaluation of the adequacy of the database for quantitative risk assessment purposes. If a sufficient quantitative basis exists for risk estimation, the assessment could benefit from a discussion of how different assumptions or results may influence the estimation process. This sensitivity analysis should provide an appreciation for the robustness of the estimates. Because the approach to quantitative assessment is different for threshold and non-threshold effects, such derivations must be done separately. Criteria discussed in this section concern the basis for the quantitative adjustments made to the data described in the Hazard Identification section.

In the final, Risk Characterization, section on the assessment (Section VI), estimates of the actual health risks associated with substance exposure are presented. These estimates are developed from information reviewed or derived in the previous sections. This overall assessment of health risk represents criteria from which decisions may be made regarding risk communication and risk management. Thus, the risk characterization section presents information which should be considered within the broader contexts of public health and social policy, where strictly scientific criteria are not the only areas of interest.

4. Selection of Studies.

4.1 Selection Procedure for Chemical Mixtures.

Risk assessments are generally conducted on individual chemicals. Knowledge of the toxicological properties of individual chemicals may not be sufficient indicators of toxicity, however, when chemicals are present in a mixture. Chemicals may interact in a number of ways. They may react chemically in the environment, thus possibly producing additional toxic substances (USEPA, 1986c). They may also modify each other's pharmacokinetics and activity at biological receptor sites (USEPA, 1986c).

Because these chemical interactions can be complex, the desired approach regarding the assessment of chemical mixtures is to evaluate information on the mixture as a whole. If the information on the mixture is found to be inadequate for risk assessment, subsequent efforts should be directed towards the assessment of the most toxic components of the mixture. This approach has been used in the chemical ranking system and formal risk assessment process developed for Maine's Hazardous Air Pollutant Program (Anderson, 1986).

4.2 Identification and Retrieval of Information.

Once the relevant steps and parameters of the risk assessment have been identified, the method for identifying and selecting applicable information should be described. Initial efforts usually focus on surveys of pertinent databases, bibliographic references in recent secondary sources, and information contained in the files of federal regulatory agencies. Relevant studies may then be identified from these information sources.

Risk assessment policy implications are associated with the manner in which the information sources are identified and selected. Failure to identify and retrieve critical studies may limit the ability of the assessment to reflect current scientific understanding. Consequently, findings based on a limited data set may be substantially different from those based on a more complete information base, even though the same risk assessment assumptions are applied in both cases. On the other hand, a thorough review of all relevant studies may easily overwhelm the risk assessor's capacity. In this case, such an expectation may preclude an assessment from being done in a timely fashion.

The level of effort devoted to the identification of risk assessment information is subject to the scope and purpose of the assessment, as well as to the amount of time, money, and personnel that can be committed to the search. Factors influencing this process include the number and nature of the available databases, and how they are searched. The identification process also depends on the quality and comprehensiveness of the secondary sources, the extent to which they are reviewed, and the amount and availability of information contained in the government files. The overall degree of identification is thus a function of the data availability and the resources that can be devoted to identifying the data. As the resulting level of effort may change with each assessment, the procedure by which information was identified should be described in each case.

A description of the procedure for selecting studies is also needed. The selection of studies should be consistent with the purpose and scope of the assessment. When many relevant studies have been identified, a hierarchy should be established to ensure that a manageable amount of the most important information is reviewed. Specifically, decisions must be made regarding which issues can be adequately addressed by review of secondary sources, and which issues require and in depth analysis of primary sources. As a result, the process of selecting studies is apt to be a dynamic one; as the review of the scientific literature progresses, issues are identified which warrant particular scrutiny.

In certain situations, studies critical to the assessment are not available for review. These situations may occur when the studies are considered to be proprietary information, when they have been written in a foreign language for which no translation is available, or when they are still works in progress. Problems associated with access to proprietary information may be overcome if arrangements to maintain confidentiality can be made with the sources of that information. Because the inability to retrieve information may significantly influence the findings of the assessment, mention should be made when such critical studies have been identified but could not be retrieved.

4.3 Types of Exposure Assessment Studies.

Data in the Exposure Assessment (Section 3) are analyzed in order to derive quantitative estimates of exposure for each exposure medium and exposure route. This information comes from two areas of investigation: monitoring and modelling. Precision varies with the analytical methods employed and the extent of the analysis.

Monitoring information is generally considered to be more important that modelling information, in that monitored values reflect actual exposure concentrations. Varying degrees of uncertainty are associated with these values, however. Apart from factors influencing the precision of the sampling and analytical procedures, substance levels may vary temporally and spatially. They may also be present in media not where no sampling has been conducted. These potential gaps and fluctuations introduce uncertainty as to whether or not the monitored values represent the true exposure patterns. Selection criteria should therefore address the need to reduce this uncertainty.

When monitoring data do not adequately reflect the true exposure patterns, assessments must rely on modelling information in order to estimate the extent and magnitude of exposure. Different types of modelling information information may be necessary for assessments which consider exposures from more than one pathway. Models are not able to consider all of the factors which influence the fate and transport of the substance in the environment, however. Therefore, the correlations between model results and actual exposure conditions may vary substantially from one location to the next. Consequently, selection criteria for modelling information should therefore consider the types of exposure information required in the assessments and the extent to which monitoring data must be supplemented by modelling information.

4.4 Types of Hazard Identification Studies.

Health effects information on a chemical comes from four principal areas of study: observational epidemiology, experimental epidemiology (that is, environmentally controlled studies with human volunteers), whole animal (in vivo) toxicology, and in vitro experiments. Each area of study has its own set of advantages and disadvantages, and each can provide information on an aspect of a chemical's toxicity unobtainable from other areas of research. In some instances, experiments are performed using both in vivo and in vitro procedures, such as the host mediated assay (IARC, 1980). Thus, evaluation of information from all study categories should contribute to the overall understanding of a substance's toxicity.

4.4.1 Observational Epidemiological Studies.

Observational epidemiological investigations concern the study of disease patterns in human populations. Observational studies may either be predominantly descriptive or predominantly analytical. Descriptive studies are usually undertaken when very little is known about the risk factors for a particular disease. In its simplest form, a descriptive study may involve the identification of a potential risk factor in a case or cluster of cases. Ecologic studies involve larger groups of people, most often defined geographically. Disease patterns are studied either over time or in relation to different exposure parameters. Both types of descriptive studies are useful for generating specific hypotheses regarding the association between a potential risk factor and disease. Generally, however, they do not provide enough information on the exposed and non-exposed populations from which a causal inference may be made.

Analytical studies include cross-sectional studies, case-control studies, and cohort studies (Kleinbaum et al., 1982; MacMahon and Pugh, 1970). In a cross-sectional study, prevalence rates at a particular time are compared among populations with different risk factor characteristics. Cross-sectional studies are particularly useful when the risk factors are stable over time and when the diseases are those which occur frequently and which have long durations (Kleinbaum et al., 1982). By only considering prevalence, however, cross-sectional studies cannot establish a temporal association between the risk factor and disease.

In a case-control study, individuals with a disease are compared with individuals without a disease to identify potential risk factors within the diseased individuals. A case-control study is especially useful when the disease of interest rarely occurs. It is limited, however, by several factors. Firstly, risk factor information is obtained in this type of study after the occurrence of the disease (Kleinbaum et al., 1982). Also, it is difficult to ensure that the cases and non-cases in the study population are similar with respect to extraneous risk factors (Kleinbaum et al., 1982). Finally, because it concerns a specific disease outcome, a case-control study may not be appropriate when a variety of possible health effects are being investigated (Kleinbaum et al., 1982).

In a cohort study, populations are identified in relation to an independent study variable (for example, exposed and non-exposed populations). A retrospective study investigates relevant data bases to determine whether the incidence of a particular disease differs significantly between the exposure groups. A prospective study follows these individuals over time to determine whether disease incidence varies significantly between the groups. Cohort studies, particularly prospective studies, are useful for establishing temporal associations between a risk factor and disease. On the other hand, these types of investigations are likely to be expensive, especially if the disease of interest occurs with a rare frequency. Also, certain kinds of studies, such as prospective studies, may take many years to complete. In the meantime, many people may suffer disease as a result of exposure, and many more may be at risk of developing disease even after the exposure is reduced or discontinued.

For risk assessment purposes, analytical studies are most useful for establishing causal associations between chemical exposure and disease. They thus represent the strongest evidence for a toxic effect. Establishing causal relationships is difficult, however, because there are usually other agents in the environment which have the potential to cause the same health effects. Also, differences in lifestyles and exposure patterns among members of the study and control populations preclude any precise characterization of these cohorts. In addition, exposure information is rarely determined with precision. This last shortcoming affects the ability of epidemiological studies to establish dose-response relationships in the populations under study. Because of these limitations, most epidemiological studies are inherently capable of detecting only comparatively large increases (50 percent) in the relative risks of chronic health effects (OSHA, 1980).

4.4.2 Experimental Epidemiology Studies.

Some of the problems associated with observational epidemiological studies can be eliminated if controlled studies with human volunteers are undertaken. In these studies, direct causal relationships can be more easily established because potential confounding factors are removed. Dose-response relationships can be precisely derived because the exposure to the substance is monitored and controlled.

Traditionally, controlled human exposure studies have been undertaken to test the effects of pharmaceuticals. In these cases, the study populations were likely to benefit greatly from a successful experiment. Human volunteers do not benefit from exposure to toxic, non-therapeutic chemicals, however. Thus, the use of experimental epidemiology for assessing the human health impacts of chemical contaminants is very circumscribed. For ethical reasons, therefore, experimentation with human subjects is not carried out when the chemical in question may produce irreversible effects, or at dose levels at which severe reversible effects may occur. Thus, controlled human studies have limited usefulness in risk assessment, which is mainly concerned with the irreversible effects of long-term exposure. Rather, environmental health studies using controlled human exposures are primarily designed to detect subtle and reversible acute effects.

4.4.3 Whole Animal Studies.

Whole animal (or in vivo) studies provide much of the substance-specific information on toxicity. This is particularly true when information is needed on irreversible effects, or when toxicological and pharmacological information can only be obtained by invasive means. Animal studies can be performed under carefully monitored and controlled conditions and are therefore capable of being reproduced or compared. They also can provide precise dose-response information. They are generally cheaper than large-scale epidemiological studies and can be completed in relatively short periods of time.

For new substances, toxicological studies can detect potentially irreversible effects, such as cancer, and thus serve as a screen for the introduction of potentially hazardous substances into the marketplace. For substances currently in use, they may identify harmful substances not identified through epidemiological investigations. On the other hand, significant interspecies differences may exist in a substance's toxicity, and there is generally no a priori way of knowing whether the response of any particular species is the most predictive of the response in human beings. Thus, it is possible that an effect found to occur in any laboratory animal may also occur in human beings. Similarly, if the same effect occurs in different species or strains, it is possible that the most potent response observed in the test animals is also representative of a human response. Exceptions to the use of the most sensitive species may be considered if the experiment contained significant flaws or if it can be demonstrated that the response is not qualitatively or quantitatively similar to the response in human beings. Generally, however, these findings can only be made during the course of the assessment. The study selection process, therefore, should be especially directed towards the retrieval of studies which indicate the lowest observed effect levels without regard as to whether or not the studies are relevant to human health. One initial source of this information is the Registry of Toxic Effects of Chemical Substances, or RTECS (NIOSH, 1986). RTECS lists by effect and species the study in which the most sensitive response was observed.

4.4.4 In Vitro Studies.

Given the large numbers of chemicals in and entering the marketplace, it is unlikely that existing laboratory facilities will be able, financially or logistically, to provide an adequate data base on chemical hazards (OTA, 1981). Because in vitro studies can be done quickly and inexpensively, they are critical for setting priorities on which chemicals should undergo more extensive testing, or for providing a toxicological basis for concern in the absence of adequate data on human beings or laboratory animals. Also, use of in vitro techniques can help to elucidate the biochemical basis for toxicity, and to predict structure- activity relationships. Understanding the mechanisms by which a toxic effect is produced provides the conceptual basis for interpreting the results of epidemiological or animal toxicology studies, or for predicting the toxicity of a substance in the absence of such data.

5. Analysis of Study Results.

5.1 Biological and Statistical Significance.

Health effects studies may demonstrate a positive association between chemical exposure and an adverse response. They cannot, however, prove the absence of an adverse response. At most, they can only demonstrate that the magnitude of the response in the exposed population was below the level at which the study could detect a statistically positive association. When a significantly positive association between exposure and effect is found in a properly designed and conducted study (that is, when there is a sufficiently low probability that the association may have occurred by chance), the study is referred to as a "positive" study for that health parameter. Studies which fail to demonstrate a significantly positive association are referred to as "non-positive" studies. Non-positive studies may include "negative" studies, or those studies in which clearly no positive associations were observed. There may also be studies in which positive responses or trends were suggested, but whose findings lacked statistical significance ("suggestive" studies). Finally, apart from issues of statistical significance, studies which had deficiencies in their design or conduct are limited in their abilities to produce biologically meaningful results. Depending on the nature of these deficiencies, these studies may be referred to as either "inconclusive" or "inadequate."

From a statistical standpoint, the ability of a study to detect a significantly positive association depends on the significance level used, the size of the exposed and control populations, and the background incidence of the effect. Most cancer epidemiology studies, for example, are unable to identify a positive association unless the response in the exposed population is 50 percent above the background rates (OSHA, 1980). In the case of animal bioassays, the finding of no tumors in a test population of 100 animals does not demonstrate that a zero cancer risk is associated with exposure to the chemical. Instead, a statistical analysis would demonstrate that we can be 95 percent confident that the actual incidence of tumors is no more than 4.5 percent (OTA, 1981). Thus, statistical methods may be employed on non-positive studies to estimate an upper bound on potential risk.

In addition to statistical analyses, study results may be evaluated on the basis of their biological significance. Information related to the assessment of biological significance may come directly from the epidemiological or animal bioassay studies in which the effect was investigated. It may also come from other toxicological areas, such as genetic toxicity, biochemical assays, and structure-activity relationships. As an example of such an assessment, an elevated incidence for a rare tumor in an exposed population may be judged to be biologically significant despite a lack of statistical significance relative to concurrent controls. Similarly, several studies showing non-significant elevations in the same tumor type may indicate a biologically significant effect. When considering such suggestive findings, the assessment should also be made concerning whether or not the design and conduct of the experiment were likely to increase or decrease the study's ability to detect a positive response. Conversely, a study may identify a

statistically significant association between exposure and effect which has limited biological significance. For example, a positive teratogenicity study in which the mother was administered toxic doses of a chemical may not be able to dissociate the direct effects on the embryo from the secondary effects resulting from maternal toxicity.

5.2 The Use of Worst Case Risk Assessment.

Estimates of the health risks associated with exposure to a substance depend on the assumptions employed in the risk assessment. A worst case assessment of public health risks is defined by the consistent use of reasonable worst case assumptions. If the current or projected exposure levels are below the levels derived from worst case risk assessment, then the conclusion can be drawn that, on the basis of the available information, exposure to the substance does not present a significant public health risk. If current or projected levels are at or above the levels derived from a worst case assessment, however, it may become necessary for risk assessments to more rigorously investigate the worst case assumptions. The intent of such an evaluation is to determine if less extreme assumptions are justified based on the best available data for the chemical in question. It is also possible that closer scrutiny may determine that assumptions more extreme than the worst case should be applied.

The worst case assumptions are used to set a plausible upper bound on the level of health risk associated with a particular exposure, or a plausible lower bound on the exposure level below which no significant health effects are expected to occur. If worst case assumptions are replaced by assumptions more reflective of the substance's observed toxicity, the confidence in the resulting degree of protection provided to the population depends on how closely the response of the study population reflects the response of the general population. Justification for this procedure may come from pertinent data on the exposure-response relationships, variations in population susceptibility and sensitivity, or the mechanisms of action. If greater certainty can be achieved, the confidence belt surrounding a risk estimate can be narrowed with no loss of public health protection. If, however, the confidence belt surrounding a particular estimate of risk is narrowed in the absence of greater certainty, a loss of public health protection may result.

Worst case assessment is, therefore, the primary means by which risks associated with chemical exposures are identified. If predicted exposure levels are less than those associated with health risks under worst case assumptions, a more rigorous analysis of these assumptions may be unnecessary. In the absence of a more thorough analysis, or if the data are inadequate to precisely characterize the uncertainty regarding the toxic response to exposure, the worst case risk risk assessment represents the procedure upon which public health policy recommendations must rely.

5.3 Derivation of Action Levels.

The analysis of the relationship between exposure and response results in an estimate of an "action level," or an exposure level above which there is a basis for health concern. The nature of the health concern is a function of the degree of analysis performed in the assessment, the severity and nature of the adverse effect, and the characteristics of the exposed population. Action levels are determined by considering the most sensitive effect of biological significance for the various exposure durations of concern. Action levels protective against the most sensitive effects consequently provide protection against the less sensitive effects as well. Depending on the risk extrapolation process, or the degree of uncertainty associated with the assessment of human risk, a range of potential action levels may be presented. Public health recommendations can then be developed using these action levels as a basis.

The method by which action levels are derived depends on whether or not a threshold is presumed to exist. If a threshold is presumed, exposure to a substance below a certain level should not increase health risks. Non-threshold effects may occur through even a single interaction between a substance and a biologically critical molecule, and no exposure level can be estimated for which there is an absence of risk. Most health effects are considered to exhibit thresholds, although the existence of thresholds cannot be experimentally determined solely on the basis of quantal responses from whole animal or epidemiological data (IRLG, 1986; Klaasen, 1986). A non-threshold response is presumed when an effect may occur through a single irreversible lesion, as could occur when a substance interacts with DNA. This is a presumed mechanism of action for many chemical carcinogens (Flamm and Lorentzen, 1985; OTA, 1981; USEPA, 1986a; Tomatis et al., 1982). Some carcinogens, referred to as "epigenetic" carcinogens (Shank and Barrows, 1985), may exert their effects through indirect mechanisms for which thresholds may exist. There is currently much uncertainty regarding the characterization of epigenetic carcinogens. Unless the dose-response relationship is developed to the degree that a threshold model can be presumed, therefore, action levels for carcinogenic substances should be derived using non-threshold assumptions.

For threshold effects, the action level should be the estimated no adverse effect level (ENAEL). This estimation process involves the application of various uncertainty factors and adjustment factors to a response level determined from a health effects study. For non-threshold effects, any exposure is associated with some degree of health risk. Thus, the ENAEL for a chemical carcinogen is zero unless it can be demonstrated that its effect is produced through a threshold mechanism. In some cases, such as when decisions are made concerning whether or not a chemical should be produced commercially, a qualitative finding of carcinogenicity implying a zero action level may be sufficient. In the majority of cases, however, a non-zero estimate of an action level is needed. Instead of an estimated no adverse effect level, a level is specified for non-threshold effects at which the risk is considered to be insignificant. As a general principle, a lifetime risk one per one hundred thousand is used as a reference. Risk levels above one per one hundred thousand indicate the presence of a health concern, and that decisions regarding risk communication or risk management should reflect this concern.

Science policy decisions to use assumptions less extreme than the worst case in intraspecies and interspecies extrapolations result in higher estimates for recommended action levels. Action levels would be higher, for example, if the decision is made to use the median response from animal bioassays, rather than the most sensitive reponse, as a basis for estimating human risk. Action levels would also be higher if the dose equivalency among animals is based on body weight when there is reason to believe that a surface area adjustment is warranted. In keeping with the principles of this policy, action levels based on less extreme toxicological assumptions than the worst case should be accompanied by documentation as to why these less extreme assumptions were chosen.

Action levels may also increase as the assumptions regarding the degree and duration of exposure become less extreme. Cancer risk estimates, for example, are commonly expressed in terms of continuous exposure over a 70 year lifetime. Less extreme assumptions may sometimes be warranted if the exposures are intermittent or if the exposure occurs only during a fraction of one's lifetime. Other risk assessments may evaluate the degree of risk associated with food contaminants. In these cases, worst case estimates of risk are based on assumptions of maximum contaminant levels found or predicted to be found in food. These worst case risk estimates may be lowered if data are available from which to derive more certain estimates of exposure. As with the toxicological assessment, adequate documentation is needed if action levels are based on exposure assumptions less extreme than the worst case.

It should be emphasized, however, that the cornerstone of public health policy is the prevention of disease. It is not the intention of the Maine Bureau of Health, therefore, to advocate a deterioration in environmental quality to the level of the derived action level. Conversely, if background levels of a substance are high enough to present a health concern, public health policy would indicate that exposure levels not be significantly increased over background concentrations, and even that they be decreased when it is feasible.

5.4 Peer Review of Risk Assessments.

Peer review is necessary to the proper evaluation of a scientific issue. Peer review of risk assessments is especially important in that these assessments often involve consideration of information from several scientific disciplines. Agencies or organizations conducting risk assessments rarely have expertise in all relevant disciplines. Instead, assurance that the studies are properly conducted and that sound judgments are drawn from them can be accomplished by peer review from appropriate scientists.

The Maine Bureau of Health has a peer review committee, the Scientific Advisory Panel (SAP), which was established by the same legislation which created the Hazardous Air Pollutant Program (M.R.S.A., 1984). The SAP reviews the Bureau of Health's assessments on hazardous air pollutants. The Panel may also review the health risks pertaining to other environmental issues, if such additional reviews are determined by the Bureau of Health to be justified. In carrying out its review, the SAP may request further assistance of experts in particular scientific disciplines.

Because of its principal role as a peer review committee for risk assessments, the Scientific Advisory Panel has been asked, and has agreed, to review this risk assessment policy and may review all subsequent additions and revisions. By separating the science policy issues from issues pertaining to the interpretation of scientific studies, the attention given to the development and review of risk assessments can be more clearly focused. This increase in the efficiency of the risk assessment process greatly assists the Bureau of Health and the SAP in their efforts to have the assessments reflect the best scientific understanding of the issue.

6. Risk Communication.

Risk assessment findings provide the health criteria for the assessment of environmental quality. As such, they influence both individual and societal decisions regarding the control of exposures to toxic substances. Because they must often be considered in a such contexts, these findings must be communicated in a way that is comprehensible to the public. Communication is made both to individual citizens who must make decisions regarding the management of their own health status, and to risk managers who must make societal decisions regarding acceptable health risks.

The precise communication of risk assessment findings is perhaps the most difficult part of the entire risk evaluation process. There are few areas of consensus among the scientific community on risk assessment policy issues. While this lack of agreement is beneficial to the development of better risk assessment methods, it may result in much confusion when risk assessment findings provide a basis for political decisions. Furthermore, many of the studies useful to risk assessment are not directly translatable into a description of human health risks. This limitation is particularly relevant when the health risks of low-level, long term exposures are to be estimated. Finally, given the increasing numbers of chemicals present in the environment with the potential to cause adverse effects, as well as the growing recognition of the hazards associated with accidents involving toxic substances, risk communication efforts must recognize the cumulative impacts on society's perception of risk resulting from the identification of new environmental health threats. In order for the scientific process to remain credible, the limitations of the underlying database must be described. Also, the extent of our knowledge and ignorance about a particular exposure situation should be adequately translated into an appropriate measure of health concern.

SECTION III. EXPOSURE ASSESSMENT

The objective of the exposure assessment is to characterize as fully as possible the potential for, and degree of human exposure to a substance. This includes a qualitative or quantitative estimate of the exposure level, as well as estimates of the exposure durations associated with each exposure level. As will be demonstrated later, this exposure information applies directly to the quantitative assessment and to the ultimate determination of whether or not exposure to this substance presents a public health threat.

Information on exposure may be direct or indirect. Direct information includes actual measurements of chemical concentrations in ambient or indoor environments, biological tissues, or food. Indirect sources of information include data related to the sources of chemical release into the environment. Another indirect estimate of exposure can be obtained through the use of various theoretical models.

7. Sources of Exposure

The source assessment provides information on the sources of exposure to the substance under investigation. Its purpose is to identify all potential avenues of human exposure to the substance. It begins with a general description of the substance's production and use, followed by a quantitative description of the releases of this substance into the indoor and ambient environments.

7.1 Production and Use.

Information presented in this subchapter should include a description of the amount of the substance produced and how it is produced. Special mention should be made of any production facilities in Maine. Production and use information also includes a description of how and to what extent this substance is used, again emphasizing usage information in Maine. From this information, an assessment can be made of prevalence of exposure, as well as the areas of the state where significant exposures are most likely to occur.

7.2 Emissions.

Information presented here describes the substance's ability to be emitted into various environmental media. Emissions can be described both qualitatively and quantitatively. A qualitative description involves a review of the substance's physical and chemical properties, as well as production and usage characteristics. The quantitative description is a presentation of either emission factors or actual test results regarding substance levels in air or water effluents. These kinds of quantitative data are generally unavailable, but could be very useful for modelling purposes if they are obtained. The Maine Bureau of Health recognizes the potential health risks associated with occupational environments, and the growing concern that non-occupational indoor exposure to toxic substances may present significant health risks. Thus, these exposure situations may need to be discussed in addition to the discussion of the emissions into the ambient environment.

7.2.1 Methods of Estimation or Measurement.

Quantitative estimates of human exposure to a toxic substance may depend, at least in part, on estimates of its emissions into the environment. These emissions estimates may vary according to the measurement or estimation method method employed. Thus, in developing its exposure assessment findings, the risk assessment may consider the methodologies used to produce the quantitative estimates.

Depending on the importance of the emissions estimate to the exposure assessment, a discussion of the methods used to derive emissions may require discussion. This discussion may include a description of what emission factors are available, the methods associated with the actual measurement of emissions, or how physical and chemical properties are to be evaluated when emissions estimates cannot be derived by more direct means. A discussion concerning the quality of the estimation or measurement methods should also be included.

7.2.2 Emissions into Occupational Environments.

Little quantitative information exists on substance emissions in occupational environments. Much of the available information is qualitative. Based on the physical and chemical properties of the substance, and the industrial processes involved in its use or formation, some assumptions may be made regarding emissions. Volatile substances, as well as combustion or comminution processes leading to air entrainment of particles, are of particular concern in this regard.

The usefulness of these emissions estimates depends on the methods used, and the uncertainties associated with their design and conduct. A discussion of these uncertainties may be necessary in order to place the quantitative findings into a qualitative perspective of their significance. The uncertainties associated with emissions estimates in occupational environments may be relatively unimportant if sufficient exposure monitoring data are available. When, however, exposure monitoring data are inadequate, emissions estimates from specific sources may provide valuable information for use in the estimation of exposure levels.

7.2.3 Emissions into Non-Occupational Indoor Environments.

Toxic substance emissions in non-occupational indoor environments may come from a variety of sources. Air contamination of indoor environments may result from the same volatilization, or from combustion or communition processes, just as they do in occupational environments. Volatilization may occur from either solid or liquid surfaces. Combustion emissions may result from cigarette smoking, or from utilization of gas stoves or indoor heating devices. emissions may also result through the use of a particular consumer product. Generally, quantitative information is limited with regard to toxic substance emissions of this type. Qualitative information, based on the substance's physical and chemical properties or product usage rates, may provide general guidance for the estimation of emission potential in the absence of quantitative data.

Uncertainties associated with the estimation of emissions in these indoor environments need to to discussed. This discussion is particularly relevant when the air emissions of volatile substances in drinking water are estimated. These emissions, while as yet poorly quantified, may contibute significantly to the human exposure (NRC, 1986).

7.2.4 Emissions into the Ambient Environment.

Emission factors or emission test results may be available for certain substances emitted into the ambient environment from specific sources. In addition, generic emission factors may exist from which environmental releases can be estimated in the absence of source-specific data. In general, however, few data exist for emission sources of toxic substances. Thus, the same concerns regarding the adequacy of emissions data on occupational or indoor environments apply as well to the discussion of most chemical emissions into the ambient environment.

8. Parameters for Deriving Quantitative Estimates of Exposure

8.1 Monitoring Information

Monitoring information is the key exposure component for quantitatively assessing risks presented by chemical substances. It is an important factor in the evaluation of epidemiological studies and is needed in order to assess the health significance of current exposures. Monitoring data provide actual measurements of exposure, and thus provide concrete evidence concerning the extent and magnitude of an environmental problem. For this reason, monitoring data are highly relevant to risk assessment.

Unfortunately, several limitations must be considered when evaluating monitoring information. The adequacy of the monitoring data base must be considered, therefore, when deriving conclusions concerning the correlation between exposure and effect. Monitoring data are rarely collected in any systematic way. Except when published in scientific journals, they are difficult to access. Moreover, consideration needs to be made that the presence of other substances in the environment could confound the health assessment. This is true for assessments of situations in which not all of the potentially toxic substances may have been identified, including those which must make use of historical information for chronic disease investigations.

In addition to design considerations, attention should be given to limitations in the sampling and analytical procedures. The assessment should specify negative data within the context of sampling and analytical limitations, and the degree of assurance that the monitoring was conducted in the right place or during the right time periods. For example, even if monitoring indicates the presence of a substance, it may be possible that higher levels of the substance were present in an area which was not monitored. Also, detection limits and analytical sensitivities vary for different compounds or when different sampling and analytical protocols are used.

8.1.1 Sampling and Analytical Methodologies.

There are various methodologies available for measuring substance concentrations. It is important that these methodologies be described for the particular substance under investigation. Of primary importance are the sampling time, and the accuracy, precision, reliability, and detection limits of the analysis with regard to different environmental media. Also of concern are the potential for losses during sampling and analysis and the possible interferences resulting from the presence of other substances. Many of these concerns could be addressed by suitable quality control and quality assurance procedures. Depending on the purpose and scope of the risk assessment, these procedures may warrant discussion.

8.1.2 Levels in Occupational Environments.

Monitoring data from occupational environments is used to relate exposure to the observed health effects. Relevant information includes the type of industry monitored, the type of job monitored within that industry, the range of measured values, the averaging times, the analytical methodology used, and the number of samples taken. Because many studies examine chronic effects, this criterion should also reflect historical trends in these parameters. For chemicals which show significant acute effects, such as the aggravation of asthma, exposure data should include short-term or peak concentrations. For chronic effects, such as cancer, information relevant to the determination of long-term average concentrations is needed.

8.1.3 Levels in Non-Occupational Indoor Environments.

Monitoring information from non-occupational indoor environments includes data on contaminant levels in water supplies, indoor air, and on surfaces. Contaminated water supplies have received considerable attention, as evidenced by the passage of the Safe Drinking Water Act and the development of health advisories and standards by such agencies as the U.S. Environmental Protection Agency, the National Research Council, and the World Health Organization. Attention to indoor air pollution problems has been considerably less involved. For certain chemicals, such as radon and formaldehyde, epidemiological studies have been conducted in non-occupational indoor environments (see, for example, Vaughan et al., 1986). Such monitoring data thus provides a needed exposure database for conducting risk assessments on these environments, as well as a quantitative basis for comparing the health risks in non-occupational environments to occupational settings. Even when monitoring data from non-occupational indoor environments are available, however, they are not likely to be accompanied by relevant or useful health data. Nonetheless, this monitoring information could still be used to identify potentially significant sources of exposure and, possibly, sources of potential health risk.

The presence of toxic substances in residential environments may be of particular concern with respect to infants and small children. Skin absorption of substances may be high in infants, especially premature infants, relative to other human populations (Dugard, 1987). In addition to pharmacological concerns, there are also behavioral factors associated with infants, whose oral exploration of their environment may include a wide variety of substances not normally regarded as food (WHO, 1986).

8.1.4 Levels in the Ambient Environment.

Exposure to toxic substances in the ambient environment may come from the air, water, or soil. A single emissions source may have impacts on more than one environmental medium. For example, air emissions of persistent compounds may be deposited on land, where they can accumulate in the soil and, possibly, throughout the food chain. Conversely, chemicals adsorbed onto soil particles may create an air pollution concern through particle re-entrainment.

Environmental sampling may be conducted to assess impacts from particular point sources, or to assess the general environmental quality of urban, rural, and remote locations. Ambient air concentrations for a particular substance depend on the meteorological conditions and the rate at which the substance is released into the ambient air. Thus, several samples may be needed to assess the peak and average air concentrations of a substance at a particular location. Ambient water measurements are done to assess the quality of surface or groundwater supplies. Concentrations of toxic substances in water may vary depending on the proximity of these substances to a contamination source, the rate at which they are released into the environment, and the depth at which water samples are taken. Finally, surface and subsurface soil samples may be taken. Surface samples measure levels resulting from spills, intentional land spreading, or the deposition of airborne substances on land or water. Measurement of soil contamination is especially important when considering risks to young children, who may ingest or absorb through their skin potentially harmful levels of these chemicals (WHO, 1986). Subsurface samples are taken to measure contamination resulting from leaking underground storage or waste sites, or the percolation of contaminants through the soil.

8.1.5 Levels in Food.

Certain substances have the potential to be present in food. This may result from direct or inadvertant application of pesticides, through treatment of animals with antibiotics or hormones, or through food processing. Quantitative exposure information on these substances come predominantly from the U.S. Food and Drug Administration. Other data sources may include market basket surveys from industry or state agencies. All of these data, unfortunately, are generally quite limited (GAO, 1986).

In addition to raw and processed foods, a particularly important source of food contamination is human milk (WHO, 1986). Levels of highly persistent, fat soluble chemicals may accumulate in breast milk to levels which are potentially harmful to the infant. As with chemical monitoring of foods, more data are needed to properly assess this important route exposure to potentially harmful chemicals.

8.1.6 Levels in Biological Tissues.

Compounds which bioaccumulate in biological tissues can indicate their presence when levels in other environmental media are too low to be detectable. Monitoring tissue levels of these compounds can also indicate exposures which may be missed by ambient sampling due to spatial or temporal variations in environmental levels.

In addition to data on environmental levels, monitoring can also help to provide direct estimates of body burden. These include the measurement of compounds in plants and animals consumed by human beings. Data on human tissue levels are also direct parameters of exposure. Levels of the parent compound or its metabolite in blood, exhaled air, or urine are indicative of current or past exposure. Levels of substances in fatty tissues may be used to assess cumulative exposures. These measurements are particularly useful for fat soluble compounds, such as dioxins and polychlorinated biphenyls, especially when assessing the exposure of breast fed infants to toxic substances.

8.1.7 Levels of Substances Occurring in Mixtures.

The presence of chemicals in addition to the substance of concern may also have an impact on the risk assessment. In some cases, these additional compounds may actually be responsible for the observed or suspected health effect. Should this occur, a health effect may be ascribed to a particular substance simply because the presence of these other substances was not adequately considered. This concern was raised, for example, in the interpretation of epidemiological studies which indicated a correlation between skin diseases and arsenic levels in drinking water (NAS, 1977). Similarly, if an analysis of a water supply does not indicate the presence of priority pollutants, a potential cause may be missed if the causative agent is a non-priority pollutant. A potential cause may also be missed if the pollutant scan indicates the presence of compounds that may eventually prove to be the causative agents, but which, because of their known toxicological characteristics or concentrations, are not be considered to be associated with the health effect of interest.

Consideration of exposure to other substances is also important to the assessment of possible combined or interactive effects. Exposure to one chemical may produce additive, synergistic, or antagonistic effects on the toxicity of another chemical. Monitoring information on compounds which have the potential to interact with the substance under investigation is therefore an important component in assessing the public health threat presented by a particular exposure situation.

8.2 Modelling Information.

For situations in which adequate emissions data are available, population exposures may be estimated using mathematical models. When monitoring data are available, modelling data may be used in conjunction with the actual exposure information. Modelling data may, for example, identify probable areas of high pollution impacts from air pollution sources, and thus be useful in the siting of monitoring devices. For cases in which monitoring data are not available, modelling data may provide the only quantitative estimate of exposure. Modelling data are thus particularly useful for estimating exposures resulting from anticipated emissions. They may also help in the estimation of impacts which are not measurable by available monitoring techniques.

Modelling data, however, are subject to limitations. Although they help to overcome some of the uncertainties of spatial and temporal fluctuations which limit the precision of contaminant impact estimates based on monitoring data, the reliability of modelling results also depends upon a number of assumptions. The degree of confidence placed on these assumptions depend on the quality of the input data, the appropriateness of the model to the assessment of the exposure situation of concern, and the extent to which the models have been verified in similar field situations. Of particular importance is the ability of the models to produce reasonable exposure estimates consistent with the assumptions of worst case risk assessment.

8.3 Assessment of Body Burden from Different Exposure Routes

Estimation of the total intake of a toxic substance is needed in order to determine whether a critical dose or a reference risk level is reached. The body burden criterion is thus important in the proper derivation of action levels, and in determining whether or not an action level has been exceeded in a particular exposure situation. Exposure from a combination of sources may produce a health effect or significant health risk, while exposure to any single one of these sources would not. One such circumstance would be in derivation of acceptable drinking water levels for contaminants ingested by human beings through inhalation or dermal exposure routes. Also, in the assessment of the health impacts from a source of environmental contamination, indirect as well as direct sources of exposure may need to be considered. Lead emitted in the air exhaust of a smelter, for example, will deposit on the ground, and thus may present a health risk for children playing in the surrounding area. Therefore, the health risk could be underestimated if the risk assessment evaluated only the impacts of the air emissions.

The information obtained from the monitoring and modelling subchapters may be used to estimate the body burden from different routes of exposure and from different environmental media. As was mentioned above, uncertainties are associated with both of the exposure estimation approaches. Typically, therefore, exposure estimates are presented as a range of values. Consideration of the uncertainty provides the upper and lower limits of this range.

9. Environmental Fate and Transport.

9.1 Environmental Dispersion and Chemical Transformations.

Once a substance is released into the environment from an emission source, it may undergo modifications which could influence its toxicity. Dispersion through air, water, or soil may reduce the concentrations to which human beings are exposed, and thus, the level of health risk. The substance may also be chemically transformed in these media to a form that is either less toxic or more toxic than the parent compound. The extent to which these substances persist in the environment should also be considered, as persistence influences the potential for cumulative exposure impacts over time. Knowledge of a substance's environmental persistence or its transformation products is also important in assessing the degree to which such processes may influence population exposure.

9.2 Bioconcentration and Bioaccumulation.

In addition to processes which disperse a substance once it is released into the environment, there are processes which result in its concentration. Substances such as metals and lipid soluble compounds can bioconcentrate in plant and animal tissues. Furthermore, these substances may bioaccumulate through the food chain. Bioaccumulation enhances the health risks for human beings, who occupy the highest level of the food chain. Consideration of bioconcentration and bioaccumulation is especially important when assessing long-term, low level emissions of substances into the environment; exposures which may initially appear to have negligible environmental importance may be of significant health concern when these concentrating mechanisms are evaluated.

SECTION IV. HAZARD IDENTIFICATION

Information relevant to the identification of a substance's toxicity is evaluated in the Hazard Identification section. The manifestation of an adverse health effect is a function of the concentration of the toxic substance at the site of action and its duration and biological activity at the site (Renwick, 1982; Tichy, 1983). The toxicity of a substance may vary substantially between different species, strains, or sexes. It may also vary within the same species, sex, or strain when the substance is administered through different routes of exposure or at different levels of exposure. Variability is attributable to differences in the bioavailability of a toxin, as well as to differences in the nature of the lesions formed by the interaction of the toxin with its biological receptor. The issue of bioavailability is important to risk assessment in that a toxin's concentration and duration at the site of action is related the pharmacokinetic characteristics in the organism involved. Biological activity is influenced by specific genetic and environmental factors. The kinds of biological activity required to produce and adverse health response may be simple or complex. Also, some kinds of biological activity may only require short periods of time before a toxic response occurs, whereas others may require longer periods of exposure.

The Hazard Identification, therefore, reviews data on the substance's pharmacokinetics, as well as on the kinds of biological effects it produces at various dose levels and exposure durations. In this section, criteria may need to be followed in order to determine the potential health significance of particular study findings. Such criteria are especially important when determining the statistical significance of dose-response patterns between exposed and non-exposed study populations. The overall health significance of these toxicological investigations is evaluated in the Hazard Assessment section (Section V).

10. Identification of Relevant Pharmacokinetic Parameters

Assessment of the pharmacokinetic literature is needed in order to understand how and to what extent a substance is absorbed, distributed, metabolized, and excreted. These processes determine the dose of a substance (or its metabolites) at the target site. These parameters may, in turn, be influenced by the physiological characteristics of the organism, and by the exposure regimen. Once at the target site, factors influencing cellular uptake of a chemical include its rate of diffusion across and between cellular membranes, and active or facilitated transport processes. Figure 10.1 presents a summary of the routes by which chemicals are absorbed, distributed, and excreted in the body. Figure 10.2 presents a specialized summary pertaining to the pharmacokinetic pathways in pregnant women.

Factors which affect any of the pharmacokinetic parameters may also affect a substance's toxicity. Thus, comparisons of pharmacokinetic parameters may help to identify reasons for interspecies and intraspecies differences in the toxic response to chemical exposure. These comparisons may also provide a basis for assessing the impacts of different exposure doses, durations, and routes of administration on toxicity. Such analyses are needed for the Hazard Assessment (Section V). In the Hazard Assessment, evaluations are made of

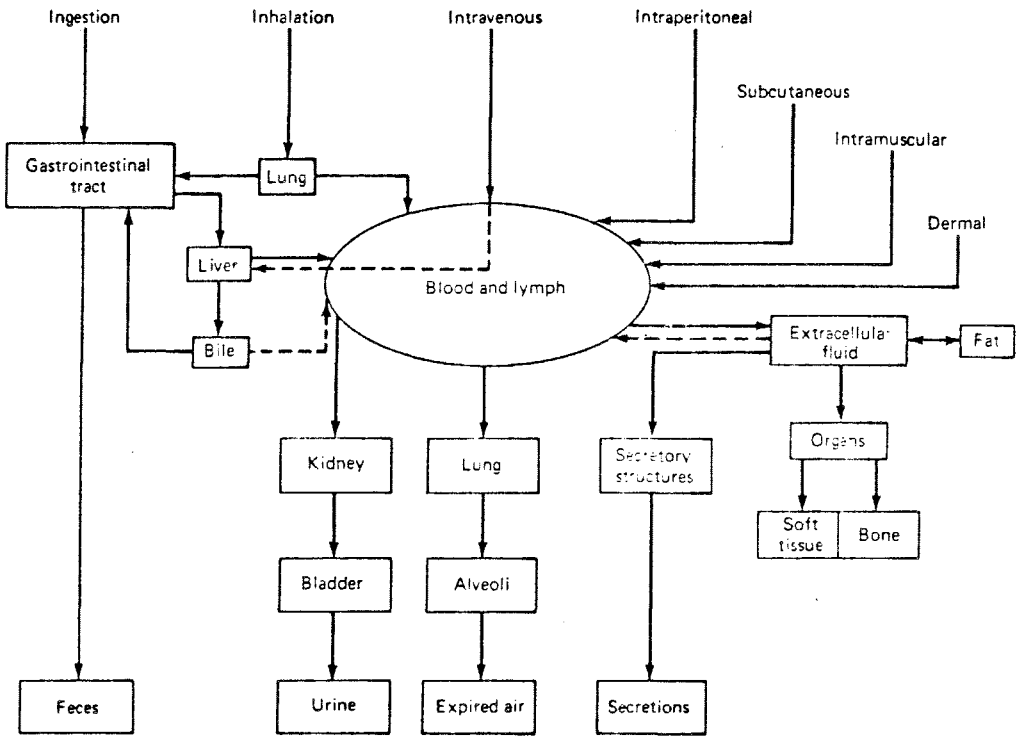


Figure 10.1. Routes of Absorption, Distribution, and Excretion of Toxicants in the Body.

Source: Klaasen, C.D., 1986, "Distribution, Excretion, and Absorption of Toxicants," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, Third Edition, MacMillan Publishing Company, New York. p. 33.

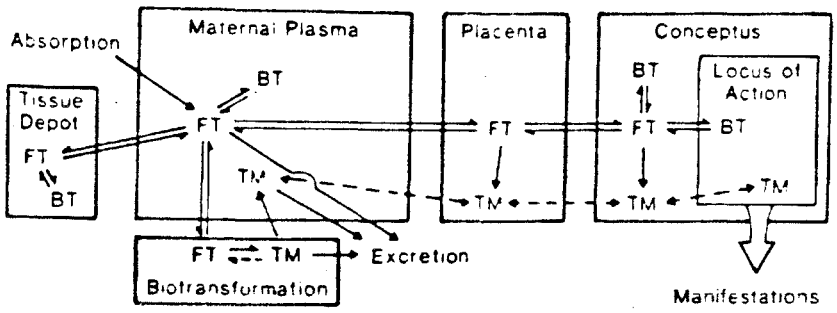


Figure 10.2. Representation of the Relation between Pharmacokinetics and Pharmacodynamics. FT = free toxicant; BT = bound toxicant; TM = toxicant metabolite.

Source: IRLG, 1982, Interagency Regulatory Liaison Group Workshop on Reproductive Toxicity Risk Assessment, Environ. Health Perspect., Vol. 66, pp. 193-221.

findings from different studies in order to determine whether or not the findings from specific studies are directly relevant to human beings at the exposure levels and exposure routes of concern.

The toxic response to a chemical exposure may be due to the action of the parent compound, one or more of its metabolites, or a combination of both. The neurotoxicity associated with exposure to trichloroethylene, for example, is caused both by exposure to trichloroethylene itself, and one of its metabolites, trichloroethanol (WHO, 1984). Of all pharmacokinetic properties, metabolism is considered to be the most variable within and between species (Rall, 1969). Therefore, intraspecies and interspecies comparisons should carefully consider differences in metabolism. Particularly, differences in the rate of metabolism and in the pathways leading to different metabolites should be analyzed.

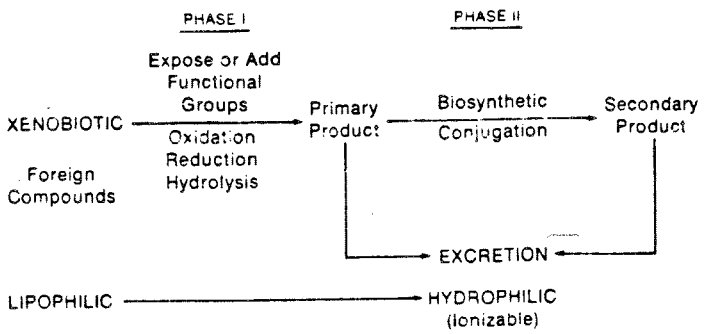


Figure 10.3. Integration of Phase I and Phase II Biotransformation Reactions.

Source: Sipes, I.G., and Gandolfi, A.J., 1986, "Biotransformation of Toxicants," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, Third Edition, MacMillan Publishing Company, New York. p. 64.

The nature and degree of metabolism depends on the concentration of the substance at the site of metabolism, the biochemical characteristics of the metabolizing tissue, and the presence of potentially modifying factors. The liver is the principal metabolizing organ, although metabolism also occurs at other sites, such as the lung, kidney, skin, and gastrointestinal mucosa (Sipes and Gandolfi, 1986). The metabolic rates and capacities may not be as high in these extrahepatic tissues. Their contribution to the total pharmacokinetic profile, however, may be significant for low level chemical exposures over long periods of time (Sipes and Gandolfi, 1986).

Metabolism is especially important with regard to lipophilic chemicals, which are easily absorbed but poorly excreted from the body. The primary purpose of metabolism, therefore, is to make these compounds more water soluble and available for excretion. The two general types of metabolic (or biotransformation) reactions are presented in Figure 10.3. The first type, or Phase I reactions, involve oxidation, reduction, or hydrolysis. In addition to making these compounds more water soluble, Phase I reactions add or expose functional groups which prepare the chemical for Phase II reactons (Sipes and Gandolfi, 1986). The chemical is conjugated with endogenous molecules in the Phase II reactions, which further increases its water solubility.

Whether or not a metabolite is toxic depends on the chemical and structural properties of the parent compound, as well as on the biochemical reactions it undergoes. In addition, other chemicals may modify the metabolism of the chemical under study. Several pathways may exist, each of which may be described by parameters concerning its enzymatic affinity and capacity. A chemical may also possess the potential for enzyme induction or inhibition. Thus, a chemical may modify its own metabolism over time through its effects on metabolizing enzymes.

11. Identification of Relevant Toxicity Endpoints.

All known or suspected adverse health effects resulting from exposure to the substance undergoing investigation should be identified. This process includes the identification of frank, clinical health effects (for example, asthma), as well as subtler indications of toxicity (for example, small changes in pulmonary function measurements). For risk assessment purposes, it is important that the most sensitive effects of exposure are identified. Identification of other, less sensitive responses should also de done. This further effort provides the Hazard Assessment section (Section V) with the basis to evaluate the constellation of health effects associated with exposure, and the progression of effects which may occur as the exposure dose increases.

In the health effects identification, therefore, the toxicological and epidemiological literature on a particular substance is reviewed. This chapter is divided into three parts: Identification of Critical Health Effects (genetic toxicity, carcinogenicity, reproductive and developmental toxicity, acute/chronic toxicity); Identification of Multiple Chemical Exposure Effects; and Identification of Sensitive Populations. The types of effects requiring the extensive analysis in most risk assessments have been genetic toxicity, carcinogenicity, and reproductive and developmental toxicity. Because these effects are of national public health significance, and have been recognized as representing particular concern with regard to toxic substance exposure, they have been segregated from the general acute/chronic toxicity category.

The classification system for assessing the different types of health effects follows the one developed by the National Institute for Occupational Safety and Health (NIOSH) to organize health effects information in its Registry of Toxic Effects of Chemical Substances (NIOSH, 1986). It also conforms to the paradigm developed by the Maine Bureau of Health for ranking hazardous air pollutants (Anderson, 1986). These four health effects categories may be further divided into subcategories when appropriate. For example, as studies more fully document the low level exposure effects regarding other health endpoints, such as immunotoxicity and neurotoxicity, this chapter should be modified accordingly.

The occurrence of a specific health effect in a study population depends largely on the exposure dose and the exposure period. The exposure conditions are thus important considerations in evaluating all health effects studies. Consequently, evaluation of these studies should be segregated on the basis of the exposure periods. The types of responses observed during these exposures can be classified according to whether they are reversible or irreversible, and as to whether they are immediate, delayed, or latent. A summary of this classification system is presented in Table 11.1. For typical animal bioassays, exposure periods follow this general classification system: acute effects (one day or less), subacute effects (one month or less), subchronic effects (one to three months), and chronic effects (more than three months) (Klaasen, 1986). These categories are generally applicable to the evaluation of human responses as well. Differences may occur, however, especially in the assessment of chronic exposure effects. Thus, when defining the exposure periods for a particular substance, the assessment should consider the underlying health database.

Table 11.1

Health Effects Categories

I. ACUTE EFFECTS

- A. Reversible
  - 1. Immediate
  - 2. Delayed
- B. Irreversible
  - 1. Immediate
  - 2. Latent

II. SUBACUTE EFFECTS

- A. Reversible
- B. Irreversible
  - 1. Immediate
  - 2. Latent

III. SUBCHRONIC EFFECTS

- A. Reversible
- B. Irreversible
  - 1. Immediate
  - 2. Latent

IV. CHRONIC EFFECTS

- A. Reversible
- B. Irreversible
  - 1. Immediate
  - 2. Latent

The inclusion of the subchapters on interactive effects and sensitive populations underscore the fact that most health effects may have several potential causes. Most of the risk assessment is devoted to the investigation of health effects with a focus on the toxicity of particular substance. The purpose of these two subchapters, on the other hand, is to evaluate a substance's toxicity within the context of actual or potential human exposure.

11.1 Identification of Critical Health Effects.

11.1.1 Identification of Effects on Genetic Material.

Genetic toxicity concerns the interaction of chemical and physical agents with the process of cellular heredity (Thilly and Call, 1986). Genes, which comprise the basic units of heredity, are comprised of varying lengths of deoxyribonucleic acid (DNA) and associated proteins. Genes are organized on chromosomes. The human somatic cell normally contains 23 pairs of chromosomes. The total length of DNA in all of the chromosomes is more than 5 billion nucleotides (NRC, 1983b). Only a small part of the DNA (about 1 percent) is required for known gene functions (NRC, 1983b). The function of most of the remaining DNA is unknown (NRC, 1983b).

The ability of an agent to interact with genetic material may lead to chronic irreversible effects such as cancer or effects on reproduction or development. There is a close correlation between genetic toxicity and the development of cancer in human beings and laboratory animals (IARC, 1980; OSTP, 1985). Genetic toxicity may also play a role in the development of other chronic diseases such as atherosclerosis (NRC, 1983b).

Three general types of genetic toxicity can be defined (NRC, 1983a; Thilly and Call, 1986). Gene mutation (point mutation) affects a single gene by producing small changes in the DNA sequence. A chromosomal mutation (clastogenesis) affects blocks of genes in one or more chromosomes. A genomic mutation (aneuploidy or polyploidy) affects the number of chromosomes without altering the chromosome structure itself. All three major types of genetic toxicity elicit DNA repair mechanisms (McQueen and Williams, 1985). These mechanisms are described in Table 11.2.

These types of genetic toxicity may occur through a variety of different mechanisms. Chemicals may damage DNA by covalent binding, intercalation, chromosomal protein binding, or by causing alterations in the synthesis and structure of DNA precursors (Bradley et al., 1985). Other types of events may occur by errors in DNA synthesis or repair (Bradley et al., 1985; OSTP, 1985), by chemically-induced alterations in the regulation of gene expression (Bradley et al., 1985; OSTP, 1985), or by stimulation via cytotoxic mechanisms of chemicals which are capable of damaging DNA (such as free radicals) (Halliwell, 1987).

Although an agent may effectively cause one kind of genetic damage, it will not necessarily cause all kinds (Thilly and Call, 1986; NRC, 1983b). This finding is understandable in light of the fact that these different types of mutations may be created by very different mechanisms. Genomic mutations, for example, are typically caused by a chemical interaction with spindle fibers, whereas gene and chromosomal mutations require chemical interactions with DNA (NRC, 1983b). Moreover, agents, such as radiation, can be very effective at breaking chromosomes and causing chromosomal mutations, but are less effective at producing changes in individual nucleotides (NRC, 1983b). Other chemicals may react with DNA bases to produce gene mutations, but are less effective at breaking chromosomes (NRC, 1983b).

Several tests for genetic toxicity have been developed to detect various types of genetic damage. A summary of these tests is presented in Table 11.3. In many cases, a substance may require metabolic activation to become biologically active. Therefore, certain metabolizing enzymes may be incorporated into in vitro genetic toxicity tests to address this possibility (WHO, 1985). Criteria for evaluating genetic toxicity studies have been developed by several academic and science policy organizations (NRC, 1983b; IARC, 1980; WHO, 1985; Jackson and Pertel, 1986). These criteria do not differ substantially from each other, except in the level of attention they devote to the specific genetic toxicity endpoints. Appropriate criteria should be followed, therefore, when the findings of these tests are to be evaluated in the risk assessment.

Table 11.2

Mammalian DNA Repair Mechanisms

Reaction	Action
Nucleotide excision repair	Removes bulky, noncoding lesions from the DNA in a manner similar to but not identical with bacterial nucleotide excision repair
Base excision repair	Permits reinsertion of the proper base into the gap left in the DNA by the action of enzymes that effect the excision of inappropriate bases from DNA as the free base (DNA glycosylases), avoiding scission of the DNA backbone, as well as a system similar to the bacterial one requiring strand scission. Included are direct demethylation by transfer of a methyl group to an acceptor protein and AP site repair (direct repair of a removed base)
Strand break repair	Rejoins single- and double-strand breaks with the addition of few or no additional nucleotides through the action of a sealing enzyme
Photoreactivation	Light-activated mechanism specific for the breakage of the UV-induced covalent bond attaching two pyrimidines in a cyclobutane-type ring
Recombination repair (post-replication repair)	System once believed to occur in mammalian systems but now controversial, by which bulky, noncoding lesions are transferred to DNA synthesized after damage
Replication bypass (post-replication repair)	Process that may function in mammalian cells as an alternative to recombination repair in which the bulky lesions are bypassed and the gap created filled
Inducible DNA repair systems	Still speculative (for mammalian systems) repair system in which DNA damage triggers the induction of enzyme systems to remove the damage

Source: OSTP (Office of Science and Technology Policy), 1985, "Chemical Carcinogens; A Review of the Science and Its Associated Principles, February, 1985," Federal Register, Vol. 50, pp. 10371-10442.

Most genetic toxicity tests are unable to measure DNA damage directly. Rather, they measure damage indirectly through observations of phenotypic changes. Close surrogates, however, to the direct measurement of DNA damage are tests which measure DNA repair. Because it is a measure of DNA repair, unscheduled DNA synthesis (UDS) is among those tests with the broadest sensitivity for detecting genotoxic chemicals (McQueen and Williams, 1985; IARC, 1980). UDS is a step in the excision repair process that is elicited by all major types of DNA damage (McQueen and Williams, 1985). Several other methods have also been identified for measuring excision repair (See Table 11.4), although UDS is regarded as the simplest and most generally applicable method for screening (IARC, 1980).

Elucidation of genetic toxicity results may be achieved through studies designed to assess the binding or biochemical interactions of chemicals with cellular macromolecules. Many carcinogens, for example, are electrophiles capable of interacting covalently with DNA (Miller et al., 1966; Miller and Miller, 1971). Every nucleoside has the potential for interacting covalently with chemicals (Singer, 1975). The interpretation of covalent binding studies, however, is subject to considerable uncertainties. The site of

Table 11.3

Summary of Commonly Used Genetic Toxicity Tests

1. Bacterial Mutation Assays (Salmonella typhimurium, Escherichia coli)
2. Yeast Cultures (Saccharomyces cerevisisiae, Schizosaccharomyces pombe)
3. Higher Plants
4. Mammalian Cells
   - Unscheduled DNA Synthesis (Human fibroblasts, HeLa Cells, Rat Liver Cells)
   - Cytogenetics and Sister Chromatid Exchanges (Chinese Hamster Ovary, Human Peripheral Lymphocytes)
   - Cell Mutation Assays (V79, Chinese Hamster Ovary, L4178Y Mouse Lymphoma)
5. Whole Animals
   - Sex-Linked Recessive Lethal Assay (Drosophila)
   - Cytogenetics: Bone Marrow Metaphase Analysis and Micronucleus Test
   - Dominant Lethal Assay

Source: Adapted from World Health Organization, 1985, Guide to Short-Term Tests for Detecting Mutagenic and Carcinogenic Chemicals, Environmental Health Criteria 51, World Health Organization, Geneva, pp. 24-125.

Table 11.4

Methods for Studying DNA Excision Repair in Cultured Cells

Incision in region of DNA damage	Alkaline sucrose gradients
	Alkaline elution
Excision of damaged region	Loss of damaged bases
	Mass spectral analysis
	Radioimmunoassay
	Loss of enzyme-sensitive sites
Resynthesis of excised region	^{3}H-thymidine incorporation autoradiography, liquid scintillation counting
	Isopycnic gradients
	Bromouracil photolysis
	BND-cellulose chromatography
Rejoining of strand	Alkaline sucrose gradients
	Alkaline elution

Source: McQueen, C.A., and Williams, G.M., 1985, "Mammalian Cell DNA Repair Assays fof Carcinogens," in Flamm, W.G., and Lorentzen, R.J., (eds.), Mechanisms and Toxicity of Chemical Carcinogens and Mutagens, Princeton Scientific Publishing Co., Princeton, New Jersey, pp. 129-151.

alkylation is a major determinant of a chemical's biological effect. Alkylation of the O6 postion of guanine, for instance, appears to be much better correlated with mutagenic and carcinogenic potential than alkylation of the N7 positions (Bradley et al., 1985). In addition, the quantitative binding of several polynuclear aromatic hydrocarbons does not show a good correlation between exposure to these compounds and species variability in susceptibility to carcinogenesis (Phillips et al., 1978; Kuroki and Heidelberger, 1971).

Findings such as those described above strongly indicate that it is the nature of the bio- logical interaction, rather than the overall extent to which genetic lesions are formed, that is toxicologically important. Furthermore, tissues may vary widely in their abilities to repair chemically-induced lesions and this varia- bility may also influence the overall tissue susceptibility to permanent genetic lesions (Bradley et al., 1985). Moreover, differences in binding and repair also exist between mammalian and non-mammalian species, as well as among different mammalian species (OSTP, 1985; Calabrese, 1983). These considerations should be taken into account when evaluating the results of genetic toxicity studies.

The limitations associated with individual genetic toxicity tests can be partially overcome through the use of a battery of tests. This approach has been called into question recently with regard to carcinogenicity screening (Tennant et al, 1987). It does, nonetheless, help to characterize the chemical's response spectrum by assessing a variety of genetic toxicity endpoints (USEPA, 1987). Given the potential seriousness of a genetic effect, any positive finding from a well conducted test should be considered significant, particularly if pharmacokinetic data are unable to demonstrate that the chemical or its active metabolite are excluded from the site of interaction with the DNA, or if the type of interaction is considered not to occur in human cells.

While genetic toxicity is correlated with certain chronic effects, information is currently lacking from which to predict human health risk solely on the basis of a chemical's action on genetic material. Conversely, an understanding of the mechanisms for most chronic diseases is not well enough developed to discount a laboratory or epidemiological finding in the face of negative genetic toxicity findings. The best current uses of genetic toxicity data, therefore, are to support the findings from other investigations, to help elucidate mechanisms of action, and to provide indications of potential concern in the absence of an adequate data base for health effects under investigation.

In the future, assessment of the toxicological impact from exposure to chemical mutagens may be enhanced by tests designed to measure in vivo genetic damage in human beings. These tests would be based on the premise that chemical mutagens produce characteristic patterns of mutations in human blood cells (Thilly and Call, 1986). This information could provide a crucial component to epidemiological studies designed to explore the relationships between chemical exposure, genetic toxicity, and irreversible diseases (Thilly and Call, 1986).

11.1.2 Identification of Carcinogenic Effects.

Cancer is a complex of diseases characterized by uncontrolled tissue growth (Flamm and Lorentzen, 1985; Berkow, 1982). It is the second leading cause of death in the country, next to heart disease. About one in five Americans die from cancer (OTA, 1981). As Table 11.5 shows, cancers of the lung, large bowel, and breast account for nearly half of all the cancer deaths in this country.

There are considerable variations in the patterns of cancer incidence among human populations (Doll and Peto, 1981; Higginson, 1980; Page and Asire, 1985). It has been estimated, largely on the basis of these variations, that 60 to 90 percent of the cancers are caused by factors in our living or working environments, and are thus theoretically preventable (Doll and Peto, 1981; Tomatis et al., 1982). These factors include man-made or natural chemical carcinogens, physical agents, radiation, viruses, nutritional deficiencies or excesses, age at reproduction, and a variety of other personal or cultural behavior patterns (OSTP, 1985; Doll and Peto, 1981; Higginson, 1980). The main causative factors have been identified for about half of the environmentally influenced cancers among men in North America and Europe, although the percentage is much lower for women (Higginson, 1981). A summary of the factors associated with the major cancers thus far is presented in Table 11.6. Despite its currently limited data base, this table demonstrates that various combinations of endogenous and environmental factors may contribute to the cancer risks.

Available information indicates that mortality rates for most cancers have remained stable throughout this century (Higginson, 1981; Young and Pollack, 1982). For the white population, however, there have been sharp decreases in the incidence of stomach and uterine cervix cancer, and increases in lung, breast, uterine corpus, and prostatic cancer (Young and Pollack, 1982). Also, there are also indications that incidence rates for bladder and kidney cancers are also increasing in males (Young and Pollack, 1982). Except for the influence of tobacco smoking on lung cancer, the reasons for these trends are not entirely clear (Young and Pollack, 1982). The apparent increases may be real. They may also be influenced by increased awareness and more efficient reporting methods (Young and Pollack, 1982). Conclusions regarding these increases, however, will have to wait until a more thorough analysis of cancer incidence trends can be conducted. In the meantime, the possibility that exposure to chemical carcinogens is responsible for increasing cancer rates cannot be discounted.

A brief description of the carcinogenic process is outlined in Figure 11.1. The diversity of causative agents and the various possible mechanisms by which cancers may be induced has prevented the establishment of any unified theory of carcinogenesis (Flamm and Lorentzen, 1985). It is generally thought, however, that multiple, independent cellular changes must accumulate before neoplastic transformation is expressed (Crawford, 1985; Higginson, 1981). This theory is supported by the fact that the induction of cancer in human beings and laboratory animals proceeds through a series of histologically distinct stages (Crawford, 1985; Williams and Weisburger, 1986). Each of these stages is subject to and controlled by a number of modifying factors (Williams and Weisburger, 1986), some or many of which may be at least partly reversible (Higginson, 1981). In addition, latency periods of at least 30 cell divisions between exposure to chemical carcinogens and the development of cancer have been observed in vitro (Berkow, 1982). These

Table 11.5

Numbers of Deaths Certified as Being Due to Various Types of Tumor: United States, 1978.

Type of tumor	No. of deaths	Percent of all deaths from tumors	
Cancer of the			
Lung[a]	95,086	24	
Large bowel (colon and rectum)	53,269	13	46
Breast	34,609	9	
Prostate	21,674	5	
Pancreas	20,777	5	
Stomach	14,452	4	46
29 other types or categories,[b] each contributing less than 3% of deaths	128,705	32	
Other or unspecified tumors[c]	33,383	8	
Total all tumors	401,955	100	

Source: Doll, R., and Peto, R., 1981, The Causes of Cancer: Quantitative Estimates of Avoidable Risks of Cancer in the United States Today, Oxford University Press, New York, p. 1197.

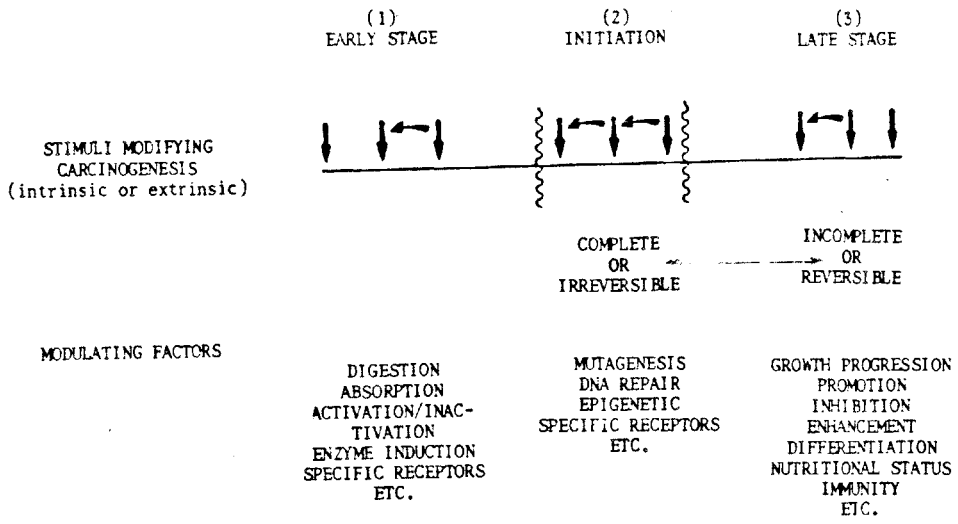


Figure 11.1. A Simplified Diagram Showing the Hypothetical Stages in Carcinogenesis and Some of the Exogenous and Endogenous Factors Possibly Involved.

Source: Higginson, J. 1981, "Rethinking the Environmental Causation of Human Cancer," Food Cosmet. Toxicol, Vol. 19, p. 544.

Table 11.6

Identified Risk Factors for the Major Cancers

Cancer Site	Risk Factors
Biliary Tract	Gallstones, hormonal changes in women
Brain and Other Nervous System Cancers	Industrial exposure to chemicals, farming, (exposure to chemicals or viruses), ranching, x-rays, head trauma, barbituates (pregnant women and children), family history
Breast	Hormonal changes in women, family history, previous breast disease, radiation, high socioeconomic status, obesity, high fat diet
Childhood cancers (leukemias, lymphomas, CNS, bone)	Genetic disorders, congenital anomalies, radiation, exposure to viruses (Burkitts lymphoma)?
Colon/Rectum	Family history, history of inflammatory bowel disease, urbanization, high fat diet
Esophagus	Tobacco and alcohol, low socioeconomic status, poor nutrition, inflammatory diseases of the esophagus
Hodgkins Disease	Exposure to viruses?
Leukemia	Genetic anomalies, exposure to radiation or industrial chemicals, exposure to viruses?
Liver	Hepatitis B infection, hormones, cirrhosis, exposure to radiation or industrial chemicals
Lung and Larynx	Tobacco, radiation, exposure to asbestos and other industrial chemicals, Vitamin A deficiency
Melanoma	Family history, hormonal factors
Multiple Myeloma	Immune system disorders, family history, exposure to radiation or industrial chemicals
Non-Hodgkins Lymphoma	Immune system disorders, exposure to viruses, exposure to pesticides
Oral Cavity and Pharynx	Tobacco, exposure to industrial chemicals
Ovary	Hormonal changes, previous breast or ovarian cancers, exposure to asbestos
Pancreas	Tobacco, diabetes?, coffee?
Prostate	High fat diet, exposure to industrial chemicals
Skin Cancer (Non melanoma)	Radiation, fair skin, family history, other diseases (tropical ulcers, burns, scars, chronic infectious, wounds)
Stomach	Diet, diseases that affect the stomach lining (pernicious anemia, atrophic gastritis), family history, radiation, Tobacco, low socioeconomic status
Testis	Congenital anomalies, hormonal drugs
Urinary Tract	Tobacco, exposure to industrial chemicals, obesity
Uterine Cervix	Multiple sex partners, age at first intercourse, venereal disease?
Uterine Corpus (Endometrium)	Same as for breast cancer

Source: Adapted from H.S. Page and A.J. Asire, 1985, Cancer Rates and Risks, Third Edition, U.S. Dept. of Health and Human Services, Washington, D.C., pp. 73-125.

findings are consistent with the observations that most cancer risks increase exponentially with age (Doll and Peto, 1981). Thus, the observation that most cancers may not appear until late in life or after long latency periods should be considered when evaluating the findings of toxicological or epidemiological studies.

The latency period associated with exposure to chemical carcinogens implies that some sort of mutational event occurs at an early stage of tumor development (Flamm and Lorentzen, 1985). This event, commonly referred to as initiation, has been well correlated with many known or suspected chemical carcinogens (OTA, 1981). It has also been confirmed by experiments which demonstrated a monoclonal origin of tumors (Yuspa and Harris, 1982).

Despite the identification of various stages and risk factors associated with carcinogenesis, many areas of uncertainty still exist as to how chemicals influence tumor development. Given the lack of knowledge regarding the underlying mechanisms of carcinogenesis (OSTP, 1985), a carcinogen or carcinogenic risk factor is identified by its association with a significantly adverse tumor response in either epidemiological or toxicological investigations. These associations are determined through the use of statistical tests which compare tumor incidences between treated and control groups, and by supporting evidence from cell culture studies.

Criteria and guidelines have been developed for deterimining the adequacy of cancer epidemiology studies (USEPA, 1986a; OSTP, 1985). Factors affecting the sensitivity of epidemiological studies include the proper selection and characterization of the exposed and control populations, the exposure duration and quality of follow-up, the proper identification and characterization of confounding factors and biases, the appropriate consideration of the latency period, the valid identification of the causes of morbidity and mortality, and the ability to detect specific tumor types (USEPA, 1986a). Statistical tests used to evaluate the significance of the exposure compare the incidence of the carcinogenic endpoint between case and control populations (for example, the odds ratio), or the relative risk of the disease between exposed and non-exposed populations (OSTP, 1985). Problems associated with the proper identification and characterization of the exposed populations, together with cost and time restraints, generally limit the sensitivity of most epidemiological studies. Thus, it is useful for these studies to include calculations of their statistical power to detect a positive response (USEPA, 1986a) and, if the results are negative, the upper confidence limits.

Various guidelines have been developed for evaluating the adequacy of cancer bioassays (USEPA, 1986a; OSHA, 1980; IRLG, 1979; OSTP, 1985; IARC, 1980; NTP, 1984; CHDS, 1985). These guidelines have been designed to ensure that the assays are well conducted, and that they have adequate sensitivity for detecting carcinogenic chemicals. Risk assessment policy issues relevant to bioassay sensitivity are described below. Issues relevant to the specificity of the bioassays, or their abilities to correctly identify non-carcinogens, are discussed in Chapter 15.5.2.

Considering the inherent limitations of epidemiological studies, many more animal carcinogens have been identified than human carcinogens. All known human carcinogens have been carcinogenic in laboratory animals (NRC, 1986). Furthermore, because animals are the closest toxicological models to human beings, results from animal bioassays are generally considered to be qualitative predictors of the human response (NRC, 1986, IARC, 1982; OSTP, 1985; OTA, 1981). Although they provide less evidence than chronic animal bioassays, in vitro cell transformation assays are also reliable qualitative indicators of carcinogenicity (IARC, 1980).

The available data have not identified any single animal species as being the most predictive model for the identification of human carcinogens (IARC, 1980). Ideally, the most appropriate animal model for predicting human carcinogenicity is one with no incidence of spontaneous tumors and a high and specific susceptibility to all human carcinogens (IARC, 1980). Because no animal species is known, it is generally recommended that the chemical be tested in at least two animal species (IARC, 1980; NTP, 1984). Most cancer bioassays have been conducted on rodents, particularly on rats, mice, and hamsters (IARC, 1980; OSTP, 1985; CDHS, 1985; IRLG, 1979). These animals have been chosen primarily for practical reasons: relatively short life span (but long enough to allow for the development of tumors), small size, and availability (CDHS, 1985; IARC, 1980; OSTP, 1985). Primates and dogs are not recommended for routine testing because their metabolic characteristics are generally no more closer to human beings than are those of rodents (IARC, 1980). Sex differences in responses to carcinogens have been found, however. Thus, it has been recommended that bioassays routine test both sexes (IARC, 1980).

Positive results from any well designed and conducted animal bioassay, therefore, may be sufficient to identify a substance as being capable of producing cancer in human beings (CDHS, 1985; USEPA, 1986a). Also, more than half of the known human carcinogens that have been adequately tested in animals produced tumors in one or more animal species at organ sites different from those produced in exposed human beings (OSHA, 1980). Thus, the specification of susceptible human tissues on the basis of bioassays of animal carcinogens is subject to significant limitations.

Most carcinogen testing has been done on inbred or hybrid animal strains. The advantages of using inbred strains is that there are generally abundant data on the tumor rates at specific organ sites in the untreated animals, and that their biological response is generally more precise and stable than those of outbred strains (OSTP, 1985). Outbred strains, on the other hand, are hardier, less expensive to maintain, less prone to genetic drift, and perhaps may more accurately reflect the human response than inbred strains (OSTP, 1985). The selection of any of these different types of strains may affect the sensitivity of the assay. It is possible for example, that a particular inbred strain may be especially resistant to the carcinogenic effects of the chemical (IARC, 1980). On the other hand, the variabilities in the responses of the outbred strains may be large enough to mask a positive response in a sensitive animal population (IARC, 1980). These issues have not been adequately resolved, although most testing continues to be done on inbred or hybrid strains. The National Toxiciology Program, for example, uses the male and female inbred Fischer 344 rat and the hybrid in B6C3F1 mouse in its bioassays (NTP, 1984). Its protocol is designed to reduce the likelihood that a carcinogen will not be identified as a result of sex and strain differences.

Concurrent control groups should be included in bioassays (IARC, 1980; USEPA, 1986a; OSTP, 1985). Use of concurrent controls reduces the potentially variability caused by different experimental conditions, fluctuations in spontaneous tumor rates, and different histopathological analyses. It has been argued that if historical rates are higher than the concurrent control rates, less weight may be given to positive results (NTP, 1984; USEPA, 1986a; Gart et al., 1979). On the other hand, it is possible that experimental conditions which depressed tumor incidence in the concurrent controls would have, in the absence of treatment effects, depressed incidence in the treated groups as well (Bickis and Krewski, 1985). Furthermore, historical tumor rates are of little value if they are unstable (Bickis and Krewski, 1985), as the variabilities in the responses are best addressed through the use of concurrent controls. Therefore, historical data should not be used to negate the significance of a positive bioassay finding unless it can be demonstrated that the concurrent controls are not suitable for analysis.

An important statistical consideration in the evaluation of bioassays is the number of animals tested. On the one hand, because these animals are serving as surrogates for approcimately 250 million people in the United States, and possible many more worldwide, and because even modest increases (for example, 0.01 - 1 percent) in cancer risks may be unacceptably high from a social welfare standpoint, large numbers of animals are needed to detect these small but significant elevations in cancer incidence. On the other hand, cost and resource concerns often limit the number of animals that can be tested (IARC, 1980; IRLG, 1979). To balance these concerns, it has been recommended that each dose and control group should contain at least fifty animals of each sex (OSTP, 1985; IARC, 1980; NTP, 1984). This initial number should be increased if interim sacrifices are planned (IARC, 1980; OSTP, 1985). With this design, the minimum detectability of the assay has been estimated to be about 10 to 15 percent (OSHA, 1980; CDHS, 1985). This is still a relatively high detection limit; however, moderate increases in sample sizes do not significantly increase the the sensitivity of the assay (OSTP, 1985; IARC, 1980).

The sensitivity of the bioassay also varies with the "spontaneous" (or background) incidence of tumors at specific tissue sites. Sites with very low spontaneous incidence rates (one percent or less) are more likely to yield false negative results than sites with higher spontaneous rates (OSTP, 1985; Fears et al., 1977; Gart et al., 1979). Five-fold, or even ten-fold increases in the tumor rate in such sites among the treated groups relative to the controls could probably go undetected under generally accepted bioassay protocols (See Table 11.7). Similar comparisons are shown in Table 11.8. For example, the false negative rate for a simple carcinogenicity screen is 87 percent when a 5 percent spontaneous tumor rate is doubled to 10 percent in the treated group. The false negative rate drops to 36 percent, however, when a 20 percent spontaneous tumor rate is doubled in the treated group. The relatively high false negative rate for sites with low spontaneous tumor rates may be counterbalanced to some degree by using historical control data. If historical control tumor rates are stable and similar to the concurrent control rates, historical control data may be used to strengthen marginally significant results at these sites (USEPA, 1986a; Gart et al., 1979; Bickis and Krewski, 1985).

Table 11.7

False Negative Error Rates for 5- and 10-Fold Increases in Spontaneous Tumor Rates for the Simple 1-Dose and 2-Dose Screens (50 Animals in the Control Group and Each of the Treated Groups)

Spontaneous Rates	5-fold increases	10-fold increases
P(FN: 1 dose)*		
0.01	0.9335	0.5761
0.02	0.6832	0.1186
0.03	0.4357	0.0157
0.04	0.2585	0.0014
0.05	0.1471	0.0001
0.06	0.0791	0.0000
0.07	0.0393	0.0000
0.08	0.0179	0.0000
0.09	0.0073	0.0000
0.10	0.0025	0.0000
P(FN: 2 doses)**		
0.01	0.9935	0.7870
0.02	0.8597	0.2073
0.03	0.6252	0.0297
0.04	0.4023	0.0028
0.05	0.2431	0.0001
0.06	0.1377	0.0000
0.07	0.0711	0.0000
0.08	0.0333	0.0000
0.09	0.0140	0.0000
0.10	0.0050	0.0000

* Probability of a false negative result with one treated group and one control. A chemical is classified as a carcinogen if it is positive at some tissue site of a 1-dose experiment with the specified sex and species.

** Probability of a false negative result with two treated groups and one control. A chemical is classified as a carcinogen if it is positive at some tissue site for both the high and low doses of a 2-dose experiment with the specified sex and species.

Source: Fears, T.R., et al., 1977, "False Positive and False Negative Rates for Carcinogenicity Screens," Cancer Res., Vol. 37, pp. 1941-1945.

Table 11.8

False Negative Rates for a Simple Carcinogenicity Screen[a]

Excess over Spontaneous Rate (%)[b]	Spontaneous Rate (%)			
	0	1	5	20
5	90	88	87	90
10	43	49	61	77
15	11	18	34	58
20	2	5	15	36
25	1	1	5	19

a. Based on Fisher's exact test with 50 animals in each of a control and test group and assuming that all animals respond independently.

b. Difference between the response rates in the test and control groups respectively.

Source: Bickis, M., and Krewski, D., 1985, "Statistical Design and Analysis of the Long-Term Carcinogenicity Bioassay," in D.B. Clayson et al. (eds.), Toxicological Risk Assessment, Vol. 1, Biological and Statistical Criteria, CRC Press, Boca Raton, FL, pp. 125-147.

Particular tumors appear with a very high spontaneous incidence in certain strains of rodents. Examples of these include lung adenomas in Strain A mice, lymphomas in strain AKR mice, liver cell tumors in C3H/HeN male mice, mammary tumors in C3H female mice, and testicular tumors in Fischer 344 male rats (IRLG, 1979; IARC, 1980). These sites are of limited usefulness for identifying increased incidences of tumors in treated animals because nearly all of the untreated controls may develop these tumors before they die. The high occurrence of spontaneous tumors at these sites also limits the applicability of total tumor bearing animals as a parameter for identifying carcinogens in chronic bioassays (IARC, 1980). This limitation may be addressed, in part, by studying the time-to-tumor response and the multiplicity of tumors at each site (IRLG, 1979). Time- to-tumor considerations may also be especially important when observing the proportion of tumor bearing animals during the earlier periods of the bioassays, before spontaneous tumors usually develop (IARC, 1980).

Because of the limited statistical sensitivity of animal bioassays, it is necessary that they be conducted at doses and under experimental conditions likely to yield the maximum tumor incidence (OSHA, 1980; NTP, 1984; CDHS, 1985; IRLG. 1979). Dosing of animals should begin soon after weaning and continue for a major portion of the animals' lifespans: 18 months for mice and hamsters and 2 years for rats (OSTP, 1985). Because of the potentially long latency periods associated with tumor development, negative results decrease

in value as the exposure and observation periods are shortened, and become practically meaningless if these periods are shorter than half the lifespans of the animals (IRLG, 1979). Little attention has been given to variations in the dosing schedules (for example, continuous versus intermittent exposures), although they may have potentially significant influences on the tumor response (NTP, 1984).

The greater the ratio between the test exposure and the human exposure, the greater is the safety margin provided by a negative result (IRLG, 1979). The doses administered to a laboratory animal strain should thus include the highest level that can be tolerated during lifetime administration without altering the animals' normal longevity from effects other than carcinogenicity (IRLG, 1979). This dose level is commonly referred to as the "maximum tolerated dose," or MTD. If the highest dose level administered is not the MTD, the sensitivity of the assay may be greatly reduced (Haseman, 1985). Levels which exceed the MTD may result in premature mortality or depressed weight gain, both of which could reduce tumor response and thus weaken the sensitivity of the assay (IARC, 1980; IRLG, 1979; NTP, 1984). The policy implications of the bioassay findings relative to increased mortality in the dosed groups are presented in Table 11.9.

Table 11.9

Interpretation of the Unadjusted Analysis of Tumor Incidence
In Light of the Survival Analysis

Outcome Type	Tumor: Association with Treatment	Mortality: Association with Treatment	Interpretation of the unadjusted test of tumor incidence[a]
A	+	+	Unadjusted test may underestimate tumorigenicity of treatment.
B	+	0	Unadjusted test gives valid picture of tumorigenicity of treatment.
C	+	-	Tumors found in treated groups may reflect longer survival of treated groups: Time-adjusted analysis is indicated.
D	-	+	Apparent negative findings in tumors may be due to shorter survival in treated groups. Time-adjusted analysis and/or retest at lower doses is indicated.
E	-	0	Unadjusted test gives a valid picture of the possible tumor-preventive capacity of treatment.
F	-	-	Unadjusted test may underestimate the possible tumor-preventive capacity of treatment.
G	0	+	High mortality in treated groups may lead to unadjusted test missing a possible tumorigen. Adjusted analysis and/or retest at lower doses is indicated.
H	0	0	Unadjusted test gives valid picture of lack of association with treatment.
I	0	-	Longer survival in treated groups may mask tumor-preventive capacity of treatment.

[a] Many of these interpretations assume that the MTD was used and that a sufficient proportion of animals survived in sufficient numbers for an appropriate length of time.

Source: Gart, J.J., et al., 1979, "Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity," J. Nat'l Cancer Inst., Vol. 62, pp. 957-974.

It has been suggested that doses in excess of the MTD may lead to tumors through either non-specific mechanisms or through mechanisms which are not operative at lower doses (See OSHA, 1980). The evidence in support of these hypotheses, however, is limited (OSHA, 1980), and sometimes even contradicted by the actual bioassay data (IRLG, 1979). Thus, the observation of significant tumor increases in animals administered a substance in excess of the MTD should be considered significant unless there is adequate evidence to the contrary. Indeed, because excessive doses are likely to decrease the sensitivity of the assay through premature mortality, a positive tumor response in this situation may increase, rather than decrease, one's concern about the substance's carcinogenic potential. Concern over decreased sensitivity can only be reduced if the carcinogenic response can be demonstrated to be an artifact of high dose administration. Generally accepted criteria for estimating the MTD from subchronic studies have been developed by the National Toxicology Program (NTP, 1984) and by the International Agency for Research on Cancer (IARC, 1980).

In addition to the MTD, bioassays should also include one or two intermediate dose levels (NTP, 1984; IARC, 1980). These additional dose levels are needed to ensure that the bioassay produces meaningful results if the MTD was exceeded. They are also useful for describing the nature of the dose-response, particularly if pharmacokinetic considerations are taken into account (NTP, 1984).

Besides statistical findings, other important toxicological information may be obtained from the cancer studies. These include the histological type of the tumor, the proportion of benign and malignant tumors and, in the case of animal bioassays, information on pre-neoplastic changes. Assessment of such qualitative information may indicate a biologically relevant response in the absence of statistical significance. Conversely, it may also find that a statistically significant result may not be biologically significant. Finally, it is hoped that qualitative information will expand the scientific understanding of the biological mechanisms associated with the development of particular tumors.

Generally, because tumors may arise from a single interaction between a chemical and DNA, no threshold dose which is free of risk can be estimated for a carcinogen which also produces genetic toxicity (NAS, 1977; USEPA, 1986a, OSHA, 1980; OSTP, 1985; OTA, 1981). Whether or not a threshold actually exists is a matter of considerable controversy. It can be argued, for example, that detoxification or DNA repair mechanisms should prevent any permanent genetic damage when the concentration of a chemical carcinogen is well below a level which would saturate these processes. On the other hand, errors may occur in this repair process. Also, threshold levels cannot be experimentally determined at present for the general population; they may vary from tissue to tissue, from individual to individual, and be influenced by other environmental agents operating through similar mechanisms. Therefore, until the biological activity of these chemicals can be better described, exposure to genotoxic carcinogens should be expressed in terms of "risk levels" rather than "safe levels" (Hogan and Hoel, 1982). Methods for estimating risks for non-threshold effects are described in the Hazard Assessment section (Section V).

There are, however, important epigenetic factors which may influence the initiation and promotion of tumors. These factors may be derived from either hereditary or environmental determinants (Higginson, 1981). The importance of initiating factors relative to epigenetic factors in tumor development is still poorly understood (Higginson, 1981). Chemicals may influence tumor development in an epigenetic fashion regardless of how they perform on genetic toxicity tests. Non-genotoxic carcinogens present special problems to the assessment of risks associated with low exposure levels. Firstly, their role in tumor development may not be detected if toxicological tests do not also include appropriate initiating factors. Secondly, even if they are identified, the conceptual basis for their mechanisms of action is different from that of the genotoxic carcinogens. Too little is known about the mechanisms of carcinogenesis, however, to adequately distinguish between the effects of genotoxic and non-genotoxic carcinogens (OSTP, 1985). Moreover, it is important to note that the three of the most potent carcinogens tested to date (2,3,7,8-tetrachorodibenzodioxin, polybrominated biphenyls, and reserpine), produce negative results in conventional genetic toxicity tests (Tennant et al., 1987). Similarly to the genotoxic carcinogens, therefore, a threshold may be assumed for these non-genotoxic carcinogens only if the critical biochemical mechanisms can be identified and quantitatively evaluated for different tissues, individuals, and species. Otherwise, in order to conform to the principle of worst case risk assessment, a non-threshold mechanism should be assumed.

Because tumor development may be dependent on multiple mechanisms, studies concentrating on complete carcinogens, or on only one factor associated with carcinogenesis, may lead to a misrepresentation of risk if potential modifying factors are not adequately identified and addressed. These factors could be better identified through improved methods of epidemiological surveillance. Such improvements could result from an expanded expanding cancer registry which includes country-wide information on exposure to potentially carcinogenic substances. Such information may include occupational history proximity to hazardous chemical industries or waste sites, and history of exposure to water or indoor air contaminants. Because of the multistage process of carcinogenisis, as well as its potentially long latency period, however, identification of human carcinogens will continue to be a difficult process. Because of the many similarities in tumor development between laboratory animals and human beings, however, much consideration should continue to be placed on the findings from animal experiments.

11.1.3 Identification of Reproductive and Developmental Effects.

An estimated one in five couples are involuntarily sterile (Dixon, 1986). Over one-third of the early embryos die and about 15 percent of recognized pregnancies abort spontaneously (Dixon, 1986). Over 1 percent of the approximately 1.3 million live infants born in the United States annually die within the first year (Manson, 1986). Seven percent have low birth weights (Manson, 1986). Approximately 2 to 4 percent are born with major congenital malformations (Manson, 1986; NRC, 1983a). Another 3 percent are found to have serious developmental effects by the end of the first year (USEPA, 1984). When defects that only become apparent later in life are included, the estimated frequency of major and minor malformations increases to about 16 percent (Manson, 1986). The contribution of toxic chemical exposure to the incidence of these adverse health effects is not well known.

Reproductive toxicity may be defined as a "dysfunction induced by chemical (as well as biological and physical) agents that affect the process of gametogenesis from its earliest stage to implantation of the conceptus in the endometrium" (Dixon, 1976). As Figure 11.2 and Table 11.10 demonstrate, successful reproduction is dependent on the interaction of several biological processes. Reproductive toxins may interfere with one or more of these processes. Direct associations between chemical exposure and reproductive toxicity are difficult to establish, however, in human populations. This is particularly true for environmental chemicals. Exposure patterns and levels are seldom estimated with precision. In addition, the influence of potentially confounding factors has to be considered in the analysis. Moreover, methods to reliably estimate damage to human fertility are not readily available (Dixon, 1986), and animal models may not be sufficiently sensitive to detect adverse reproductive effects in human beings (Dixon, 1986; IRLG, 1986). Thus, a limited number of environmental chemicals have been linked with reproductive effects in human beings.

Given the general paucity of information on this health parameter, reproductive toxicity is generally evaluated along with the assessment of developmental effects. This combination is not meant to imply that the mechanisms of toxicity are similar. Indeed, they are not, as reproductive toxicity concerns the parent and developmental toxicity concerns the conceptus. Thus, when the toxicological database is sufficient to separate reproductive effects from developmental effects, a separate subchapter should be devoted to the assessment of reproductive toxicity.

Once the conceptus is implanted in the endometrium, factors affecting intrauterine growth are referred to as "developmental toxicity" (Manson, 1986). There are four general categories of developmental toxicity: death, malformation, growth retardation, and functional deficit (Wilson, 1980; Schardein, 1985; USEPA, 1986). Developmental toxicity may result directly from the effect of a toxin on the conceptus in the absence of maternal toxicity, indirectly as a secondary effect of maternal toxicity, or from a combination of both. All categories of developmental toxicity should be investigated when assessing the potential developmental hazards, especially as there are associations between these health effects categories (Wilson, 1973). Growth retardation, for example, may be indicative of more severe developmental effects (van den Berg and Yerushalmy, 1966; Scott and Usher, 1966; Low and Galbraith, 1974). Also, the prevalence of embryolethality may be explained by the increased number of several malformations which predispose the embryo to death (Wilson, 1980). Early toxic effects to the embryo may increase the risk of lethality, possibly masking later developmental effects (Wilson, 1973).

In general, embryotoxic effects have a characteristic distribution along the dose-response curve (Wilson, 1980). At low doses, no effects are observed. As the dosage increases, both lethal and non-lethal effects begin to appear, with lethal effects becoming increasingly more prevalent at higher doses. At the upper end of the dose-response curve, the possible developmental effects of chemical exposure cannot be easily distinguished from secondary effects resulting from maternal toxicity. This relationship is described in Figure 11.3.

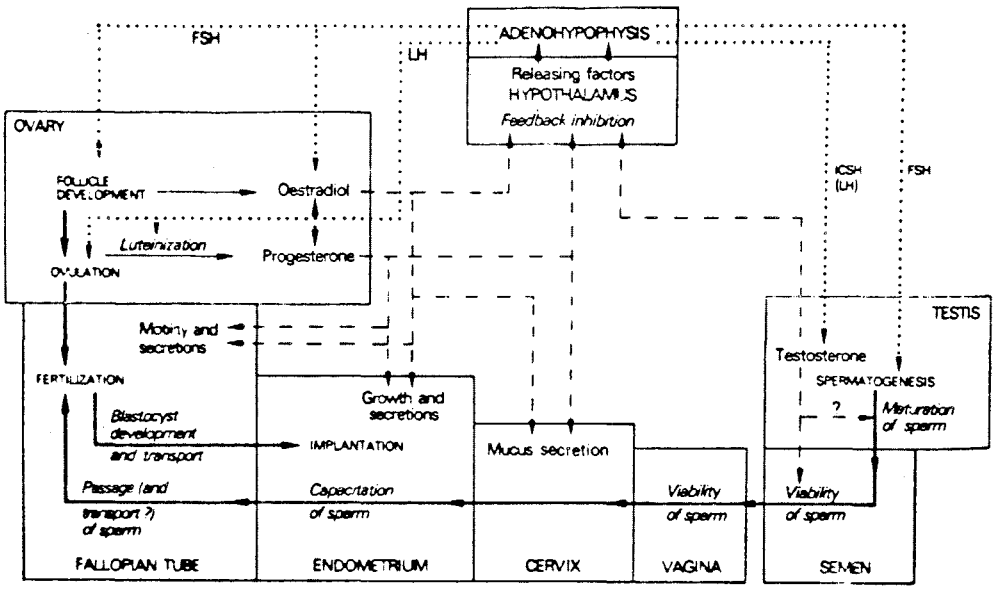


Figure 11.2. Schematic representation of the biologic processes that are essential for normal reproduction.

Source: Dixon, R.L., 1986, "Toxic Responses of the Reproductive System," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. p. 448.

Table 11.10

Various Reproductive Functions Susceptible to Toxic Chemicals

Process	
Hormonal	Endocrine Cell Hormone Synthesis Hormone Intracellular Transport and Release Feedback Mechanisms Hormone Transport and Metabolism Hormone Mediated Response Membrane Transport Cytoplasmic Receptors Nuclear Translocation Chromatin Interactions (DNA included) mRNA Protein Synthesis Structural Integrity of Protein Functional Integrity of Protein
Cellular	Maintenance Excretion Metabolism Replication Differentiation Growth
Secretory	Ion Transport Protein
Smooth Muscle Function	Microvasculature
Neurobehavioral	Tubular Function

Source: IRLG (Interagency Regulatory Liaison Group), 1986, "Interagency Regulatory Liaison Group Workshop on Reproductive Toxicity Risk Assessment," Environ. Health Perspect., Vol. 66, p. 202.

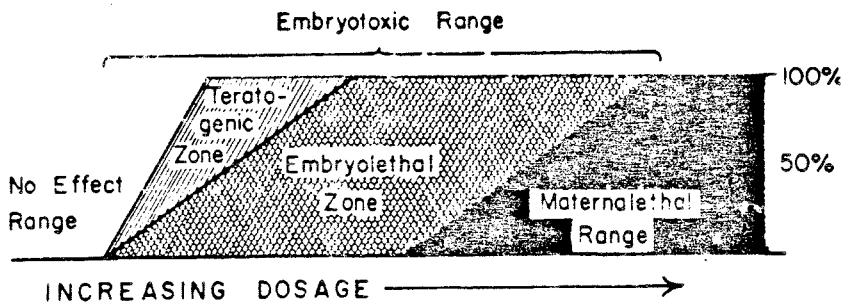


Figure 11.3. Types of Responses that Occur when Pregnant Animals Are Subjected to Increasing Dosage of a Biologically Active Chemical or Physical Agent

Source: Wilson, J.G., 1980, "Environmental Effects on Intrauterine Death in Animals," in I.H. Porter and E.B. Hook (eds.), Human Embryonic and Fetal Death, Academic Press, New York, p. 23.

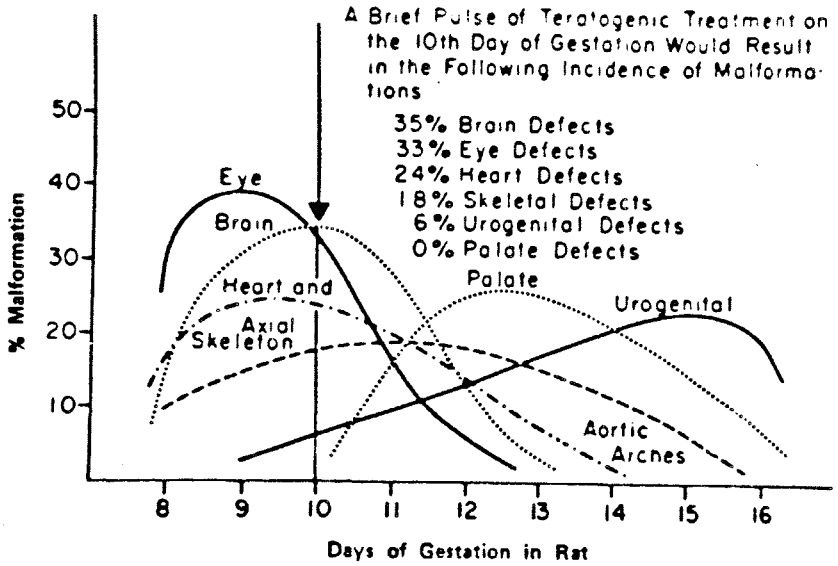


Figure 11.4. Groups of Curves Representing the Susceptibility of Particular Organs and Organ Systems in Rat Embryos to a Hypothetical Teratogenic Agent Given on Different Days of Gestation.

Source: Wilson, J.G., 1973, Environment and Birth Defects, Academic Press, New York.

Much emphasis in reproductive and developmental toxicology has been placed on the teratogenic potential of a chemical, or the potential of a chemical to produce malformations (Brown and Nigel, 1983). As Figure 11.4 indicates, organ systems differ in the intervals within the gestation periods at which they are most vulnerable to teratogenic insult. Overall, the most critical period of susceptibility ranges from shortly after the time of implantation to the end of the embryonic stage of development (approximately the first two to three months of gestation in human beings (Schardein, 1985)). It is during this period that that most organogenesis takes place. Some tissues, however, continue to differentiate beyond this time. The cerebrum and cerebellum continue differentiation even into the postnatal period (NRC, 1986), and the lung continues to differentiate throughout childhood (Kattan, 1979). The immune, reproductive, gastrointestinal, and endocrine systems are also incompletely developed at birth (WHO, 1986). These tissues, therefore, may remain sensitive to developmental toxicants well beyond the embryonic stage.

A breakdown of malformations by etiology is presented in Table 11.11. As can be seen from the table, most causes of malformations in human beings are unknown. Furthermore, subtler manifestations of developmental toxicity may also occur, but generally go unreported (USEPA, 1986d). Epidemiological studies currently lack an adequate database from which human developmental toxins may be identified or compared with animal teratogens. It is unfortunate that adequate surveillance in this area is not being implemented, as there are arguments that implementation would be relatively cost-effective (H.R., 1986). Along with the enormous public health benefits associated with the identification and control of developmental toxins, particularly those which are able to produce genetic toxicity, such efforts may also reduce risks of cancer and other chronic diseases.

The guidelines for assessing reproductive and developmental risks from epidemiological investigations are similar to those associated with the identification of carcinogens (IRLG, 1986). In light of the limited database available for epidemiological studies, the rapid association between a teratogen and its effects in human beings depends either on a appreciable number of cases at one time or place, or an unusual if not unique defect or association of defects (Wilson, 1973). Yet, the pattern of defects observed depends on the dose received as well as the time frame in which the dose is administered. Consequently, it is possible that teratogen may only be sporadically expressed in human populations or may be expressed without a well-defined pattern of effect (Wilson, 1973). Thus, it has not been recommended that epidemiological investigations depend on sentinel effects solely in the recognition of chemically-induced teratogenesis (Wilson, 1973).

Despite the lack of comparative data, there is a basis for using animal tests to identify potential human developmental toxins. Most known human teratogens have also been teratogenic in at least one animal species (See Table 11.12 (NRC, 1986; Schardein et al., 1985). The mouse and rat tests produce the highest percentage of positive responses. Also, the types of effects produced in humans and those produced in laboratory animals are correlated, as indicated by Table 11.13. Only specific anticancer drugs, anticonvulsants, and lithium failed to produce a pattern of defects in at least one animal species that is similar to the human response. (Schardein, 1985).

Table 11.11

Causes of Malformations in Human Beings

Known genetic transmission	20%
Chromosomal aberration	3-5%
Environmental causes:	
radiations	1%
infections	2-3%
maternal metabolic imbalance	1-2%
drugs and environmental chemicals	4-6%
Potentiative interactions	?
Unknown	65-70%

Source: Schardein, J.L., Chemically Induced Birth Defects, Marcel Dekker, Inc., New York, 1985. p. 2.

Table 11.12

Predictability of Laboratory Animal Models for Putative Human Teratogens[a]

Teratogen/group	Mouse	Rat	Rabbit	Hamster	Primate	Dog	Cat	Pig	Ferret	Guinea Pig
Alcohol	+	+	-		+	+		+		+
Androgenic hormones[b]	+	+	+	+	+	+		+		+
Anticancer antimetabolites[c]	±	+	+	-	±	±	±	±		+
Anticancer alkylating agents[d]	+	+	+		+				+	
Anticonvulsants[e]	+	±	+	-	±	-	-			
Coumarin anticoagulants	-	-	-					-		
Antithyroid agents[b]	+	+	+							+
Progestrogenic hormones[b]	+	±	±			+				+
DES[b]	+	+	-	±	+				+	
Methylmercury	+	+	-	+	-	-	+	-		
Thalidomide	±	±	+	±	+	±	±	±	+	-
Lithium	+	-	-		+			-		
D-Penicillamine		+		+						
Streptomycin antibiotics[b]	-	+	-							-
Vitamin A analogues	+	+	+	+	+	+		+		+

[a] Legend: (+) teratogenic; (±) variably teratogenic; (-) not teratogenic.
[b] Defects related to functional activity.
[c] Includes azauridine, aminopterin, fluorouracil, methotrexate, cytarabine.
[d] Includes busulfan, chlorambucil, cyclophosphamide, and mechlorethamine.
[e] Includes hydantoin and dione groups, and valproate.

Source: Schardein, J.L., et al., 1985, "Species Sensitivities and Prediction of Teratogenic Potential," Environ. Health Perspect., Vol. 61, p. 60.

Although there was a concordance found between the effects in human beings and the effects in some laboratory animal, there was far less concordance when the responses of individual species were compared to the human response (Schardein et al., 1985). These comparisons are presented in Table 11.13. There is, therefore, no a priori way of determining which animal species is most reflective of the human response to a particular chemical (Schardein, 1985). This drawback may be addressed in the future more insight is gained into the toxicological mechanisms and the pharmacokinetic differences among species (IRLG, 1986). Unless data are available to support an alternative animal model, however, a human health risk should be presumed if an adverse reproductive outcome occurs in a well conducted animal study.

Table 11.13

Comparisons of Concordant Malformations between Human Beings and Laboratory Animals

Teratogen	Reference Malformation	Concordant	Nonconcordant
Alcohol	Craniofacial, limb, CV	Mouse, dog	Rat, guinea pig, pig
Androgenic/progestogenic hormones	Pseudohermaphrodism (fem.)	Mouse, rat, guinea pig, pig hamster, rabbit, dog, primate	
Anticancer antimetabolites			
Aminopterin	Skeletal	Rat	Dog, pig
Fluorouracil	Multiple visceral	Mouse, rat, guinea pig	Rabbit, primate
Methotrexate	Skeletal	Rabbit, cat	Mouse, rat, primate
Cytarabine	Limb, ear	Rat	Mouse
Anticancer alkylating agents			
Busulfan	Multiple visceral		Mouse, rat
Chlorambucil	Urogenital		Mouse, rat
Cyclophosphamide	Digits	Mouse, rat	Rabbit, primate
Mechlorethamine	Renal, Limb, ear	Rat, rabbit, ferret	Mouse
Anticonvulsants			
Hydantoins	Facial, mental	Mouse	Rat, rabbit, rabbit
Diones	Facial, mental		Mouse, primate
Valproate	CNS		Mouse, rat, rabbit
Antithyroid agents	Hypothyroidism	Mouse, rat, rabbit, guinea pig	
DES	Uterine lesions	Mouse, rat, primate, ferret	
Methylmercury	Microcephaly, mental	Mouse, rat, cat	Hamster
Thalidomide	Limb	Rabbit, primate	Mouse, rat, hamster, dog, cat, pig, ferret
Lithium	CV		Mouse
D-Penicillamine	Skin lesion	Rat	Hamster
Streptomycin antibiotics	Inner ear	Rat	
Vitamin A analogues	CV, ear, brain	Rat, mouse, hamster, dog, primate	Rabbit, guinea pig, pig

Source: Schardein, J.L., et al., 1985, "Species Sensitivities and Prediction of Teratogenic Potential," Environ. Health Perspect., Vol. 61, p. 61.

As with chemical carcinogens, guidelines have been developed to evaluate the adequacy of reproductive and developmental toxicity studies (USEPA, 1986d; IRLG, 1986; Schardein, 1985; Wilson, 1973, 1980). Many of the same concerns associated with sensitivity of the carcinogenicity bioassay apply to the developmental toxicity studies as well. A summary of test endpoints for evaluating developmental and maternal toxicity are presented in Tables 11.14 and 11.15. Tests to measure the transmission of hereditable changes have also been developed, and are described elsewhere (USEPA, 1986b; IARC, 1983; NRC, 1983b).

Enough animals should be included in the study to ensure adequate sensitivity for detecting an effect. This typically includes at least twenty animals per dose group (USEPA, 1986d). Findings of positive associations between dose and developmental effects with lower numbers of animals per dose group should be seriously considered, as such findings indicate that a much stronger response would have been observed had adequate numbers of animals been used.

Also, it is recognized that in all animal species there is a detectable incidence of spontaneously occurring developmental effects (Wilson, 1980; IRLG, 1979). Concurrent controls should therefore be evaluated as part of the assay. As with controls for chemical carcinogenicity studies, primary attention in developmental effects studies should be given to the effects in the treated groups relative to the concurrent controls, especially if there is a difference in the incidence rates between the concurrent and historical controls.

Table 11.14

Endpoints of Maternal Toxicity

- Mortality
- Gestation Index (no. with implants/no. with sperm or seminal plugs
- Body Weight
  - Treatment days (at least first, middle, and last treatment days
- Organ Weights (in cases of suspected organ toxicity)
  - Absolute
  - Relative to body weight
- Gross Necropsy and Histopathology
- Fertility Index (no. with seminal plugs or sperm/no. mated)
- Gestation Length (when allowed to deliver pups
- Body Weight Change
  - Throughout Gestation
  - During Treatment (including increments of time within treatment period)
  - Post-treatment to sacrifice
  - Corrected maternal body weight (change throughout gestation minus gravid uterine weight or litter weight at sacrifice)
- Food/Water Consumption (where relevant)
- Clinical Evaluations (on days of treatment and at sacrifice)
  - Types and Incidence of Clinical Signs
  - Enzyme Markers
  - Clinical Chemistries

Source: USEPA (U.S. Environmental Protection Agency). 1986d, "Guidelines for the Health Assessment of Suspect Developmental Toxicants," Federal Register, Vol. 51, pp. 34027-34040.

Table 11.15

Endpoints of Developmental Toxicity

Litters with implants	Litters with live offspring[b]
No implantation sites/dam	No. and percent litters with live offspring
No. corpora lutea (CL)/dam[a]	Sex ratio/litter
Percent preimplantation loss $\frac{(CL - implantations) \times 100}{CL}$[a]	No. and percent live offspring/litter
No. and percent live offspring/litter	Viability of offspring[c]
No. and percent resorptions/litter	Mean offspring body weight/litter[c]
No. and percent litters with resorptions	Mean male body weight/litter[c]
No. and percent late fetal deaths/litter	Mean female body weight/litter[c]
No. & percent nonlive (late fetal deaths & resorptions) implants/litter	No. and percent externally malformed offspring/litter
No. and percent litters with non-live implants	No. and percent viscerally malfromed offspring/litter
No. and percent affected (non-live & malformed) implants/litter	No. and percent skeletally malformed offspring/litter
No. and percent litters with affected implants	No. and percent malformed offspring/litter
No. and percent litters with total resorptions	No. and percent litters with malformed offspring
No. and percent stillbirths/litter	No. and percent malformed males/litter
	No. and percent malformed females/litter
	No. and percent offspring with variations/litter
	No. and percent litters having offspring with variations
	Types and incidence of individual malformations
	Types and incidence of individual variations
	Individual offspring and their malformations and variations (grouped accouding to litter and dose)
	Clinical signs[c]
	Gross necropsy and histopathology

[a]Important when treatment begins prior to implantation. May be difficult in mice

[b]Offspring refers both to fetuses observed prior to term or to pups following birth. The end points examined depend on the protocol used for each study.

[c]Measured at selected intervals until termination of the study.

Source: USEPA (U.S. Environmental Protection Agency). 1986d, "Guidelines for the Health Assessment of Suspect Developmental Toxicants," Federal Register, Vol. 51, pp. 34027-34040.

The highest dose level should result in some maternal toxicity, in order to ensure that the test has the maximum sensitivity of detecting a positive response (USEPA, 1986d). Lower dose levels should not be toxic to the dam, and be sufficient to provide information from which a dose-response relationship may be established (USEPA, 1986d). These levels should include a no observed effect level as well as one or more dose levels in between the no observed effect level and the highest dose level (USEPA, 1986d). When evaluating the dose response data, the greatest concern should be with those agents which produce developmental toxicity at a dose that is not toxic to the adult, as this implies that the developing organism is selectively affected or more sensitive than the adult (USEPA, 1986d). Findings of developmental effects at dose levels which are also toxic to the mother should be interpreted cautiously, as it is possible that the developmental effects are secondary to maternal toxicity (USEPA, 1986d). Should this be the case, the study could imply that dose levels which do not cause toxicity in the mother should not cause harm to the conceptus either. Findings of developmental toxicity only at maternally toxic doses should not be summarily discounted, however, as current information is inadequate to assume that developmental effects at maternally toxic doses result only from maternal toxicity (USEPA, 1986d).

Because of the differences in species sensitivity to developmental toxins, it is possible that testing on only one species may fail to identify a substance that is toxic in other species, including human beings. It has been recommended, therefore, that at least two species be tested (Wilson, 1973). This is consistent with the protocol recommended for the animal testing of carcinogens (IARC, 1980; NTP, 1984). It has also been recommended that at least one species should be a non-rodent species (Schardein, 1985), preferably a non-rodent/non-rabbit species (Calabrese, 1983; Wilson, 1973). These recommendations are based on the concern that the reproductive systems in rodents and rabbits are similar, and differ significantly from the human reproductive system.

The dosing schedule is generally restricted to the early part of gestation, when most organogenesis is taking place and the organisms are most susceptible to teratogenic insult (See Table 11.16). Yet, even within this restricted interval, effects of repeated dosing may produce different embryotoxic outcomes than a single dose administered at a critical time in organogenesis. The ways in which these differences might occur are presented in Table 11.17. Several of these mechanisms include chemically-induced alterations in metabolic rates. A list of some chemicals known to influence metabolic rates is presented in Table 11.18. Of primary importance with regard to study sensitivity is the possible induction of catabolizing enzymes by high doses before the time of maximum susceptibility (Wilson, 1973). It has been recommended that, to address this concern, treatment intervals be subdivided into 3-4 day time spans (Wilson, 1973).

Multigeneration studies have been recommended to identify cumulative and genetic effects (Dixon, 1986; Manson, 1986). These studies are designed to last for three generations and involve the administration of the test agent to the first two generations (Dixon, 1986). The dosing of the F_0 generation begins as soon as possible after weaning and acclimation, and continues until all the F_1 animals selected for the next phase of the study have been weaned. Dosing of the F_1 generation selected for breeding continues until 30 days after the F_2 animals have been weaned. Consequently, this protocol increases the sensitivity of the assay to identify recessive genetic effects and effects which may occur at any stage of the life cycle.

Table 11.16

Critical Period of Organogenesis in Various Species

Species	Days of Organogenesis[a]
Mouse	7-16
Hamster	7-14
Rat	9-17
Guinea Pig	11-25
Armadillo	1-30
Ferret	12-28
Rabbit	7-20
Cat	14-26
Rhesus Monkey	20-45
Baboon	22-47
Dog	14-30
Ovine	14-36
Bovine	8-25
Porcine	12-34
Human	20-55[b]

a. Following Fertilization
b. Also may be given as day 35-70 after last menstrual period.

Source: Schardein, J.L., Chemically Induced Birth Defects, Marcel Dekker, Inc., New York, 1985.

Table 11.17

Ways in Which Repeated Treatment Prior to the Peak Susceptible Period of the Embryo May Produce Misleading Results

Time of treatment	Primary effect	Secondary effect capable of altering test results
1. Before implantation	Interference with implantation	No issue
2. Early organogenesis	Early embryonic death	No issue
3. Before peak susceptibility	Induction of catabolizing enzymes	Reduced blood level during susceptible period
4. Before peak susceptibility	Inhibition of catabolizing enzymes	Increased blood level during susceptible period
5. Before peak susceptibility	Liver pathology or reduced function	Increased blood level during susceptible period
6. Before peak susceptibility	Kidney pathology or reduced function	Increased blood level during susceptible period
7. Before peak susceptibility	Saturation of protein-binding sites	Increased blood level during susceptible period

Source: Wilson, J.G., 1973, Environment and Birth Defects, Academic Press, New York, pp. 137-171.

Table 11.18

Some Chemical Agents Known to Influence Rates of Metabolic Degradation of Themselves And/Or Other Compounds After Repeated Dosage

Increase metabolic degradation	Decrease metabolic degradation
Barbiturates	SKF 525-A
Thyroxine	Chlorthione
Some insecticides (DDT, chlordane, aldrin, dieldrin, heptachlor)	Iproniazid
Some tranquilizers and antipsychotics (meprobamate, Librium, chlorpromazine)	Metopirone
Some antihistamines (chlorcyclizine, diphenhydramine)	Actinomycin D
Several hypoglycemic agents	Puromycin
3,4-Benzpyrene	CCl_4
3-Methylcholanthrene	Triparanol
Steroid hormones	Chloramphenicol
	Any that competitively inhibit catabolic enzymes

Source: Wilson, J.G., 1973, Environment and Birth Defects, Academic Press, New York, pp. 137-171.

In addition to teratogenicity, other measures of developmental or reproductive toxicity could be utilized in the identification of substances which could produce these hazards. Such additional parameters are often associated with teratogenicity, and occur more frequently and consistently in animal experiments (Schardein, 1985). For example, low pregnancy rate, reduced litter size, and poor viability were observed in the early animal tests with thalidomide in rodents in the absence of teratogenicity (Schardein, 1985). Also, these parameters are more sensitive indicators of developmental effects than are gross malformations (USEPA, 1986d). For example, under a testing protocol which uses 20 animals per dose group, it is possible to detect an increased incidence of malformations in the range of 5 to 12 times above control levels, an increase of 3 to 6 times in the in utero death rate, and a decrease of 0.15 to 0.25 times the fetal weight (USEPA, 1986d). Other factors may need to be considered, however, when evaluating changes in fetal weight. In polytocous animals, for example, fetal and neonatal weights are usually inversely correlated with litter size (USEPA, 1986d). Also, the average body weight of males is greater than that of females in the more commonly used laboratory animals (USEPA, 1986d), thus warranting consideration of the sex ratios.

Sensitive developmental toxicity endpoints which are not commonly investigated include those that measure subtle functional defects in organs or organ systems (USEPA, 1986d). Often, these effects occur at levels below those which cause gross malformations (USEPA, 1986d). Much of the work that has been done has focused on behavioral effects, although the cardiopulmonary, immune, endocrine, digestive, urinary, nervous, and reproductive systems are also subject to alterations in functional competence (USEPA, 1986d). Neurotoxicity is a significant concern in this regard, as laboratory animals are poor surrogates for the identification of effects on complex human neurological function.

Tests to measure effects on fertility have also been developed (Dixon, 1986, IRLG, 1986). Positive findings from these studies may have special significance in that the fertility of humans is more susceptible to environmental chemicals than is the fertility of laboratory animals (Dixon, 1986; IRLG, 1986). Laboratory animals, for example, produce sperm in considerable excess over that required for normal reproductive function (IRLG, 1986).

11.1.4 Identification of Acute/Chronic Toxicity Effects.

Any organ or organ system in the body may experience toxic effects from chemical exposure. For most situations subject to risk assessment, the predominant concern is for the identification of chronic health effects. Of particular importance are those effects resulting from long-term exposure to toxic substances at levels below those which produce acute toxicity. This exposure situation warrants considerable attention for several reasons. Firstly, the exposed population, as well as the number of potentially hazardous chemicals to which this population is exposed, increases as the exposure dose decreases. Also, as the exposure duration sufficient to produce a toxic effect increases, so does the probability that those experiencing the toxic effect are exposed to an increasing number of other risk factors as well. This problem limits the ability of epidemiology to attribute an adverse health outcome to a particular substance. Moreover, registries of most chronic diseases (other than cancer) are not kept. This further limits the ability to identify toxic hazards even in the absence of other risk factors. Finally, even if a toxic outcome is established in association with exposure to a toxic substance, many years may have elapsed between the time of the chemical's identification and the time of its introduction into the market. The shortcomings of disease surveillance, combined with the inherent limitations of animal studies (See Chapter 4.4), contribute to much of the uncertainty and anxiety concerning the hazard identification of toxic substances.

Traditional measurements of acute and chronic health effects have focused on common manifestations of toxicity. These include such overt acute effects as irritation, dermatological disorders, behavioral disturbances, central nervous system depression, cardiopulmonary depression, and death. They also include common biochemical indicators of visceral organ toxicity (such as serum liver enzyme levels). As the scientific understanding of chemically-induced health effects has expanded, however, an increasing number of toxicological endpoints has been investigated. These endpoints include effects associated with genotoxic, carcinogenic, and reproductive/developmental risks discussed earlier. Such effects are not discussed in this subchapter, except insofar as they were observed in studies along with other

health effects not previously described. Additional endpoints also include subtler toxic effects associated with reversible or irreversible tissue damage (for example, neuropathic or immunosuppressive effects). A summary of selected non-carcinogenic health endpoints for various target organs is presented in Table 11.19.

Because different health effects are likely to result from different toxicological mechanisms, there are limitations associated with the use of one endpoint or group of endpoints as a surrogate for a substance's overall toxic potential. This caveat is particularly relevant when evaluating lethality studies (for example, LD50, LC50, LD10), as there may be many toxicologically significant endpoints which are not well correlated with acute lethality. For example, because of the body's homeostatic responses to acute exposures, and its ability to respond to insult following these exposures, the concerns associated with chronic exposure may be quite different from those associated with acute exposure. Chronic effects may occur as a result of the accumulation of subtle tissue damage through repetitive exposures, as a consequence of adaptive mechanisms induced by the exposures, or from an accumulation of a substance within the body to a level sufficient to cause an acute response (Figure 11. 5).

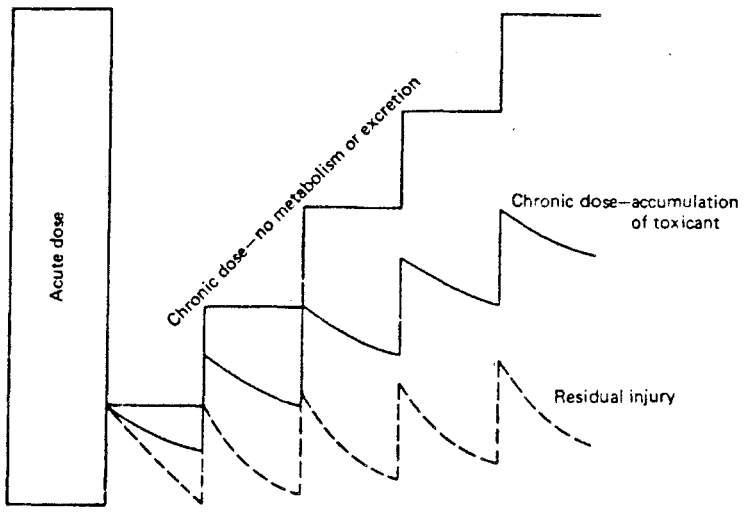


Figure 11.5. Diagrammatic View of Dose and Corresponding Measure of Effect.

Source: Klaasen, C.D., 1986, "Distribution, Excretion, and Absorption of Toxicants," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 33-63.

Table 11.19

Selected Parameters Associated
with Toxicity to Various Organs and Organ Systems

Organ/Organ System	Parameters Indicative of an Adverse Effect
Hematopoietic	Bone marrow depression, Abnormally high reticulocyte count, Hypoxia, abnormal hemoglobin
Immune	Alterations in lymphoid organ weight or histology, quantitative changes in perpheral leukocyte counts and differentials, depressed cellularity of lymphoid tissues, increased susceptability to infections by opportunistic organisms, increased incidence of allergy and autoimmunity
Cardiovascular	Changes in heart rate, conductivity, excitability, and contractility, changes in cardiac output, changes in arterial blood pressure, hemorrhage, thrombosis, structural changes, hypersensitivity
Respiratory System	Irritation of the air passages, damage to cells lining the airways, production of fibrosis, constriction of airways through allergic responses
Nervous System	Alterations in functions: coguitive, sensory, somatosensory, motor, antonomic, immune effects; structural damage
Eye	Lesions of the cornea, lens, and retina; tear secretion, ocular pressure, electroretinogram; biochemical assays of lens and aqueous humor; eye reflexes
Liver	Interference with bilirubin uptake, excretion and conjugation; cytotoxic injury; cholestatic injury; fatty liver; cirrhosis; phospholipidosis, vascular lesions; chronic active hepatitis; subacute hepatic neurosis
Kidney	Decreased elimination of wastes; changes in extracellular fluid volume and electrolyte composition; changes in hormone levels involved in systematic metabolic functions; changes in prostaglandin and kinin levels

General Sources: C.D. Klaasen et al. (eds.), 1986, Casarett and Doull's Toxicology, Third Edition, MacMillan Publishing Company, New York, 974 pp.; G.M. Cohen (ed.), 1986 Target Organ Toxicity, Volumes I and II, CRC Press, Boca Raton, Florida; National Research Council, Drinking Water and Health, Vol. 6, National Academy Press, Washington, D.C., pp. 105-138.

Most toxicological studies have concerned effects on organs or organ systems that receive the most intense exposure, either through direct contact or through transport via the blood. Examples of these sites are the blood itself, the skin and external organs, the pulmonary and cardiovascular systems, the liver, and the kidney. Exposure intensity is not the only parameter of concern, however, as tissues may vary widely in their response to a given dose of a toxic substance. For example, certain tissues, such as heart tissue (Balazs et al., 1986), are less able than the liver and kidney tissues to protect themselves from the toxic effects of reactive metabolites.

Generally, the epidemiological assessment of health effects depends on the ability of the studies to identify the populations exposed directly relate the observed effects to actual health impairment. Difficulties exist, however, in the identification of sensitive health endpoints, in the adequate assessment of exposure levels, and in separating the effects of a particular exposure from those induced by exposure to other chemical, physical, or biological agents. Despite these shortcomings, however, an important focus of risk assessments should be the identification of sensitive indicators of adverse health effects. Such limitations are overcome through the use of laboratory tests. As the health effect parameters become more sensitive, however, their association with clinical effects becomes less clear. Thus, a distinction has to be made between the significance of a change in a particular health parameter and the overall health significance of this change. Any chemically-induced perturbation in an organism, regardless of its clinical significance, may contribute to a toxic outcome. Therefore, any subtle but significant biological change may constitute an adverse health effect. Evaluation of this change within the context of the substance's other toxicological manifestations is required before conclusions can be made regarding its importance to human health. This evaluation is conducted in the Hazard Assessment section (Section V).

11.2 Identification of Multiple Chemical Exposure Effects.

Human beings are not typically exposed to a single compound of toxicological concern. Rather, they are exposed to varying numbers of such substances. These substances may enter the body through oral, pulmonary, and dermal exposure routes. The human response to any one chemical in this exposure mixture may be either independent or dependent on the properties of the other chemicals in the mixture. Independent responses which affect the same toxicological endpoint are additive in nature. The severity of the response is a function of the toxic potencies of the individual chemicals. Dependent responses are interactive in nature. When the response is greater than what would be predicted by adding the potencies of the individual chemicals, the response is considered synergistic. A special kind of this interaction takes place when a chemical with no known contribution to a particular toxicological effect increases the potency of another chemical relative to that effect. This particular synergistic response is potentiation. When the interaction of chemicals results in a response which is less than what would be predicted from an additive model, the interaction is considered to be antagonistic.

The toxicological importance of a substance may also be altered significantly when it precedes or follows exposure to another substance. In addition to temporal variability, the nature of the interaction also depends on the concentrations of the substances to which individuals are exposed. Toluene, for example, may antagonize the toxicity of benzene when co-exposure occurs at high concentrations (Goldstein, 1983). It does this by competing with benzene for enzymes responsible for benzene's metabolic activation. At lower concentrations, however, this antagonism should not occur because metabolism would not be saturated. To the contrary, it is possible that toluene could enhance benzene's toxicity at these lower concentrations, through the induction of metabolizing enzymes.

Chemical exposures associated with such lifestyle factors as diet and smoking habits may also interact with a substance's toxicity. For example, the synergistic effects of cigarette smoking on the carcinogenicity of asbestos have been widely discussed (See, for example, OSHA, 1980). A low carbohydrate diet increases hepatic metabolism of various chemicals, whereas a high carbohydrate diet decreases it (Sato et al., 1983). Opposite effects to those of carbohydrates were observed with ethanol (Sato et al., 1983). These lifestyle factors, therefore, may either enhance or decrease the toxicity of a substance, depending on whether the effect of concern is caused by the parent compound or its metabolite.

Consideration of interactive effects is a requisite element in the proper assessment of specific exposure situations, and in the proper identification of sensitive populations or exposure cohorts. Assessment of interactive effects has been hampered, however, by the focus on the toxicity of individual compounds rather than on a matrix of risk factors. Nonetheless, this chemical-specific approach is taken because the assessment of the many potential risk factors represents a complicated, time-consuming, and costly process.

In the future, problems associated with the assessment of multiple chemical exposures may be overcome by a clearer understanding of the mechanisms of toxicity, and the influences which various environmental stresses have on these mechanisms. Meanwhile, the key risk assessment issue in this regard concerns the underprediction of the toxic effects associated with these chemical exposures. The problem is addressed to some extent by the policy that assessments should first consider the toxicological effects of chemical mixtures when they can be identified with an adequate database. If this cannot be done satisfactorily, attention should be given to the interactive effects observed in the studies on individual chemicals. Synergistic interactions, while infrequently reported, should be assessed when information is available. Antagonistic interactions, although potentially relevant with regard to specific exposure situations or population cohorts, may not be generally applicable for risk assessment purposes.

Even when information on interactive effects is available, however, it may be difficult to describe these interactions quantitatively. For most situations, therefore, independence of effects is usually assumed unless adequate information exists to support a different assumption. While these effects may either be additive or non-additive in nature, proper hazard identification warrants that they should be assumed additive unless the available data indicate otherwise.

11.3 Identification of Special Risk Populations

For any given substance, populations exist which may be especially vulnerable to its toxic effects. These populations may be at special risk either because of their exposure circumstances or because of their inherent sensitivity to the substance's biological effects. Individuals may thus vary widely in their responses to toxic chemicals.

Specific factors which may influence a chemical's toxicity include genetics, age, sex, health status, diet and nutrition, lifestyle, occupation, and environmental setting. Infants and elderly people, and those with compromised liver function, have diminished metabolic capability (Sipes and Gandolfi, 1986). They also have altered susceptibility to neurotoxic disorders (NRC, 1986). Hereditary factors may have a direct or indirect role in the development of many chronic diseases (NRC, 1983b). Certain occupational toxicants may interact in unpredictable ways with other chemicals. Malnutrition is factor in the development of neurotoxic responses (NRC, 1986). Malnutrition and infection diminish one's immune system. A simplified representation of this variability is presented in Figure 11.6.

Estimation of an exposure level which protects the population from the harmful effects of a substance must consider these potentially wide differences in susceptibility. Variations in sensitivity may be caused by a combination of pharmacokinetic and pharmacodynamic factors, leading to potentially marked differences in the shape of the dose-response curve for different human subpopulations (Gillette, 1985).

Sensitive populations may also exist by virtue of increased activity levels. By increasing the absorption rate for inhaled substances, and decreasing the blood flow rate to the liver, increased activity levels modify the delivered dose of a substance. This modification in pharmacokinetics is especially important when considering threshold effects associated with short-term inhalation exposures; increased activity would decrease the time necessary to achieve equilibrium in blood and tissue concentrations.

Infants represent a special risk population for several reasons. Relative to older children and adults, they absorb chemicals more readily from the gastrointestinal tract, but have lower protein binding efficiencies are less able to metabolize and excrete them (WHO, 1986). Toxic substances may also exert adverse effects on organ systems still undergoing development, such as the nervous, immunologic, gastrointestinal, and respiratory systems. Furthermore, as Figure 11.7 shows, weight gain in infants is higher than any other age group. Weight gain is evidenced by a period of rapid cell division. This period of rapid growth also takes place at a time when the body's metabolic and immunologic capabilities are not fully developed (WHO, 1986), thus leaving the infant more potentially vulnerable to tumor inititators and other genetic toxins. Infants may also be at special risk by virtue of their relatively high degree of oral and dermal contact with toxic substances in the environment. This issue was raised in Chapter 8.

Unfortunately, lack of data, as well as an inadequate conceptual understanding of the mechanisms behind many chemically-induced effects, usually prevents a rigorous determination of risk in sensitive human populations. To compensate for these shortcomings, conservative assumptions and methodologies must be incorporated into the risk assessment. Among these is the assumption that some proportion of the human population will be at least as sensitive to the effects of a chemical as the most sensitive animal species investigated in a laboratory setting (Gillette, 1985).

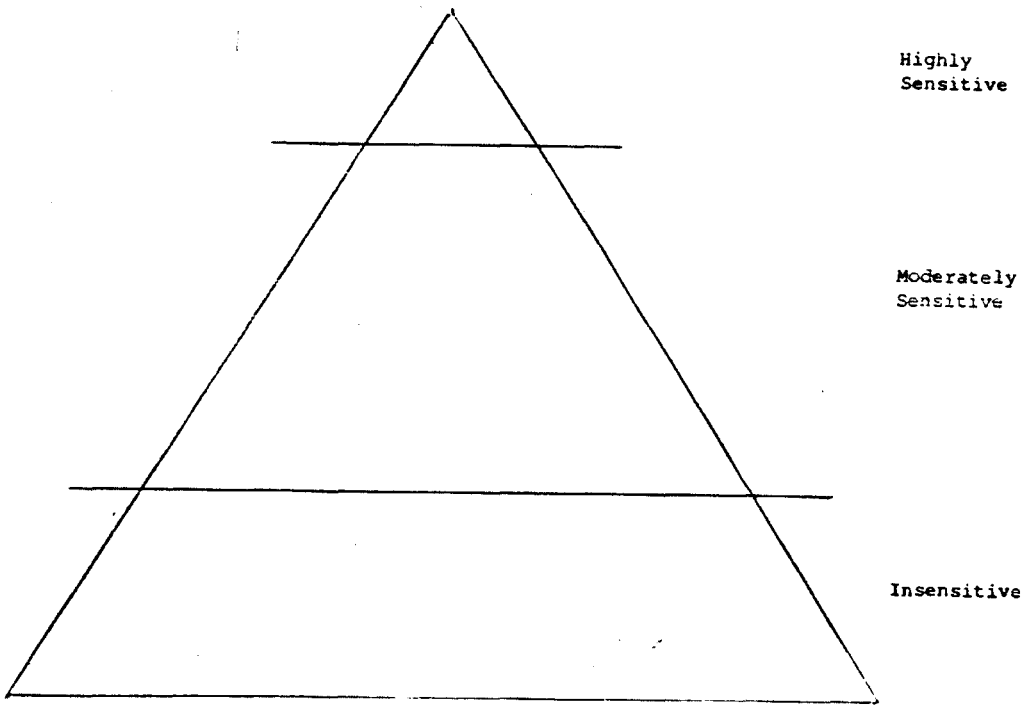


Figure 11.6. Variations in Response to Environmental Stress Among the General Population.

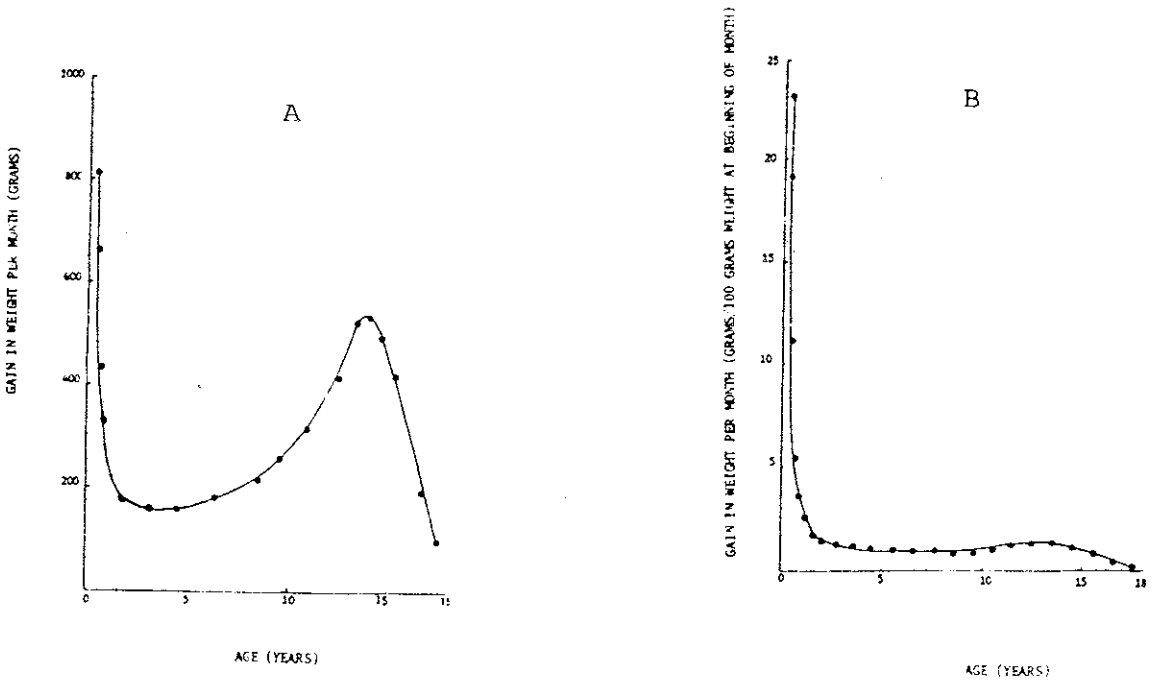


Figure 11.7. Gain Weight Per Month in Boys. a) Absolute weight gain; b) Weight gain in grams/100 grams of body weight.

Source: WHO (World Health Organization), 1986, Principles for Evaluating Health Risks from Chemicals during Infancy and Early Childhood: The Need for a Special Approach, Environmental Health Criteria 59, World Health Organization, Geneva, pp. 17-18.

12. Health Effects of Reaction Products

Some chemicals may be transformed into toxic substances as a result of chemical or biological reactions in the environment. A popular example of this is the atmospheric transformation of sulfur dioxide into sulfates. Reaction products, such as sulfates, are identified in the Exposure Assessment section (Chapter 9.1). The health effects of these reaction products should be reviewed, as the assessment should consider situations in which human exposure to a chemical's breakdown products occurs. The findings from this review of the health data, along with the findings from the exposure assessment, may form the basis for recommending that these reaction products undergo formal assessment themselves. The framework for this process has already been established in the Maine Bureau of Health's Hazardous Air Pollutant Program (Anderson, 1986). Consideration of breakdown products also has importance relative to Maine's groundwater monitoring program for pesticides. Ethylene thiourea, for example, an animal carcinogen and teratogen, is a breakdown product of widely used dithiocarbamate pesticides.

13. Environmental Effects

Most risk assessment issues concern the direct human health effects of chemical exposure. There may be impacts, however, which indirectly affect human health by disrupting the environmental support systems. These disruptions may either be associated with alterations in environmental chemistry, or with impacts on biological systems. While these modifications could be very subtle, they could also be the most significant of all effects caused by toxic substances. For impacts on the climate, large segments of the world's population may be affected, but with no direct relationship between exposure and effect. Some impacts, such as the extinction of a species or depletion of the ozone layer, may be irreversible or require centuries to reverse. Because of the difficulty in interpreting the parameters which measure these effects, however, there is a significant concern that the significance of such effects will not be understood in time to prevent serious environmental damage.

13.1 Effects on Environmental Chemistry.

The chemical equilibrium of the environment is maintained by a variety of chemical reaction cycles. Substances which are emitted in large enough quantities, or whose environmental persistence is relatively long, may cause significant alterations in these reaction cycles. For example, hydrocarbon and nitrogen oxide emissions in the lower atmosphere can lead to the build-up of ozone to toxic levels by inhibiting reactions which destroy this gas. On the other hand, emissions of chlorofluorocarbons into the upper atmosphere may enhance the destruction of ozone, thus depleting the shield which protects the earth from harmful ultraviolet rays. In addition to identifying sensitive reaction cycles, the assessment of the impacts on environmental chemistry is further complicated by the possibility that some chemicals may enhance, whereas others may inhibit, the same reaction cycle.

13.2 Ecosystem Effects.

Substances may affect ecological systems, either by direct effects on sensitive plant and animal species, or by disrupting the biological interactions which occur among these species. Such ecosytem effects involve agricultural, terrestrial, freshwater, saltwater, and aerial species. For example, depletion of certain lichen species around London was an indicator of the air pollution in that city as far back as the mid nineteenth century. Surveys of chemical effects on plants and animals have also helped in the identification of potential and actual exposures to hazardous wastes. The effects of chlorinated pesticides, such as DDT, on the survival of terrestrial wildlife has been well documented.

Investigations concerning the effects of a substance on the environment may thus serve a variety of purposes. Firstly, they may identify adverse effects not related to human toxicity, but which significantly alter the functioning of systems which support human life. They may also indicate a source of exposure with potentially adverse implications for human health. Finally, they may provide additional toxicological insight regarding how certain substances may affect human health. For risk assessment purposes, the findings of these studies have the greatest relevance when the qualitative assessment of chemical hazard is conducted (See Chapter 15).

14. Structure-Activity Relationships

Most relevant information concerning health effects is drawn from studies of individual compounds. Toxicological information on certain compounds, however, may be supplemented through the study of related compounds. These compounds are related in the abilities to induce the same biological effect, and may either be structurally similar or dissimilar (Tichy, 1983). The magnitude of certain toxic effects may depend more on their physical properties than on their specific biochemical reactivities (Tichy, 1983). The correlation between the threshold limit values (TLVs) and blood air partition coefficients for a number of chlorinated hydrocarbons is described in Figure 14.1.

This relationship suggests that the general toxicity of these compounds may be predicted on the basis of their relative solubilities. In other cases, however, basing predictions of toxicity on structural similarities may lead to misleading conclusions. Benzo-a-pyrene, for example, is a potent mutagen and animal carcinogen whereas a structurally similar compound, benzo-e-pyrene, has not been found to be carcinogenic (IARC, 1983). This difference may be attributable to differences in the ways in which the compounds are activated to toxic metabolites (Selkirk and MacLeod, 1979). Conversely, structurally dissimilar compounds may produce similar biological effects, as is the case with the wide variety of compounds which can induce proliferation of peroxisomal enzymes (Reddy and Lalwai, 1983). Thus, while information on structure-activity relationships may be useful in risk assessment, the data should be cautiously interpreted.

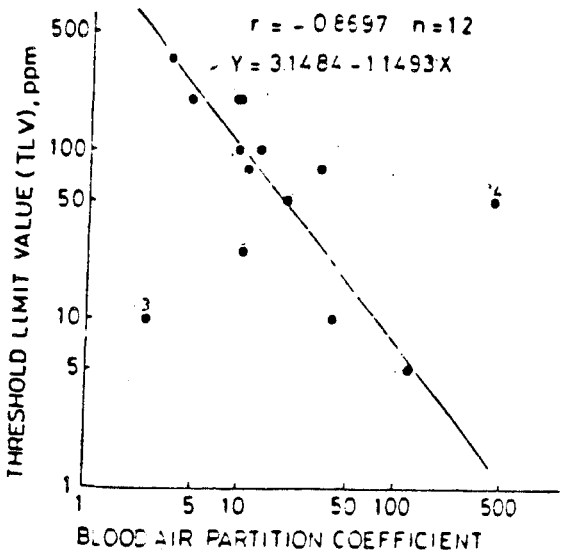


Figure 14.1. Relationship Between Threshold Limit Value and Partition Coefficients for Various Chlorinated Solvents

Source: Sato, A., and Nakajima, T., 1979, "A Structure-Activity Relationship of Some Chlorinated Hydrocarbons," Arch. Environ. Health, March/April, pp. 69-75.

SECTION V: HAZARD ASSESSMENT

The Hazard Assessment comprises both a qualitative and quantitative evaluation of the health risks associated with exposure to toxic substances. Of the two, the qualitative assessment involves a discussion of hazard that is more specific to the substance of interest. Due to limited databases for most chemicals, quantitative assessments usually must rely on generally accepted procedures which are applicable to a wide variety of chemicals. These include the use of default uncertainty factors when estimating the degree of health hazard in sensitive human populations, as well as default low dose extrapolation models for estimating the health risks associated with non-threshold effects. Despite its limitations, however, the quantitative risk assessment findings have more practical relevance for risk management and risk communication purposes than do the findings of the qualitative risk assessments. Issues associated with this situation are discussed in the Risk Characterization section (Section VI).

15. Appraisal of Hazard Identification Studies.

In the appraisal of the Hazard Identification studies, all potential health effects of concern are evaluated with reference to the overall biological significance of chemical exposure. By comparing the findings of the studies, the consistency of the toxic response pattern can be evaluated. In this comparison, the severities of the toxicological endpoints are assessed, along with the abilities of the studies to detect these endpoints and the dose-response relationships found. This evaluation then forms the basis for a discussion of the mechanisms of action. After the empirical and conceptual basis for a chemically-induced health effect is evaluated, the strength of the association (or weight of evidence) between exposure and the occurrence of that effect in human beings is determined.

15.1 Comparison of Study Findings.

It is important that all the toxicologically relevant information identified in the assessment be evaluated. This evaluation should include a discussion of the non-positive as well as positive studies. Differences in study findings may result from differences in study design, study quality, the species, strain, or sex under investigation, the level of histopathological examination, the endpoints evaluated, and the likelihood of detecting an effect. This evaluation should also include a discussion of other known or suspected risk factors for the health effects identified in the previous section.

Two factors affecting the results observed in the health studies are the design of the studies and the particular health endpoints measured. The first factor can be addressed by statistical comparisons of the various studies to determine their abilities to detect a particular effect. The second factor is important because various tests have been developed to identify and measure the potential for a substance to cause particular health effects. These tests may vary significantly in their sensitivities to detect these effects.

Spirometry tests of lung function, for example, may not be sensitive enough to detect subtle changes in airway resistance (Berkow, 1982). Because the lower respiratory tract is particularly susceptible to damage, failure to detect an effect from spirometry tests does not necessarily imply the absence of an effect. On the other hand, if an effect is observed from spirometry tests, it is a good indication that significant lung damage has occurred. Thus, consideration of test sensitivity is important when determining the effect severity and in subsequent derivations of action levels.

Aside from the evaluation of the individual health endpoints, the consideration of the spectrum of identified health effects is also a relevant area of discussion in the appraisal of health hazards. Assessment of the full range of health effects which a substance may produce is helpful when postulating a mechanism of action (See Chapter 15.2). Different health effects may be correlated with one another (for example, genetic toxicity, carcinogenicity, and developmental toxicity). Therefore, this particular assessment may provide important evidence for weight of evidence determinations, even in the absence of a precisely defined toxicological mechanism.

The dose-response relationship is still another important area of investigation with regard to the appraisal of the hazard identification studies. Generally stated, as the dose of the chemical increases, so should its toxicity. The most obvious explanation for the lack of a consistently positive relationship between chemical exposure and health effect is that the chemical causes no toxicity. Yet, the absence of a consistent dose-response relationship may also be caused by a number of other factors. These include competing mechanisms, pharmacokinetic characteristics, and statistical limitations created by a small sample size. Such factors should be assessed when evaluating the nature of the dose-response curve.

The shape of the dose response curve may provide valuable qualitative information concerning the nature of the response. The steepness of the curve may be useful in estimating threshold or no effect levels. Comparisons between different dose-response curves may also provide information concerning the relationships between pharmacokinetic and toxicological responses, or the inter-relationships between the different toxic responses. Similar to the benefit gained by comparing the spectrum of toxic responses, comparisons of dose-response relationships may also be useful in describing possible mechanisms of action, as well as in the overall weight of evidence determinations. An example of a relationship between the metabolism and hepatotoxicity of tetrachloroethylene is described in Figure 15.1. The parallels between metabolite formation and hepatotoxicity suggest that a metabolite of tetracloroethylene might responsible for its toxicity to the liver.

15.2 Evaluation of Health Effect Severity.

An assessment of the hazard identification findings should include an analysis of health effect severity. This analysis is necessary for two principal reasons. Firstly, it could provide much toxicologically relevant information from a relatively limited database, such as might be associated with a single well-conducted animal bioassay for carcinogenicity. Secondly, it should provide a basis for gauging a quantitative degree of health concern, as is needed when action levels are derived for chemical exposures using various uncertainty and modifying factors.

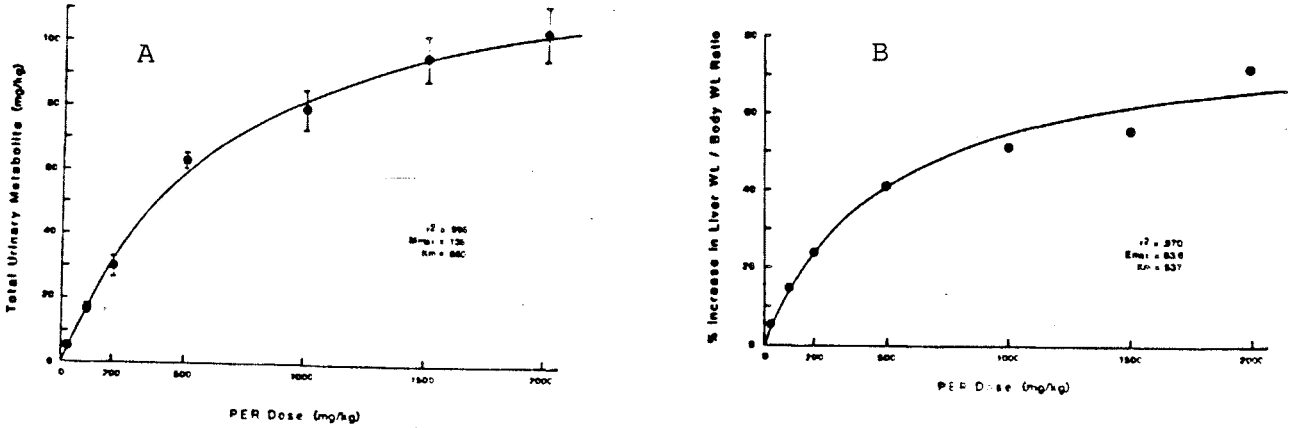


Figure 15.1. Relationship Between Tetrachloroethylene (PER) Dose and Toxicokinetic Parameters in Mice. a) relationship between PER dose and the amount of total urinary metabolite excreted per day; b) relationship between PER dose and the increase in liver weight.

Source: Buben, J.A. and O'Flaherty, E.J., 1985, "Delineation of the Role of Metabolism in the Hepatotoxicity of Trichloroethylene and Perchloroethylene: A Dose-Effect Study," Toxicol. Appl. Pharmacol., Vol. 78, pp. 105-122.

There are several parameters associated with cancer bioassays which may provide a qualitative concern regarding the potency of the carcinogenic response. These parameters are described in Table 15.1. Not all qualitative indicators of carcinogenic potency correlate well with quantitative potency estimates (Gold et al., 1986). Many potent carcinogens, however, have been found to cause tumors in multiple species and multiple sites, with relatively short latency periods, and with a large proportion of malignant tumors (Munro and Krewski, 1981). They have also demonstrated significant biological activity in short-term tests (Munro and Krewski, 1981).

Squire (1981) developed a quantitative ranking system for animal carcinogens based on qualitative indicators of potency. According to his system, qualitative indicators of carcinogenicity, such as the ones listed in Table 15.1, are given numerical weights. The maximum score a chemical could receive is 100. Recommended regulatory actions would become progressively more severe as the numerical ranking becomes closer to 100. A score of 100, for example, may be very likely to result in a recommendation to ban the chemical. Although this system does not directly translate bioassay result into a measure of a human cancer risk level, it does provide a method for assessing the relative potencies of carcinogens based on qualitative considerations. At the present time, however, this quantitative approach appears to be too arbitrary for risk assessment purposes.

Qualitative criteria have also been developed for comparing the relative severities of developmental effects. The severity of the effects generally follows the paradigm established by Wilson (1973): in order of increasing potency, effects range from reversible embryotoxicity and fetotoxicity, to teratogenicity, to fetal death. Severity indices for developmental effects can be derived by taking the logarithm of the ratio between the lowest adult toxic (or lethal) dose and the lowest developmental dose (NRC, 1986).

Finally, indices have been developed from which chemicals can be compared on the basis of their acute or chronic toxicity. For example, indices which compare relative general toxicity potencies have been developed by several regulatory agencies (USEPA, 1981). A summary of these systems is presented in Table 15.2. While this table only includes values based on oral and inhalation exposure, criteria regarding dermal exposures have also been established (USEPA, 1981), and should be followed accordingly. Criteria for evaluating the severity of an acute/chronic effect have been developed by the Massachusetts Department of Environmental Quality Engineering (MDEQE, 1985), and are presented in Table 15.3.

15.3 Evaluation of the Evidence Regarding Mechanisms of Action.

Ideally, all of the toxicological effects identified in the Hazard Identification section should be explained by some sort of biochemical mechanism. The extent to which these mechanisms can be derived from scientific studies is limited by an incomplete understanding of normal biological processes, and by an incomplete understanding of the biochemical processes leading to chemically-induced illness. Therefore, an adequate understanding of the toxicological mechanisms is not a prerequisite to the assessment of chemically-induced adverse health effects.

When sufficient data are available to describe the mechanism of action, however, they can provide a sound scientific basis for interpreting the results of the health effects studies. This information can also represent a starting point for evaluating the effects of a large number of chemicals on a particular health parameter. Furthermore, it may provide the basis for predicting how certain chemicals may interact with one another. As important and promising as information regarding the mechanism of action can be to risk assessment, it should be used with caution. Specifically, hypotheses regarding mechanisms of action should not be used to discount the findings of well conducted toxicological or epidemiological studies, unless there is reasonable certainty that the toxic effect only occurs through those mechanisms.

Table 15.1.

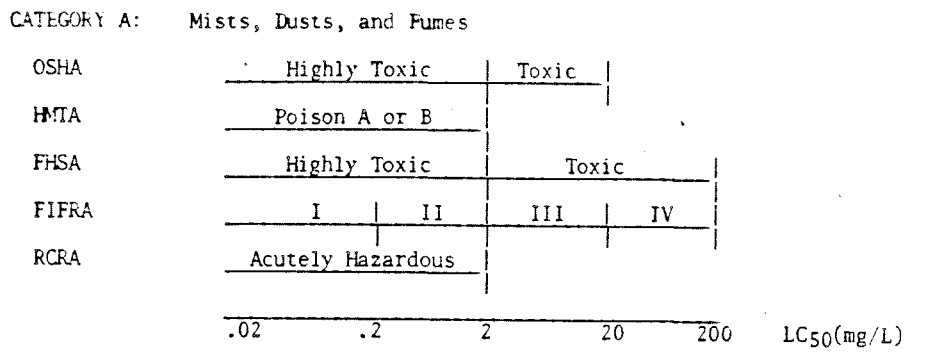
Important Characteristics of Carcinogenic Potency.

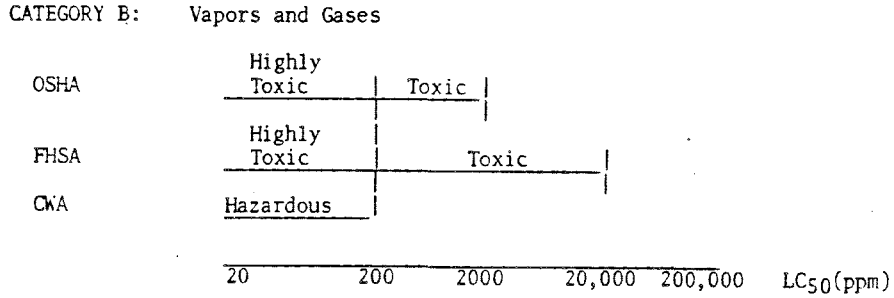
a) The proportion of animals bearing neoplasms at each exposure level.
   The number of neoplasms per animal.
   The number of different types of neoplasm.
   The number of species affected.
b) The magnitude of the dose at which the carcinogenic response occurs.
c) The proportion of malignant and benign lesions.
d) The nature and degree of other pathological changes.
e) The organ or target tissue in which the carcinogenic response occurs. It should be recognized that an increase in the number of tumors of a type which occurs spontaneously in a high proportion of the strain of animal being used (e.g., liver tumors or pulmonary adenomas in certain strains of mice) carries less weight in the estimation of potency than does the appearance of tumors in other organs.
f) The latency period before tumor development. The shorter the latency period the more potent is the chemical.
g) The sensitivity of the experimental model.
h) Chemical similarity to other known carcinogens.
i) Genetic toxicity and activity in short-term tests for carcinogenicity.
j) Biochemical reactivity with DNA, RNA, and protein.
k) Further information obtained from other toxicological studies such as kinetic and metabolic data. The significance of these in the estimation of potency to man is not clear in every case. Fundamental differences in genetic make-up between animals and man, which can lead to wide variations in response to the action of chemicals, include differences in immune and hormonal status, among others.

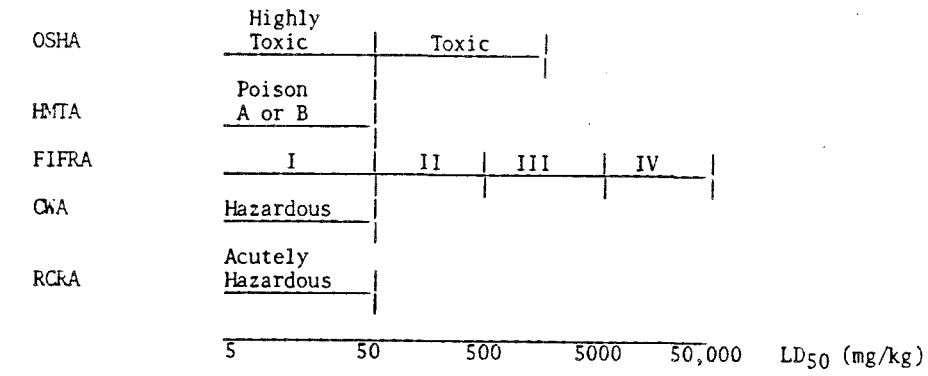
Sources: Munro, I.C., and Krewski, D.R., 1981, "Risk Assessment and Regulatory Decision Making," Food Cosmetic Toxicol., Vol. 19, p. 556; Purchase, I.F.H., 1985, "Carcinogenic Risk Assessment: A Toxicologist's View," in D.G. Hoel et al. (eds.), Risk Quantification and Regulatory Policy, Cold Spring Harbor Laboratory, pp. 181-182.

Table 15.2

Acute Toxicity Tests for Selected Federal Regulatory Programs







OSHA Occupational Safety and Health Act
HMTA Hazardous Materials Transportation Act
FHSA Federal Hazardous Sustances Act
FIFRA Federal Insecticide, Fungicide, and Rodenticide Act
CWA Clean Water Act
RCRA Resource Conservation and Recovery Act

Source; USEPA (U.S. Environmental Protection Agency), 1981, Chemical Substances Designation, Vol. I: Overview and Analysis, Washington, D.C., pp 108-109.

Table 15.3.

Severity Ranking for Acute/Chronic Effects

One Point:	Mild or Transient Irritant Effects (e.g. runny nose, eye irritation, coughing).
Two Points:	Moderate to Severe Irritant Effects; Mild to Moderate Transient systemic effects; effects generally considered to be reversible (e.g., bronchitis, anoxia, incoordination, fatigue, dizziness, reversible kidney effects).
Three Points:	Irreversible Pulmonary Effects; Serious Systemic Effects; chronic or persistent effects; cumulative effects; effects involving multiple sites or organ systems (e.g., cirrhosis, emphysema, epilepsy, peripheral nerve damage).

Source: MDEQE (Massachusetts Department of Environmental Quality Engineering), 1985, Chemical Health Effects Assessment Methodology, Massachusetts Department of Environmental Quality Engineering, Boston, Mass., p. 99.

15.4 Weight of Evidence Determinations.

Weight of evidence refers to likelihood that particular health effects occur in human populations as a result of chemical exposure. Weight of evidence determinations do not directly consider potency. Rather, they determine the data adequacy concerning the association between toxic substance exposure and a particular adverse health effect. Judgments are made with respect to both human and animal studies, and can be classified as "sufficient," "limited," "inconclusive," "inadequate," or "negative." These judgments, therefore, reflect the conclusions made in the evaluation of individual studies (See Chapter 5.1).

The weight of evidence from epidemiological and toxicological studies increases with the strength of the association, the presence of a dose-response relationship, and the number and proportion of corroborating studies. Replicated studies reduce the possibility that apparently positive results were observed by chance alone. Replicated studies done by different investigators further add to the weight of evidence by addressing any intralaboratory biases. The strongest evidence is derived from studies on human populations. With respect to animal studies, if an effect can be induced in different sexes, strains, or species, the likelihood that the effect will also occur in human populations is increased. In vitro studies are also used in weight of evidence determinations. Generally, they are used as evidence in support of epidemiological or toxicological studies. They are useful when assessing similarities in activity among structurally related compounds or in postulating possible mechanisms of action.

For all health effects, weight of evidence determinations should distinguish between primary effects and secondary effects which may result from an initial toxic effect on the cell. These determinations also need to consider the possibility that significant toxicity may prevent the effect of concern from being detectable or becoming manifest. Preliminary assays in genetic toxicity experiments should be able to determine substance dose levels which result in significant cytotoxicity (WHO, 1985). Likewise, subchronic assays should be run to determine maximum tolerated doses for cancer bioassays and maternally toxic doses for experiments on developmental toxicity.

15.4.1 Genetic Toxicity.

Weight of evidence for genetic toxicity refers to the likelihood that exposure to a chemical will result in some sort of DNA damage. The weight of evidence increases as the test systems showing a positive response more closely reflect actual human exposures. Within this context, two issues warrant special consideration. The first issue concerns the preference for experimental conditions which reflect the response of the entire organism, as opposed to the response of cells cultured in vitro. This preference is based on the fact that there are physiological influences on toxicity which cannot be studied in vitro.

The second issue concerns the preference for species most closely related phylogenetically to human beings. This preference is based on the fact that potential differences in cell biology may exist which could lead to misleading interpretations of the test results. Of particular importance is the difference between procaryotic and eucaryotic cell types. Eucaryotes differ from procaryotes with respect to many characteristics. For example, eucaryotic cells contain a nuclear membrane, have multiple chromosomes, have histones bound to these chromosomes, have large numbers of repetitive nucleotide sequences, divide by mitosis, and contain several organelles such as lysosomes (Calabrese, 1983; Stanier et al., 1976). These differences could account for significant variations in the responses of these two cell types to genetic toxins.

Eucaryotic organisms may also differ in their responses to genetic toxins. For example, human beings exhibit substantially greater DNA excision repair capacity than do rodents (Calabrese, 1983). On the other hand, human peripheral lymphocytic chromosomes have been found to be twice as sensitive than those of mice to the induction of translocations by ionizing radiation (Calabrese, 1983). Techniques to directly and accurately measure the effects of suspected genetic toxins on human populations are currently being developed, and may soon be available for widespread use (Thilly and Call, 1986; OTA, 1986).

In addition to the overall preference for phylogenetic closeness and in vivo experiments, findings from all well conducted experiments should be examined for reproducibility, a consistent dose-response curve, and the type of genetic lesion produced (WHO, 1985). Whole animal experiments may lack sufficient sensitivity for detecting potentially genotoxic compounds, despite the fact that these experiments consider in vivo conditions (WHO, 1985). Furthermore, although reflective of responses in procaryotic organisms, results from bacterial assays correlate strongly with known carcinogens and non-carcinogens in eucaryotic organisms (McCann et al., 1975; Purchase et al., 1978). McCann et al. (1975) found that approximately 90 percent of

carcinogens were also bacterial mutagens and that 90 percent of non-carcinogens failed to show mutagenic activity, when Salmonella typhimurium was used as the test organism. These findings were corroborated by Purchase et al. (1978). The ability of the Salmonella assay to detect carcinogens may even be greater if certain classes of compounds are excluded (OTA, 1981). This high degree of correlation has been recently questioned, as more chemicals have been tested and found not to conform to this paradigm (Tennant et al., 1987). Nonetheless, the correspondence between the Salmonella assay and the animal cancer bioassay findings reinforces the need to consider bacterial assays in carcinogenicity assessments. Conversely, a lack of mutagenicity in the bacterial assays should not necessarily alter the weight of evidence conclusions drawn from well conducted animal bioassays.

The Massachusetts Department of Environmental Quality Engineering developed weight of evidence guidelines based on various national and international risk assessment committees (MDEQE, 1985). These guidelines are presented in Table 15.4, with expanded documentation presented in Table 15.5. While the U.S. Environmental Protection Agency has no specific recommendations for the evaluation of mutagenicity test results relevant to chronic diseases such as cancer (USEPA, 1986a), it has developed guidelines relating specifically to germ cell mutations (USEPA, 1986b). These guidelines are presented in Table 15.6.

A separate weight of evidence determination for genetic toxicity may be unnecessary when positive toxicological or epidemiological data are available on carcinogenicity, developmental toxicity, and other potentially relevant chronic effects. In the absence of such information regarding these effects, the guidelines presented above provide a useful basis for making weight of evidence determinations regarding genetic toxicity. The presentation of this evidence should also be accompanied by statement regarding the level of testing. By doing so, the assessment could distinguish between classifications based on the weighing of positive and negative results and those based on lack of data.

15.4.2 Carcinogenicity.

The weight of evidence determination for carcinogenicity concerns whether or not a substance can be considered as a human carcinogen. Epidemiological studies and animal bioassays are the primary areas of investigation from which weight of evidence determinations are made. Both areas, however, are subject to limitations regarding their abilities to demonstrate causal connections between human exposure and carcinogenesis. Epidemiological studies have limited abilities to rigorously characterize exposed and non-exposed populations, to identify enough of these individuals to produce statistically meaningful results, and to separate the potential effects of a particular substance from those of other risk factors. Studies designed to address these obstacles generally require large amounts of time and resources to follow study populations over time.

Animal bioassays are limited by the fact that they are not direct qualitative or quantitative measures of a human response. Such limitations are becoming increasingly relevant in light of the growing numbers of chemicals identified in the animal studies as being potentially carcinogenic. Only about 30 chemicals or chemically-related processes have been positively associated with human cancer, while a few hundred have been found to be carcinogenic in laboratory animals (IARC, 1982; NTP, 1985).

Table 15.4

Scoring for Mutagenicity According to the MDEQE'S Chemical Health Effects Assessment Methodology

Category	Test Type and Number of Positive Results	Letter Code Score
Sufficient Evidence	Group I: Two or More	A
	or	
	Group II: Four or More	A
	or	
	Group III: Six or More	A
	or	
	Group I: One	A
	and	
	Group II: One or More	
	or	
	Group I: One	A
	and	
	Group III: Two	
Substantial Evidence	Group I: One	B
	or	
	Group II: Three	B
	or	
	Group III: Four or Five	B
	or	
	Group II: One or Two	B
	and	
	Group III: Three	
Suggestive Evidence	Group II: One or Two	C
	or	
	Group III: Two or Three	C
	or	
	Group II: One or Two	C
	and	
	Group III: One or Two	
Limited Evidence	Group III: One	D
Inadequate Evidence	No Data or Non-positive Data	E

MDEQE (Massachusetts Department of Environmental Quality Engineering), 1985, Chemical Health Effects Assessment Methodology, Massachusetts Department of Environmental Quality Engineering, Boston, Mass., p. 167.

Table 15.5

Categories of Mutagenicity Tests Used in MDEQE'S Weight of Evidence Protocol

TEST DESCRIPTION/TYPE
Group I: Mammalian, In Vivo
Mouse Specific Locus Test
Mouse Spot Test
Dominant Skeletal Mutation
Dominant Cataract Assay
Dominant Lethal Test - Rodents
Heritable Translocation Test - Rodents
Micronucleus Test - Mouse
Group II: Primary Short-Term Tests
Chinese Hamster Lung (V79) Cells, All Loci
Chinese Hamster Ovary (CHO) Cells
Mouse Lymphoma (L5178Y) Cells, TK Locus
S. typhimurium, histidine reversion (Ames Test - TA98, TA100, TA1535, TA1537-8)
E. coli (WP2/WP2 uvra) - reverse mutation
Sex-Linked Recessive Lethal Test - Drosophila m.
Host-Mediated Assay Studies
Mammalian Cytogenetics, bone marrow/lymphocyte or leucocyte
Mammalian Cytogenetics, oocyte, early embryo/male germ cell
Mammalian Cytogenetics, lymphocyte, leucocyte/cell culture
Mammalian Cytogenetics, all mammalian
Micronucleus Test, lymphocyte
Micronucleus Test, mammalian cell
Heretible (reciprocal) Translocation Test - Drosophila
Sister Chromatid Exchange - lymphocyte
Sister Chromatid Exchange - cells/embryonic lung fibroblasts (WI-38)/lymphocyte
Sister Chromatid Exchange - in vivo/in vitro
A. nidulans - cross over studies
S. cerevisiae, homozygous - recombination/gene conversion
E. coli pol A (W3110-P3478) - with S9/without S9
B. subtilis rec (H17-M45/17A45T) - spot test
Human Sperm Morphology
Cell Transformation Studies - BALB/C-3T3 / C3H/10T1/2
Cell Transformation Studies - mouse prostate
Cell Transformation Studies - Syrian hamster embryo
Cell Transformation Studies - SA7 Fischer rat cells
Group III: Secondary Short-Term Tests
Forward/Reverse Mutation, S. cerevisiae (YEF/YER) & S. pombe (YEY/YEZ)
Forward/Reverse Mutation, A. nidulans
Forward/Reverse Mutation, N. crassa
Plant Gene Mutation Studies
Body Fluid Assay - urine
Aneuploidy Studies, whole sex chromosome - loss/gain
Aneuploidy Studies, S. cerevisiae/ A. nidulans/ N. crassa
Micronucleus Test - plants
Plant Chromosome Studies
Mammalian Sperm Morphology - mouse/rabbit/rat
Mammalian Sperm Morphology - mouse F1 assay
Unscheduled DNA Synthesis - human diploid fibroblast
Unscheduled DNA Synthesis - mouse germ cells
Unscheduled DNA Synthesis - rat primary hepatocyte

Source: MDEQE (Massachusetts Department of Environmental Quality Engineering), 1985, Chemical Health Effects Assessment Methodology, Massachusetts Department of Environmental Quality Engineering, Boston, Mass., p. 161.

Table 15.6

Weight of Evidence Hierarchy for Germ Cell Mutations.

1. Positive data derived from human germ-cell mutagenicity studies, when available, will constitute the highest level of evidence for human mutagenicity.

2. Valid positive results from studies on heritable mutational events (of any kind) in mammalian germ cells.

3. Valid positive results from mammalian germ-cell chromosome aberration studies that do not include an intergeneration test.

4. Sufficient evidence for a chemical's interaction with mammalian germ cells, together with valid positive mutagenicity test results from two assay systems, at least one of which is mammalian (in vitro or in vivo). The positive results may both be for gene mutations or both for chromosome aberrations; if one is for gene mutations and the other for chromosome aberrations, both must be from mammalian systems.

5. Suggestive evidence for a chemical's interaction with mammalian germ cells, together with valid positive mutagenicity evidence from two assay systems as described under 4, above. Alternatively, positive mutagenicity evidence of less strength than defined under 4, above, when combined with sufficient evidence for a chemical's interaction with mammalian germ cells.

6. Positive mutagenicity test results of less strength than defined under 4, combined with suggestive evidence for a chemical's interaction with mammalian germ cells.

7. Although definitive proof of non-mutagenicity is not possible, a chemical could be classified operationally as a non-mutagen for human germ cells, if it gives valid negative test results for all endpoints of concern.

8. Inadequate evidence bearing on either mutagenicity or chemical interaction with mammalian germ cells.

Source: USEPA (U.S. Environmental Protection Agency). 1986b, "Guidelines for Mutagenicity Risk Assessment," Federal Register, Vol. 51, pp. 34005-34012.

According to the U. S. Environmental Protection Agency (USEPA, 1986a) and the International Agency for Research on Cancer (IARC, 1982), in the absence of adequate human evidence, a chemical which is found to cause cancer in laboratory animals should be presumed to cause cancer in human beings as well. The strength of the association, however, depends on the weight given to the animal studies and to the negative epidemiological studies. Criteria for assessing the weight of evidence for chemical carcinogens have been developed by both agencies (USEPA, 1986a; IARC, 1987; IARC, 1982). The two classification systems are very similar, especially when the USEPA criteria are compared with the more recent IARC criteria (IARC, 1987). The only major difference is that IARC considers as limited evidence those "neoplasms which may occur spontaneously in high incidences in certain strains," whereas USEPA considers this evidence sufficient unless other information is available to justify a lower classification. The reasoning employed by the USEPA is more consistent with the principles of worst case risk assessment than the reasoning employed by IARC. The USEPA criteria, therefore, are adopted by the Maine Bureau of Health to provide general policy guidance. These USEPA criteria for evaluating human and animal evidence are presented in Tables 15.7 and 15.8.

Three issues pertaining to these criteria require special attention. The first issue concerns the criteria which must be met before a causal association can be inferred from epidemiological studies. Again according to the U. S. Environmental Protection Agency (USEPA, 1986b) and the International Agency for Research on Cancer (IARC, 1982), three criteria must be met : (1) there is no identified bias which could explain the association, (2) the possibility of confounding has been considered and ruled out as explaining the association, and (3) the association is unlikely to be due to chance.

The second issue concerns the distinction between benign and malignant tumors in animal studies. Although USEPA and IARC placed specific emphasis on the presence of malignant tumors, this distinction may only be warranted in special circumstances. No chemical has yet been identified that produces only benign neoplasms (Williams and Weisburger, 1986). Also, many chemicals which induce primarily benign tumors in one species or strain induce malignant tumors in other species or strains (CDHS, 1985). Furthermore, truly benign tumors in rodents are consider to be rare and may actually represent a stage in the progression to malignant tumors (OSTP, 1985). Finally, IARC has stated that the discrimination between benign and malignant tumors is not as important an issue as it was once considered to be (IARC, 1980). Thus, the incidence of only benign tumors should be considered as evidence in support of the presumption of human carcinogenicity. When benign and malignant tumors appear together, their incidences may be combined, subject to established guidelines (NTP, 1984).

The third issue concerns the occurrence of false positives in animal studies. Typical bioassay protocols require analysis of 20 to 30 sites in both sexes of two species at two dosage levels (OSTP, 1985). The possibility needs to be addressed, therefore, that a statistically significant response may occur by chance. For a random binomial distribution in which 20 sites are evaluated at a 0.05 significance level, the probability of a false positive is about 0.64 for a single dose-sex-species combination, and about 0.87 for both sexes (Gart et al., 1979). Tests done at multiple doses yield even larger probabilities for false positives, according to this random distribution assumption (Gart et al., 1979). This argument has been criticized, however, on the grounds that the bioassay data represent counts, not continuous data

Table 15.7.

Assessment of Evidence for Carcinogenicity from Studies in Human Beings

i. Sufficient evidence of carcinogenicity, which indicates that there is a causal relationship between the agent and human cancer.

ii. Limited evidence of carcinogenicity, which indicates that a causal interpretation is credible, but that alternative explanations, such as chance, bias, or confounding, could not adequately be excluded.

iii. Inadequate evidence, which indicates that one of three conditions prevailed: (a) there were few pertinent data; (b) the available studies, while showing evidence of association, did not exclude chance, bias, or confounding, and therefore a causal interpretation is not credible.

iv. No Data, which indicates that data are not available.

v. No Evidence, which indicates that no association was found between exposure and an increased risk of cancer in well-designed and well-conducted independent analytical epidemiological studies.

Source: USEPA (U.S. Environmental Protection Agency). 1986a, "Guidelines for Carcinogen Risk Assessment," Federal Register, Vol. 51, pp. 33991-34003. Adapted from: IARC (International Agency for Research on Cancer), 1982, Chemicals, Industrial Processes and Industries Associated with Cancer in Humans, IARC Monographs Supplement 4, Lyon, France. 292 pp.

Table 15.8

Assessment of Evidence for Carcinogenicity from Studies in Experimental Animals

i. Sufficient evidence of carcinogenicity, which indicates that there is an increased incidence of malignant tumours or combined malignant and benign tumors: (a) in multiple species or strains; or (b) in multiple experiments (e.g, with different routes of administration or using different dose levels); or (c) to an unusual degree with regard to incidence, site or type of tumour, or age at onset. Additional evidence may be provided by data on dose-response effects, as well as information from short-term tests or on chemical structure.

ii. Limited evidence of carcinogenicity, which means that the data suggest a carcinogenic effect but are limited because: (a) the studies involve a single species, strain or experiment and do not meet the criteria for sufficient evidence described in i(c); (b) the experiments are restricted by inadequate dosage levels, inadequate duration of exposure to the agent, inadequate period of follow-up, poor survival, too few animals, or inadequate reporting; or (c) an increase in the incidence of benign tumors only.

iii. Inadequate evidence, which indicates that because of major qualitative or quantitative limitations, the studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect.

iv. No data, which indicates that data are not available.

v. No evidence, which indicates that there is no increased incidence of neoplasms in at least two well-designed and well-conducted animal studies in different species.

Source: USEPA (U.S. Environmental Protection Agency). 1986a, "Guidelines for Carcinogen Risk Assessment," Federal Register, Vol. 51, pp. 33991-34003. Adapted from: IARC (International Agency for Research on Cancer), 1982, Chemicals, Industrial Processes and Industries Associated with Cancer in Humans, IARC Monographs Supplement 4, Lyon, France. 292 pp.

(OSTP, 1985; Gart et al., 1979). When the nature of the bioassay data is considered, many investigators believe that the overall false positive rate does not greatly exceed 0.05 (Gart et al., 1979; OSTP, 1985). One reason for this concerns the background rate of tumor formation in laboratory animals. There are relatively few sites in commonly treated laboratory animals for which enough tumor bearing animals would be observed so that one could find by chance alone a significantly increased tumor incidence at the 0.05 level (Gart et al., 1979). For these sites (for example, those sites with background tumor rates above 5 percent), certain considerations may apply to ensure that the elevated response did not occur by chance. These include the requirement of corroboration of the response from other bioassays, or the imposition of a more stringent significance level. Haseman (1985), for example, recommended that a chemical be regarded as carcinogenic if it produces a high dose increase in a common tumor that is statistically significant at the 0.01 level, or a high dose increase in in an uncommon tumor that is statistically significant at the 0.05 level. Another recommendation is the use of the Bonferroni correction, which divides the significance level by the number of dose groups (Gart, 1979). According to this modification, if the significance level for an elevated tumor response in a one dose experiment is 0.05, the significance level for an elevated tumor response in a two dose experiment would be 0.025. Such adaptations to statitistical tests should be considered when evaluating the likelihood of a false positive finding in a bioassay.

In addition to the criteria developed for evaluating human and animal evidence, criteria have also been developed by the U. S. Environmental Protection Agency (USEPA, 1986a) and the International Agency for Research on Cancer (Vaino, 1987) to determine the overall weight of evidence criteria regarding a substance's human carcinogenic potential. These classification systems, which are presented in Tables 15.9 and 15.10, rely largely on the criteria which were developed for evaluating human and animal evidence. Overall, the IARC classification system appears to place less weight on the animal evidence than does the USEPA. Both systems, however, are designed to be flexible in order to incorporate other relevant data into the weight of evidence determination. It is, therefore, difficult to establish precisely how various substances will be evaluated according to these different classification systems.

Although useful to a cancer policy, the guidance provided by the weight of evidence criteria just described may still not address all concerns relevant to a proper determination of human carcinogenic potential. Criteria are needed to help avoid arbitrary judgments. They can provide the assessment process with a basis for deciding which chemicals clearly should or should not be considered as human carcinogens. Often, however, the carcinogenicity data bases associated with a particular substance are not easily amenable to strict classification. A thorough analysis of the data is generally required in these situations.

Table 15.9

USEPA Classification System for Assessing the Overall Weight of Evidence for Human Carcinogenicity*

		Strength of Evidence from Animal Studies				
		Sufficient	Limited	Inadequate	No Data	Negative
Strength of Evidence from Human Studies	Sufficient	A	A	A	A	A
	Limited	B1	B1	B1	B1	B1
	Inadequate	B2	C	D	D	D
	No Data	B2	C	D	D	E
	Negative	B2	C	D	D	E

* Quoting the USEPA, "the above assignments are presented for illustrative purposes. There may be nuances in the classification of both animal and human data indicating that different categorizations than those given in the table should be assigned. Furthermore, these assignments are tentative and may be modified by ancillary evidence. In this regard all relevant information should be evaluated to determine if the designation of the overall weight of evidence needs to be modified. Relevant factors to be included along with the tumor data from human and animal studies include structure-activity relationships, short-term test findings, results of appropriate physiological, biochemical, and toxicological observations, and comparative metabolism and pharmacokinetic studies. The nature of these findings may cause an adjustment of the overall categorization of the weight of evidence."

Group A Human Carcinogen
Group B Probable Human Carcinogen
  Group B1 - Reserved for agents for which there is limited evidence of carcinogenicity from epidemiologic studies.
  Group B2 - Reserved for agents for which there is sufficient evidence from animal studies and for which there is inadequate evidence, or no data from epidemiologic studies.
Group C Possible Human Carcinogen
Group D Not Classifiable as to Human Carcinogenicity
Group E Evidence of Non-Carcinogenicity for Humans

Source: USEPA (U.S. Environmental Protection Agency). 1986a, "Guidelines for Carcinogen Risk Assessment," Federal Register, Vol. 51, pp. 33991-34003.

Table 15.10

IARC Classification System for Assessing the Overall Weight of Evidence for Human Carcinogenicity

		Strength of Evidence from Animal Studies				
		Sufficient	Limited	Inadequate	No Data	Negative
Strength of Evidence from Human Studies	Sufficient	1	1	1	1	1
	Limited	2A	2A*/2B	2A*/2B	2A*/2B	2A*/2B
	Inadequate	2A**/2B	2B+/3	3	3	3/4+
	No Data	2A**/2B	2B+/3	3	3	3/4+
	Negative	3	3	3/4+	3/4+	4

* = "exceptionally" solely on the basis of limited evidence in humans.
** = "exceptionally" when strengthened by supporting relevant data.
+ = "in some circumstances" when strengthened by supporting relevant data.

Group 1 Carcinogenic to Humans
Group 2A Probably Carcinogenic to Humans
Group 2B Possibly Carcinogenic to Humans
Group 3 Not Classifiable as to its Carcinogenicity to Humans
Group 4 Probably not Carcinogenic to Humans

Source: H. Vaino, Chief, Unit of Carcinogen Identification and Evaluation, International Agency for Research on Cancer, Letter to Elizabeth Bourque, Massachusetts Department of Public Health, January 20, 1987.

15.4.3 Reproductive and Developmental Toxicity.

The weight of evidence determination for reproductive effects concerns whether or not a substance can be considered as a human reproductive toxin. A similar determination is made regarding developmental effects. These determinations rely primarily on evidence from animal studies, as reproductive and developmental toxicity endpoints have been inadequately monitored in human populations (Workshop, 1986; Dixon, 1986). In contrast to the weight of evidence determinations for carcinogenicity, no international or federal criteria exist to evaluate the weight of evidence regarding these effects. As a general principle, a substance associated with a significant increase in a developmental effect in any well conducted animal bioassay should indicate a human health concern. Furthermore, the substance should be considered as a human reproductive or developmental toxin unless it can be adequately demonstrated that the response occurred by chance or that the observed response is not relevant to human risk assessment.

The probability that an adverse reproductive outcome occurred by chance depends in part on the number of endpoints evaluated. Although a general consideration for the evaluation of all health effects, chance occurrence is particularly important in relation to these effects. Several toxicological endpoints have been identified (See Chapter 11.2.3). Furthermore, many of these endpoints are not independent events. Examples of this interdependency include the relationship between fetal weight and litter size, and the competition between teratogenicity and embryolethality as the dosage levels increase. Such considerations need to be addressed when assessing the biological significance of elevations in adverse reproductive outcomes.

When assessing the relevance of an adverse reproductive or developmental effect, the degree of consistency among the studies should also be considered. With regard to animal studies, the strength of the association between exposure and effect increases with the proportion of positive studies and the severity of the response observed. Special consideration is warranted when similar studies produce conflicting results. In the absence of adequate data on human beings, however, positive findings from a well conducted animal bioassay may constitute a sufficient basis for presuming that the substance could cause adverse developmental effects in human beings.

The human data base, although generally limited, should also be considered in the weight of evidence determination. A small number of chemicals have been directly identified as posing reproductive or developmental risks to human beings (See, for example, Schardein, 1985; Dixon, 1986). Also, although laboratory animals may be considered sufficiently sensitive as far as detecting potential human teratogens, they may vary considerably in their specificity to chemicals with no known human teratogenic potential. Only 28 percent of the chemicals no known human teratogenic potential were also found to be negative in all animals tested for teratogenicity (Brown and Nigel, 1983). It is currently unknown the extent to which this reflects a true lack of specificity in the animal models, differences in dosing regimens, or the limited nature of the human data base for teratogens (Brown and Nigel, 1983; IRLG, 1986; NRC, 1986). Therefore, human evidence is usually inadeqate to negate the findings of reproductive or developmental toxicity in experimental animals.

15.4.4 Other Acute/Chronic Toxicity.

Weight of evidence determinations may be made for effects other than mutagenicity, carcinogenicity, or reproductive and developmental toxicity. Health concerns associated with exposure to toxic substances encompass more than these three endpoints. Often, however, attention is given to these effects because they represent sensitive indicators of an adverse health impact in response to chemical exposures.

Any substance can be toxic if it is administered at high enough doses, as Paracelsus noted several hundred years ago. A substance may induce different toxic responses through different mechanisms, however. Hexane, for example, may directly cause narcosis; it may also cause peripheral neuropathy after being metabolized to 1,2-hexanedione (Clayton and Clayton, 1982). Toxicological mechanisms which are responsible for acute lethality, for example, may not be relevant to the evaluation of subtler acute or chronic effects (such as sensitive, but biologically significant, neurologic or immunologic toxicity). Yet, information is not generally available regarding the subtle, non-carcinogenic, long-term effects of chemicals on different organ systems. This may represent a serious deficiency in the toxicological data base for a chemical. A discussion of the weight of evidence for such toxic interactions, therefore, should address the extent of the data base, and the sensitivity of tests in which positive results were observed.

In the absence of adequate human evidence, the strength of the association between exposure and a particular effect increases with the proportion of animals showing a positive response. It should be emphasized, however, that there is no a priori method to determine whether or not a response in a laboratory animal is indicative of a human response. As with the assessment of carcinogenic or developmental effects, animal models may differ markedly in their abilities to identify acute or chronic effects in human beings (Connelly and Bridges, 1986). While it is possible that a substance may produce an effect in laboratory animals which is not relevant to human beings, it is also possible that several, if not all, of the animals tested may not be as susceptible as human beings to a chemical exposure. Examples of cases in which animal models have failed to predict human responses are presented in Table 15.11. Unless there is sufficient information to the contrary, therefore, it must be assumed that human beings are at least as susceptible as the most susceptible animal species to chemical exposure effects.

Table 15.11

Some Failures in Laboratory Species to Predict Human Toxicity

Agent	Toxic Outcome Observed in Human Beings	Laboratory Species Testing Negative
Azauracil	CNS Disturbances	All Mammals Tested
Benzene	Leukemia	All Mammals Tested
Clioquinol	Eye Lesions	All Mammals Tested
Dinitrophenol	Cataracts	All Mammals Tested
Napthylamine	Bladder Cancer	Rats and Rabbits
Oral Contraceptives	Thrombosis	All Species Tested
Practalol	Skin, Eye, Intestinal Lesions	All Species Tested
Thalidomide	Fetal Abnormalities	Most Rat Strains
TCDD	Chloracne	All but Hairless Mouse

Source: Connelly, J.C., and Bridges, J.W., 1986, "Species Variation in Target Organ Toxicity," in G.M. Cohen (ed.), Target Organ Toxicity, Vol. I, CRC Press Boca Raton, FL, p. 91.

16. Quantitative Evaluation of Dose-Response Relationships.

The quantitative evaluation of the dose-response relationships focuses on those studies which can quantitatively correlate health effects with exposure. First, criteria are needed in order to determine which studies are adequate for quantitative risk assessment. Also, procedures are needed to estimate equivalent doses when the quantitative assessment involves extrapolations between species or between different exposure regimens. The remainder of the quantitative evaluation is concerned with the derivation of action levels. Action levels are derived from studies on relatively small populations of human beings or laboratory animals, but are expected to represent quantitative indices of health concern for the entire population. Therefore, consideration should be given to the intraspecies and interspecies variability in toxicity. This consideration is addressed through two approaches: 1) the application of uncertainty factors to the data in the experimental range and 2) low dose extrapolation of the dose-response curve by using mathematical models.

16.1 Selection of Studies for Quantitative Risk Assessment.

Once a weight of evidence determination has found that there is a positive association between chemical exposure and a particular health effect, an assessment of the degree of associated health hazard as a function of exposure dose should be conducted. Before a quantitative assessment can be done, however, a determination has to be made regarding which of the studies can be used in this assessment. Selection criteria should consider two risk assessment policy objectives: 1) the need for the assessment to provide conservative, worst-case risk estimates in the face of scientific uncertainty, and 2) the need for the assessment to reflect the best scientific understanding of the issue.

The first objective is addressed by selecting the studies which show the most sensitive toxic responses, and by selecting the experimental groups which are the most susceptible to exposure related effects. Identification and quantification of subtle effects reduces the uncertainty in the estimation of action levels. Basing the risk estimates on the most susceptible experimental group reduces the likelihood that the response in human populations is underestimated. The second objective is addressed by selecting studies which are superior in terms of design and conduct, which most closely reflect or predict the human responses, which are most applicable to the exposure route of concern, and which provide the best quantitative estimates of dose-response. The need to consider these two policy objective often results in the selection of multiple studies relevant to a specific health effect. It is a matter of scientific judgment as to which of the resulting risk estimates is the most appropriate for the characterization of human health risks.

Situations often arise in which no studies on a particular substance are considered adequate for quantitative risk assessment. In these cases, risk characterization must rely solely on qualitative considerations. Default action levels, based on the presumption that public health risks should be kept to a minimum when data are inadequate to evaluate risks, may be developed for these substances. These action levels may lack the scientific justification necessary to be considered in a risk assessment policy. They

should, nonetheless, be considered within the context of an overall public health policy.

16.2 Calculation of Equivalent Exposure Units.

In order for toxic effects to be precisely compared quantitatively, equivalent dose units have to be derived between the test population and the human population of concern. The most direct measure of chemical dose involves the concentration of either the chemical or its toxic metabolite of the chemical at a site of action. Often, however, data are insufficient to calculate this value for either population, and surrogate estimates need to be derived. Such surrogates may be the administered dose, the absorbed dose, the absorbed dose adjusted for interspecies differences in pharmacokinetics, or the metabolized dose.

The administered dose of the parent chemical is frequently used as a surrogate for the dose of the active chemical (either the parent chemical or one or more of its metabolites) at the site of action. This dose may be expressed in a variety of ways. Inhalation studies, for example, commonly express administered dose in terms of a mass of chemical per unit volume of air (e.g., milligrams/m^3) or volume of chemical per unit volume of air (e.g. parts per million). Oral studies may express administered dose in terms of the chemical concentration in food or water (e.g. parts per million).

A consideration in the calculation of an equivalent administered exposure dose between the study population and the human population of concern is the need to adjust for differences in chemical uptake. For a given concentration of a substance in air, food, or water, the actual intake differs significantly between human beings and laboratory animals, and between adults, children and infants. Mimimum food, water, and air intake requirements vary in relation to body size (Guyton, 1947; Adolph, 1949). Adjustments are therefore needed to address the relationship between body size and intake rates. Further adjustments may be needed to address other factors associated with intake differences. These include differences in activity levels, and in food and water consumption patterns. Inhalation rates for selected human and rodent populations are presented in Table 16.1. Depending on the comparisons required of the assessment, intake parameters other than those listed in Table 16.2 may need to be considered as well. With this information, the administered dose can be standardized to body mass, such as the mass of chemical administered per unit body weight per day (mg/kg/day).

While the admininstered dose is usually identified in the health effects studies, it may not correlate well with the toxicological response. Reliance on the administered dose may lead to imprecise estimates of human health risks if pharmacological differences associated with different absorption or different metabolic characteristics are not considered. The absorbed and delivered doses may vary with the exposure routes, dose levels, or species. Metabolic differences within species and between species may also signficantly influence the amount of the proximal toxin that reaches the site of action. If the proximal toxin is the parent compound, the concentration of the active compound at the site of action decreases as metabolism increases. If a metabolite is the proximal toxin, the concentration of the active compound at the site of action increases as metabolism increases. The quantitative association between dose and response can thus be greatly improved by efforts to estimate the dose of the chemical actually delivered to the site of action.

Table 16.1

Inhalation Rates for Selected Human and Rodent Populations

Population	Pulmonary Ventilation L/min	Pulmonary Ventilation m^3/day	Alveolar Ventilation L/min	Alveolar Ventilation m^3/day
70 Kilogram Adult[a]	15	22	10	15
10 Kilogram Infant[b]	2.8	4.0	1.9	2.7
Rat (300 grams)[c]	0.15	0.22	0.10	0.13
Mouse (30 grams)[c]	0.027	0.039	0.018	0.026

a Source: I. Astrand, 1983, "Effect of Physical Exercise on Uptake, Distribution and Elimination of Vapors in Man," in V. Fiserova-Bergerova (ed.), Modeling of Inhalation Exposure to Vapors: Uptake, Distribution, and Elimination, Vol. 2, pp. 107-130. These calculations assume 16 hours at rest and 8 hours at a light activity level.

b Source: International Commission on Radiological Protection, 1975, Report of the Task Group on Reference Man, Pergamon Press, Oxford, p. 346. These calculations assume 14 hours at rest and 10 hours at a light activity level.

c Source: A.C. Guyton, 1947, "Measurement of the Respiratory Volumes of Laboratory Animals," Amer. J. Physiol., Vol. 150, pp. 70-77. The allometric equation is:

$$\text{Resp. Vol. (ml/min)} = 2.10 \times BW^{3/4}$$

If information is inadequate to quantitatively estimate an absorbed or metabolized dose for either the study population or the human population of concern, the standardized administered dose represents the best basis from which equivalent exposures can be derived from the available data. When adequate, however, pharmacokinetic data may help to relate the administered dose to the absorbed dose and, further, to the dose of the active compound at the tissue exhibiting the toxic response. For certain well studied compounds, adequate pharmacokinetic data are available from animal studies. Frequently, however, information is lacking from which a qualitative and quantitative profile of human absorption and metabolism can be constructed, particularly at environmental exposure levels. Assumptions consistent with worst case assessment methodologies therefore have to be made regarding the actual extent of human absorption, distribution, and metabolism.

In certain instances, the pharmacological and toxicological data may indicate that a scaling factor applied to the absorbed dose is appropriate when dose comparisons are made between species. A commonly used scaling factor is based on interspecies differences in surface area. The typical surface area adjustment, presented in the following equation, uniformly reduces the weight-adjusted dosage for human beings compared with that determined for smaller laboratory test animals (NRC, 1986).

$$\text{Human Dose (mg/kg)} = \text{Animal Dose (mg/kg)} \times (\text{animal bw/human bw})^{1/3}$$

This adjustment is based on the assumption that surface area is a function of the body weight raised to the two-thirds power. Because of differences in body shape between animals, this allometric relationship is not a precise one (Calabrese, 1983; Lu, 1985). It has been a useful approximation, however, in certain clinical applications. Figure 16.1 demonstrates the effect of the surface area adjustment on the interspecies scaling factor.

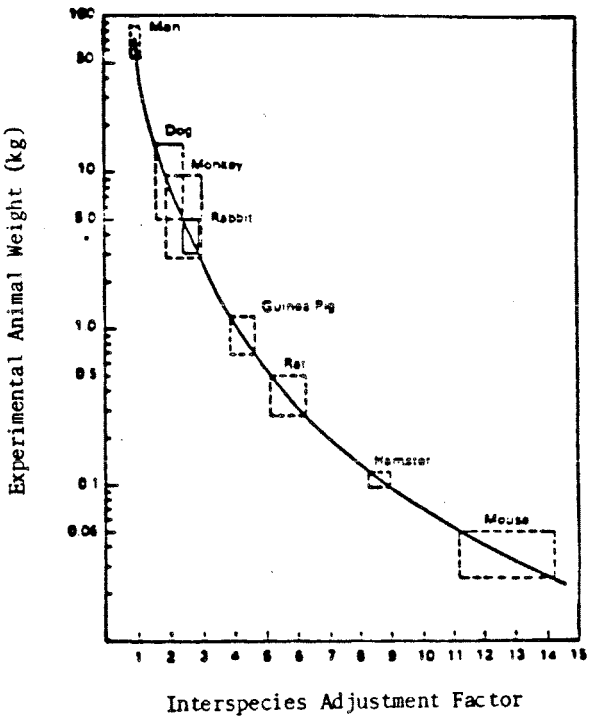


Figure 16.1. Experimental Animal Weights Versus an Interspecies Scaling Factor Using the Surface Area Relationship. Enclosed areas along the function represent general ranges of average body weights of experimental animals. Rabbit values are represented by the box with solid lines.

Source: Dourson, M.L., and Stara, J.F., 1983, "Regulatory History and Experimental Support of Uncertainty (Safety) Factors," Reg. Toxicol. Pharmacol., Vol. 3, p. 229.

The surface area adjustment has been found to remove much of the interspecies variability associated with the toxicity and effectiveness of certain non-metabolizable anticancer drugs (Freireich et al., 1966; Pinkel, 1958). It also produces a stronger correlation among species than a straight body weight extrapolation with regard to several physiological parameters, including blood flow (Ramsey and Anderson, 1984), enzyme activity (Calabrese, 1983), renal clearance (Adolph, 1949), and drug half-lives and concentrations in plasma and tissue (Calabrese, 1983). According to the National Research Council (NRC, 1986), for certain compounds, a specified effect level in human beings may be estimated from the corresponding effect level in rodents after this surface area adjustment is applied.

There are limitations, however, regarding the application of surface area adjustments in interspecies scaling. Specific interspecies differences in absorption, metabolism, distribution, or excretion may render the surface area adjustment inappropriate for quantitative risk assessment. Thus, the adjustment should only be used when there is evidence that adjusting for surface area related differences would provide a better estimate of delivered dose than alternative approaches. It is possible, for example, that a more specific scaling factor than the general surface area adjustment may be derived from the available information. Particularly, if the pharmacokinetic studies show that metabolism is significantly different between the test animals and human beings, interspecies adjustments should reflect those differences. On the other hand, if there are no data to justify the use of a scaling factor, estimates of delivered dose should rely estimates of absorbed or administered dose.

Uncertainties associated with the estimation of the delivered dose of a toxic chemical to the site of action may be reduced through the use of pharmacokinetic modelling. Before such modelling is accepted as the quantitative basis for the assessment of the dose-response relationship, however, the parameters which provide the basis for the model output should be carefully examined. Such consideration is particularly relevant if the models are developed primarily from animal data, or if the findings from high dose exposures are to be applied to low dose exposure situations. Failure to consider these differences could result in imprecise estimates of delivered dose in the human populations of concern.

16.3 Derivation of Action Levels Using Mathematical Models.

Most risk assessment applications require the estimation of health risks at dose levels below those in the experimental dose ranges. Two methods are currently used to address this issue. The first method involves the application of mathematical models to extrapolate the dose response curve below the experimental range. The second method involves the application of uncertainty factors to either the lowest observed adverse effect level (LOAEL) or the no observed adverse effect level (NOAEL) in the study population in order to derive and estimated no adverse effect level (ENAEL) for sensitive human populations.

Of the two low dose risk estimation approaches, the mathematical modelling approach is preferable if the effect is presumed to exhibit no threshold, or if adequate experimental data are available and enough is known about the toxic effect to presume the shape of the dose response curve at sub-experimental doses. This approach is preferable because it makes use of more dose-response data. Therefore, it conforms more closely to the key risk assessment policy objectives of considering all the available data and ensuring that the risk assessment findings reflect the best scientific understanding of a chemical's toxicity. When information is inadequate to extrapolate the dose-response into the low dose region, however, the uncertainty factor method must be employed. While the uncertainty factor approach may not make as much use of the toxicological data as the mathematical modelling approach, it is also true that the available information in these situations does not permit more sophisticated risk assessment techniques to be employed. Thus, when data are inadequate for low dose extrapolation using mathematical models, the best scientific understanding is limited to the relatively crude uncertainty factor method. This method is described following the discussion on mathematical models.

Most practical applications of low dose mathematical modelling have involved the estimation of cancer risks. Cancer risk assessment has relied on this method because its non-threshold assumption is inconsistent with the population threshold assumption implicit in the uncertainty factor method. On the other hand, models have been developed which employ a population threshold assumption with regard to both quantal and continuous data (Crump, 1984; NRC, 1986). These models also consider all the dose-response data and experimental factors such as sample size. They may prove to be very useful if adequate dose-response data are available, particularly if uncertainty exists regarding the determination of the no observed adverse effect level. They are not generally used at the present time, however, primarily because sufficient data rarely exist for such an extrapolation (Dourson et al., 1986; NRC, 1977).

At the present time, carcinogenicity is the only health effect for which a non-threshold assumption has been generally applied. Although genetic toxicity is also presumed exhibit no threshold, this health endpoint is not usually employed in the derivation of action levels. Findings from laboratory studies indicate that intrauterine death, as well as other embryotoxic effects, follow a dose-response relationship with an apparent threshold at some level below the experimental dose range (Wilson, 1980). Therefore, no effect levels for developmental effects in human beings may be estimated from the dose response data or from the application of uncertainty factors to no observed effect or lowest observed effect levels in the animal studies (IRLG, 1986).

No criteria have been developed to determine when mathematical models can be applied to non-carcinogenic effects data. Conversely, criteria have not been sufficiently developed to determine when the uncertainty factor method should be applied to carcinogenicity data. Thus, in the absence of sufficient information to the contrary, low dose extrapolation for carcinogenicity should involve mathematical modelling while low dose extrapolation for non-carcinogenic effects should involve the use of uncertainty factors.

16.3.1 The Selection of the Appropriate Mathematical Model.

Two general types of mathematical models are commonly used in low dose extrapolation for carcinogenic effects: tolerance distribution models (including the probit, logit, and Weibull models), and mechanistic or stochastic models (including the one-hit, multi-hit, and multi-stage models). The tolerance distribution models are based on the assumption that each individual in the population has his own tolerance to the substance, and that these tolerances follow some distribution function, such as the normal distribution. The mechanistic models are premised on the assumption that a positive response for each animal is the result of the random occurrence of one or more biological events. Although these two modelling concepts rest on different premises, it is possible for a mechanistic argument to lead to a tolerance distribution (USEPA, 1987). For example, the gamma multi-hit model follows the same distribution regardless of whether one assumes a random occurrence of "hits", or whether one assumes that each individual in the population has a particular tolerance level to a chemical (USEPA, 1987). Thus, the distinction between the mechanistic and tolerance distribution models is not always clear (USEPA, 1987).

Adaptations of the tolerance distribution and mechanistic model types have been developed for certain situations. For example, the population threshold models mentioned above are variations on these models (Crump, 1984). Also, time-to-tumor models (including the Weibull distribution, Armitage-Doll, Hartley-Sielken models) are adaptions of the tolerance distribution models, with adjustments made for latency periods (Hogan and Hoel, 1982; Park and Snee, 1983; Munro and Krewski, 1981; Krewski et al., 1983). The time-to-tumor models have the potential to provide more complete information regarding the actual carcinogenic process. Sufficient uncertainty currently exists regarding the ability of bioassays to consistently determine time-to-tumor, however, that their applicability may be limited in most risk assessments (OSTP, 1985; USEPA, 1986a).

The mechanisms of carcinogenicity are still not precisely understood, despite an increased understanding regarding the overall initiation, development, and progression of neoplasms. Thus, both tolerance distribution and mechanistic models lack empirical justification. All models may fit the experimental range generally well, although they may produce widely divergent estimates of risk in the low dose region (USEPA, 1986a; OSTP, 1985). A general description of the basic curves generated by low dose extrapolation models is presented in Figure 16.2. The threshold curve (Curve A) is not currently considered to adequately reflect the risks associated with low level exposure to carcinogens. This conclusion may change for at least certain carcinogens, however, when issues associated with the mechanisms of chemically-induced carcinogenesis become better resolved. Supralinear models, such as those conforming to Curve D, are not generally considered to be biologically plausible (OTA, 1981). Therefore, the most commonly discussed low dose extrapolation models are characterized by either a linear or sublinear curve in the low dose region.

Both linear and sublinear low dose extrapolation models generate curves which cross the origin and therefore, are consistent with a non-threshold assumption for the population. Determining the shape of the curve in the sub-experimental range, however, requires consideration of the uncertainty associated with the observed dose-response. Most extrapolation models depend

on few data points. Cancer bioassay results, for example, generally include only two data points in addition to the control. Even when adequate data are available, extrapolation models must consider variations in individual sensitivity. There is, therefore, uncertainty as to the true shape of the dose-response curve. The upper 95 percent confidence limit is generally used to estimate the upper bound on what the true dose-response relationship is. If the normal distribution of responses in the exposed population reflects the general population distribution, then there is only a 5 percent probability that the response of the general population is greater than that predicted by using the upper 95 percent confidence limit of the dose response curve.

Another concern which must be addressed when determining the shape of the cancer dose-response curve at sub-experimental doses is whether the chemical is acting independently or additively to the background response. Because chemical carcinogens do not behave uniquely, it is likely that any exposure could act in a simple additive way with background risks (Brown, 1975; Crump, 1985). Although certain studies (e.g. Russell et al., 1982) indicate that sublinear responses to low doses of carcinogens may occur, the shape of the dose-response curve in the low dose region is an still unresolved issue (NRC, 1986). After a review of this issue, the National Research Council has concluded that, as a general rule, low dose extrapolation models for chemical carcinogens should incorporate the assumption of background additivity (NRC, 1986). The assumption of additivity generally leads to the assumption of low dose linearity as well (Munro and Krewski, 1981).

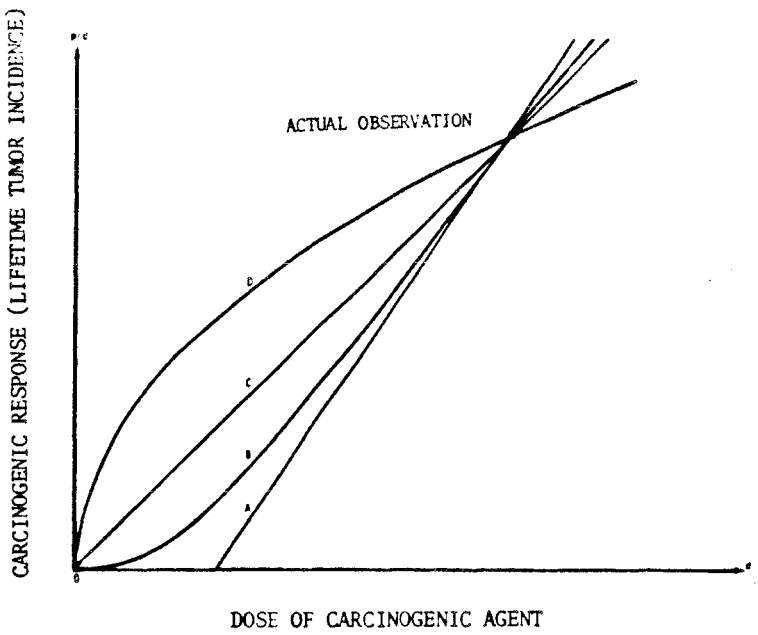


Figure 16.2. Four Dose-Response Relationships Postulated for Low-Dose Extrapolation. a) Threshold; b) Probit (Sublinear); c) Linear; d) Gamma - k = 0.5 (Supralinear).

Source: Tomatis, L., et al., 1982, "Experimental Studies in the Assessment of Human Risk," in D. Shottenfeld and J.F. Fraumeni, Jr. (eds.), Cancer Epidemiology and Prevention, W.B. Saunders Company, Philadelphia, p. 68.

Of the mechanistic models, the multistage models are the most commonly used and accepted (Peto, 1977; OSHA, 1980; USEPA, 1986a). Conceptually, they are the most consistent with our current understanding of the multistage nature of carcinogenesis (Day and Brown, 1980; Hogan and Hoel, 1982). A specific variation of these models, the linearized multistage model, incorporates the assumption of low dose linearity through restrictions placed on its upper 95 percent confidence limit. The upper 95 percent confidence limit of the dose-response curve is determined by maximizing the linear term of the model. Of the other models, the multi-hit model is generally less conservative than the multistage model (Munro and Krewski, 1981). It has also been criticized as having significant practical problems with its application, including the estimation under certain situations of a supralinear dose-response in the low dose region (OSTP, 1985). The one-hit model is more conservative than the multistage model, although it has been criticized as not adequately reflecting tumor responses in the experimental region (OSTP, 1985). A modification designed to address this shortcoming of the one hit model is a straight linear extrapolation from the lowest observed response rate to the background rate or origin (Hogan and Hoel, 1982; Hallenbeck and Cunningham, 1986; Gaylor, 1985). The only assumption needed is that the dose-response curve from which the linear extrapolation is made is convex. Like the one-hit model, however, the linear extrapolation technique has been criticized as not being sufficiently reflective of the observed data and for being overly conservative and for not making use of all the available dose-response data (Hogan and Hoel, 1982; OSTP, 1985).

The tolerance distribution models have been used in describing a variety of toxicological phenomena. The probit model, for example, is used in the detemination of median lethal dose levels levels among experimental animals (Tomatis et al., 1982). These models, therefore, may have applicability when evaluating response patterns for threshold effects. Despite the fact that it precisely reflects laboratory data in the observed response region, this model tends to become very flat at low doses, and therefore predicts risk values that are substantially lower than those calculated from the linearized mechanistic models. An adaptation of the probit model was developed by Mantel and Bryan (1961) to address this possible shortcoming. According to this model, carcinogenic risks are estimated from an upper confidence limit on an extrapolated dose-response curve with a slope of one. Although this modification was considered by the authors to add sufficient conservatism into the modelling process, this may not be necessarily true. The logit and Weibull models generally tend to be more conservative than the probit model at low doses, with the Weibull model being the most conservative of the three (USEPA, 1987).

The selection of the appropriate low dose extrapolation model for cancer risk assessment purposes thus depends on a number of considerations. Firstly, the selection must conform to the general risk assessment policy objectives of adopting a conservative approach regarding scientific uncertainty, of considering as much of the relevant available data as possible, and of ensuring that the risk assessment findings reflect the best scientific understanding of the chemical's toxicity. Also, the selection should specifically consider that the upper 95 percent confidence limit on risk be used to account for uncertainties in the dose-response, and that the dose-response curve be linear at low doses.

The tolerance distribution models may not be suitable for two reasons: they are not as conservative as the multistage model, and they may not generate a linear dose-response curve at low doses, even if the upper 95 percent confidence limit on risk is used (Munro and Krewski, 1981). With the exception of the multi-stage model, the mechanistic models warrant concerns because they generally do not consider all of the available dose-response data. On the other hand, the linearized multistage model is relatively conservative and ensures low dose linearity when risks are based on its upper 95 percent confidence limit. Also, it considers all of the available dose-response information and is conceptually consistent with the current scientific understanding of the multistage process of carcinogenicity. In the absence of sufficient to indicate the use of an alternative model for low dose cancer risk assessment, therefore, the linearized multistage model should be selected.

Despite the widespread use of mathematical models in cancer risk assessment, there is still considerable uncertainty concerning the role of low level chemical exposure in the development of human cancers. For the most part, these models rely on general curve fitting procedures with basic assumptions concerning biological variability and the activity of carcinogens in the sub-experimental dose range. It is difficult, for example, to quantitatively correlate mathematical modelling with possible biological differences between early and late stage, genotoxic and non-genotoxic, or chemical and biological carcinogens. It is also difficult to quantitatively consider other potential factors such as genetic susceptibility, hormonal balance, immunologic competence, and nutritional influences. Ongoing research is directed towards the development of mathematical models which more closely reflect human carcinogenesis than the current models (Armitage, 1985). Future risk assessment policy revisions should consider these newer models within the general policy objectives of worst case risk assessment and having the assessments reflect the best scientific understanding of a chemical's toxicity.

16.3.2 The Determination of Appropriate Model Input Information.

Several variables concerning dose and response are required as input into low dose extrapolation models for carcinogenesis. Specifically with regard to the linearized multistage model, these variables include the number of animals per dose group (including the control group), the number of animals with the tumor of interest in each group, the dose levels, the degree of the polynomial, and the type of risk to be determined (extra risk or additional risk). The process for deriving appropriate dose estimates is discussed in Chapter 16.2. Policy issues concerning the process by which appropriate response parameters are estimated focus specifically on the proper assessement of tumor incidence. Issues identified thus far by the Maine Bureau of Health pertain to the selection of the appropriate biological endpoints, to adjustments which may be necessary regarding the raw tumor incidence data, and to the use of additional or extra risk. Other policy issues associated with the choice of model inputs will be incorporated into the policy if discussion is determined to be necessary.

Selection of Appropriate Biological Endpoints. One issue associated with the estimation of tumor incidence in response to chemical exposure concerns how tumors are grouped for quantitative assessment. With regard to animal bioassays, two basic approaches are generally available. The first approach to grouping tumors considers that cancer is a group of distinct neoplastic

diseases, and that the endpoint of interest is the elevation of a specific tumor type or site. According to this approach, incidence data for any tumor site or type may be used in low dose extrapolation. The conclusion could be made from the results of this extropolation that the risk associated with any specific tumor site and type should be no greater than the risk associated with the most sensitive type/site response. The major limitation of this approach is that the resulting risk estimates may not adequately predict the increased incidence of any type of cancer associated with exposure to a potential carcinogen. On the other hand, data are generally available regarding the tumor incidence rates at specific sites, thus making this approach applicable for routine risk assessment purposes.

The second approach to tumor grouping considers that cancer is the endpoint of interest and that elevation in overall cancer incidence is the appropriate parameter for quantitatively assessing the cancer response. Measures of this response include: 1) comparisons of total tumor bearing animals among the exposed and control populations, and 2) the pooling of all animals with significantly elevated tumor sites or types relative to the untreated controls. The first measure of overall cancer response is subject to the limitation imposed by the fact that some sites (for example, the rat testes) have a very high background incidence. This high background incidence could mask the significance of a treatment related effect between treated and control populations unless, possibly, the tumor response is measured relatively early in the study (IARC, 1980). The limitation imposed on quantitative assessment by tumor types with high background rates may be overcome by using the second measure of the overall tumor response. Pooling significantly elevated tumor sites or types is recommended by the U.S. Environmental Protection Agency (USEPA, 1986a) and, when adequate data are available, represents an appropriate measure of total carcinogenic risk. On the other hand, the exact number of animals with significantly elevated tumor responses may be difficult to determine from the bioassay data. Pooling should address the requirement that an animal with a significantly elevated tumor response at multiple sites not be counted more than once. When multiple sites are significantly elevated, information is often unavailable concerning the incidence pattern in each animal. In addition, depending on the tumor rates for the individual sites or types in the control animals, pooling the data may result in a less conservative estimate of cancer risk than an estimate based on the most sensitive tumor response.

Cancer risk estimates derived from epidemiological studies may use both the tumor-specific approach and the approach of pooling significantly elevated tumors at specific sites. The issues associated with the selection of tumor incidence data from epidemiological studies are similar to those associated with the animal bioassays. Yet, before epidemiological data can be used in modelling procedures, additional concerns specific to the analysis should be addressed. These concerns involve variations in background tumor rates attributable to such factors as age, lifestyle, genetic make-up, and other potentially confounding influences.

The choice of whether or not to use the specific tumor incidences depends on the data availability. If adequate data are available to pool the tumor incidences at specific sites, the use of these data into low dose extrapolation models could provide an estimate of overall cancer risk. The process of estimating overall cancer risks when adequate information is available is consistent with basic risk assessment policy objectives.

Inclusion of pooled data does not necessarily mean, however, that the resulting cancer risks estimates are given preference to those based on tumor specific information. Such a decision should also consider the design and conduct of the study, the specific tumors involved, and the relative degree of conservatism afforded by these two approaches.

Another policy issue pertaining to the quantitative evaluation of tumor incidence is the segregation of tumors according to their stages of development. Given that benign tumors for a specific site and type usually represent a stage in the development of malignant tumors, the incidences of benign and malignant tumors generally should be combined in order to derive a conservative estimate of potential cancer risk. On the other hand, separate analyses of these tumor stages could provide a better understanding of the severity of the carcinogenic response. Both combined and separate analyses, therefore, could be complementary to each other as long as the objectives of worst case risk assessment are fulfilled. Guidelines for combining specific benign and malignant tumors have been developed by the National Toxicology Program (NTP, 1984), and should be followed so that inappropriate groupings of benign and malignant tumors can be avoided.

Adjustments of Tumor Incidence Data from Animal Bioassays. After the appropriate biological endpoints for mathematical modelling purposes have been determined, adjustments in the tumor incidence rates may sometimes be required. An important adjustment addresses premature mortality. In light of the latency period associated with carcinogenicity, it is possible that some animals may die before a tumor has had a chance to develop. Thus, when determining the actual number of animals at risk in the bioassay, these early deaths need to be excluded. Such a consideration also applies to the analysis of interim sacrifices, as it is unknown what proportion of the non-tumor bearing animals would have developed tumors had they been allowed to live until the end of the study. The issue of premature mortality can be most readily addressed by elimating from the analysis all animals which died before the appearance of the first tumor and all interim sacrifices.

Additional Risk Versus Extra Risk. The final consideration regarding model input information is whether the cancer risk estimates should be based on additional risk or extra risk. The additional risk approach estimates the cancer risk over background:

$$\text{Additional Risk} = P(d) - P(0).$$

The extra risk approach estimates the cancer risk among those individuals who would not develop cancer in the absence of exposure to the carcinogen:

$$\text{Extra Risk} = (P(d) - P(0))/(1 - P(0)).$$

Extra risk is used to account for tumors with high background rates relative to the human population (for example, liver tumors in B6C3F1 mice). In these situations, extra risk estimates should be used, as this adjustment addresses potential inherent interspecies differences in susceptibility to cancer. When background tumor rates in experimental animals are very low, there is little difference between cancer risk estimates based on additional risk and extra risk.

16.3.3 Presentation of the Mathematical Modelling Results.

Given basic input information, mathematical models may produce a variety of information outputs. The linearized multi-stage model output provides estimates of parameters describing the shape of the dose-response curve, and the upper 95 percent confidence limit on that curve. It also provides estimates of dose levels corresponding to a given level of risk, and risk levels corresponding to a given dose. The specification of action levels for carcinogens varies from agency to agency. As a general policy, however, the Maine Bureau of Health considers an exposure level corresponding to a lifetime cancer risk of one per one hundred thousand as a basis for a public health concern. At a minimum, the exposure level corresponding to this cancer risk level should be presented.

The equation for the maximum likelihood estimate (MLE) describes the shape of the dose-response curve which provides the best fit to the experimental data. This value represents only a point estimate of risk and lacks the confidence associated with the upper bound risk estimate. Nonetheless, the MLE should be presented along with the 95 percent confidence limit in order to provide an indication of the correspondence between the upper bound on risk and the experimental data. At a minimum, the cancer risk estimates presented should include those based on worst case assumptions, as such estimates are necessary if the risk assessment is to reflect a conservative approach to scientific uncertainty. It is possible, however, that more than one MLE and upper confidence limit on risk may be presented, depending on the uncertainty regarding the input variables. Inclusion of multiple risk estimates may add to the assessment by providing a quantitative indication of the range of uncertainty associated with the estimation of potential human cancer risk.

Cancer risk estimates derived from animal bioassays are usually based on tumor incidences associated with lifetime exposures. In addition, most epidemiological studies consider exposure durations which last for a significant portion of a normal human lifespan. These exposure periods should be reflected in the presentation of the human cancer risk estimate by basing the estimate on lifetime exposure. Often, cancer risk estimates may be needed for less than lifetime exposure periods. As there is generally inadequate information from which to compare cancer risks resulting from long-term exposure to those resulting from shorter term exposure, cancer risks for less than lifetime exposures should be based on the product of exposure dose and exposure duration unless a more precise estimate of cancer risk can be derived. Also, for cancer risks to be translated from a mg/kg/day dose to an exposure concentration in an environmental medium, a reference body weight is needed for the general population. A general procedure in this regard is to base cancer risks on a reference adult weight of 70 kilograms unless another reference weight is shown to be preferable to a given risk assessment.

16.4 Derivation of Action Levels Using the Uncertainty Factor Approach.

Uncertainty factors should be used when available chemical-specific health data are inadequate to estimate either a no effect or a de minimus exposure level for a threshold effect in the general population. The uncertainty factor approach rests on two basic assumptions. The first assumption is that a population threshold exists below which chemical exposure causes no adverse health impacts. The second assumption is that for any given area of

uncertainty, chemical exposure produces the most toxic response that can reasonably be estimated from the available health data. Current scientific understanding of chemical carcinogenesis does not generally allow for the assumption of a population threshold. Therefore, the uncertainty factor approach should only be used in the assessment of potential human cancer risks when sufficient scientific justification is available, or when interim guidance is needed for suspected carcinogens which lack an adequate quantitative data base for cancer risk estimation purposes (See Chapter 16.4.5).

Risk extrapolation procedures for non-carcinogenic effects involve the identification of sensitive health effects, the determination of either a lowest observed adverse effect level (LOAEL) or a no observed adverse effect level (NOAEL), and the application of appropriate uncertainty factors to address the probability that certain subgroups of the population may respond more sensitively then what is indicated by the experimental data. It then follows that a level below that which causes the most sensitive response in sensitive populations should also protect human beings from all adverse effects resulting from exposure to a particular chemical.

Four concerns may have to be addressed by the uncertainty factor approach in reference to any given health effect: 1) the estimation of a response level in sensitive human populations from the available epidemiological or animal toxicology studies, 2) the estimation of a no adverse effect level in sensitive human populations when only LOAELs are available, 3) the estimation of a chronic no adverse effect level when only subchronic data are available, and 4) the estimation of a no adverse effect level for a sensitive health effect when only data on relatively more severe health effects are available. Depending on the nature of the data base, some or all of these concerns should be addressed in the risk assessment. The final outcome of this procedure is the derivation of an estimated no adverse effect level (ENAEL) for the general population with regard to a particular sensitive health endpoint and exposure duration.

The development of appropriate criteria for use in the uncertainty factor approach is presented in the following discussion. In certain cases, the risk assessment may be unable to use the uncertainty factor approach because of an inadequate health effects data base. Interim guidance may still be needed, however, for risk management purposes. Recommendations concerning procedures for developing interim guidance on inadequately studied chemicals are discussed in Chapter 16.4.5.

16.4.1 Consideration of Intraspecies and Interspecies Variability.

The major area of uncertainty in the derivation of action levels for threshold effects concerns the potentially wide range of population variability in both susceptibility and sensitivity to the adverse effects of chemical exposure. Few chemicals have been so extensively studied that their entire range of health effects in human populations can be precisely defined. Therefore, an adverse observed in an animal population should also be presumed to occur in at least some human population, unless adequate information exists to demonstrate that the effect is a species-specific response. If human data are available, the policy issue concerns the potential range of variability between the study population and a sensitive human population. In estimating

these ranges in variability, it is reasonable to assume that more uncertainty is associated with the extrapolation of animal data than with the extrapolation of human data.

If it is assumed that there is a general qualitative correspondence between human susceptibility to a toxic effect and the occurrence of that effect in some laboratory species, risk assessment concerns then focus on the quantitative correspondence between the response in any given laboratory model or human population and the response in sensitive human populations. If it is further assumed that the response in laboratory animals is quantitatively similar to some human population, the risk assessment policy issue becomes one of determining the potential range of variability between the animal response and the response of a sensitive human population.

According to the procedure just described, consideration of intraspecies and interspecies differences begins with an estimate of the human variability in sensitivity to the toxic effects of chemical exposure. A factor of 10 has commonly be used to account for variations in sensitivity among human populations (NRC, 1977; Dourson and Stara, 1983; Hogan and Hoel, 1982). It is possible, however, that this factor may be inadequate to protect certain sensitive populations. Gillette (1985) noted that even among small groups of test subjects, clearance of environmental chemicals may easily vary 20- to 50-fold. Calabrese (1983) examined variations on human responses to certain xenobiotics as well as differences in risks to diseases for specific sensitive populations. He noted that variations in human responses may range up to two or three orders of magnitude, and that a 10-fold uncertainty factor may sometimes be inadequate for up to 20 percent of the human population.

Indirect estimates of the range of human variability have been made based on animal models. Dourson and Stara (1983) conducted a probit analysis on acute rat lethality data for 490 toxicants. The slopes ranged from 1.4 to 64. They then calculated the dose reductions required to drop the median response (e.g., the LD50) to a level which would result in deaths of only the most sensitive members of the population (LD13). The frequencies of the slopes plotted against the dose adjustments needed to reduce lethality to the LD13 level. These results are presented in Figure 16.3. The authors calculated that for approximately 92 percent of the slopes, a 10-fold reduction in dose would drop the response to below the LD13. Furthermore, they stated that average slope of 7.8 needed only a 2.4-fold reduction in dose to reach this level. On the basis of this information, they concluded that the 10-fold factor may be overly conservative with respect to the "average" chemical, but sufficient to protect the population from most chemicals.

There are limitations associated with the use of such animal models for the purpose of estimating the potential human variability in responses to toxic chemicals. Firstly, the data base used by Dourson and Stara are based on a relatively homogenous population of rats. Furthermore, the toxicological endpoint (LD50) may not correlate well with more sensitive measures of toxicity. Finally, when compared with the human studies, it appears that the animal studies may significantly underestimate the potential human variability.

In light of these considerations, an appropriate conceptual model for assessing intraspecies and interspecies differences is one that focuses on the wide variability in the human responses to chemical exposures. A hundred-fold range of variability would include most populations. In the absence of

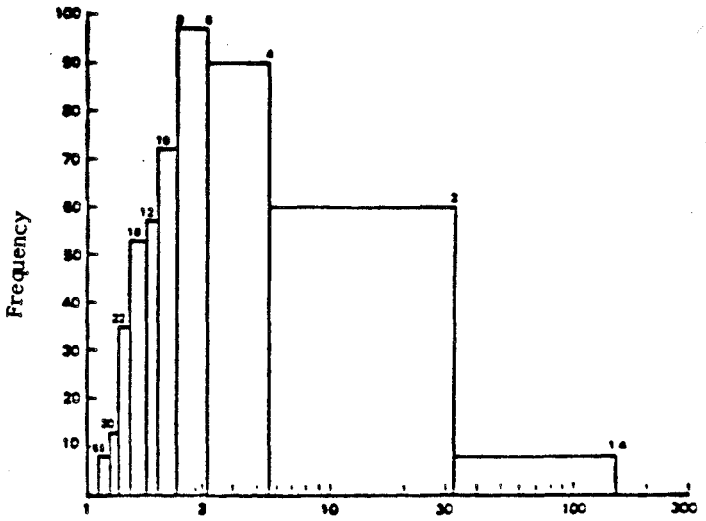


Figure 16.3. Frequency of Studies Versus the Intraspecies Adjustment Factor Required to Reduce the Median Lethal Response in Rats to That of the Most Sensitive Responders.

Source: Dourson, M.L., and Stara, J.F., 1983, "Regulatory History and Experimental Support of Uncertainty (Safety) Factors," Reg. Toxicol. Pharmacol., Vol. 3, p. 227.

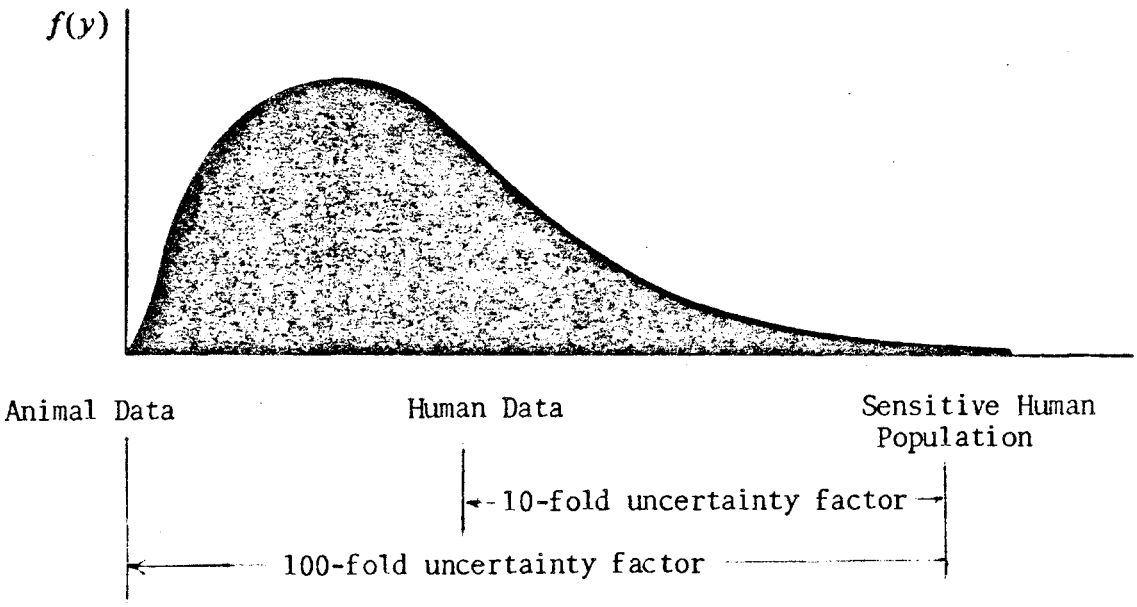


Figure 16.4. A Simplified Graphical Representation of the Relative Uncertainties Associated with Intraspecies and Interspecies Differences in Response to Toxic Chemical Exposure. The graph assumes a lognormal distribution in human sensitivity.

comparative interspecies data, it must be assumed that animals may be up to 100 times less sensitive to the effects of exposure than a sensitive human population. As there is less uncertainty associated with predictions based on human data, it is less likely that the mean response observed in an epidemiological study would represent the least sensitive response among human populations. Rather, it is more likely that a 10-fold uncertainty factor would adequately address potential intraspecies difference among human beings. An illustration of the conceptual basis for using intraspecies and intraspecies uncertainty factors is presented in Figure 16.4.

These uncertainty factors are consistent with those generally used by the National Academy of Sciences (NRC, 1977) and the U.S. Environmental Protection Agency (Dourson and Stara, 1983). The approach just described effectively provides a 10-fold interspecies uncertainty factor to estimate an average response in human populations from animal data. Justification for this factor is provided by studies which have compared similar toxic responses between human beings and laboratory animals (Dourson and Stara, 1983). Such comparisons have raised speculations that interspecies differences may result from factors correlated with animal surface area (Dourson and Stara, 1983). If this is true, it is also possible that species are equally sensitive to a given chemical when the doses at the site of action are equivalent. Thus, if adjustments have been made to account for interspecies differences in delivered dose, and it can be demonstrated that the experimental animal is as equally sensitive as human beings after such adjustments have been made, then a modification of this uncertainty factor may be warranted. Often, however, the distinction cannot be made concerning the relative contributions of delivered dose differences and inherent differences in sensitivity to the overall assessment of interspecies variability. In these cases, the uncertainty factor for animal data should not be modified, regardless of the way in which equivalent dose units are expressed.

16.4.2 Databases Which Lack NOAELS.

Another issues associated with quantitative assessment using the uncertainty factor approach is the derivation of the estimated no adverse effect level for exposed populations in studies which lack adequately derived no observed adverse effect levels (NOAELS). The NOAEL is dependent on the sample size, as the two are inversely related to one another (Munro and Krewski, 1981; Hogan and Hoel, 1982). If, for example, the actual probability of a toxic response occurring at a specific dose is 7 percent, and if 10 animals were exposed, there would be a 50 percent chance that all of the exposed animals would fail to show the response (Hogan and Hoel, 1982). Thus, if the data do not contain a NOAEL for a particular effect, or if the postulated NOAEL is in question, an uncertainty factor should be applied to the LOAEL.

Confidence in the estimation of no adverse effect levels may be increased through analysis of the shape of the dose response curve. An assessment of the curve's trend helps in some cases to distinguish between a lowest observed adverse effect level (LOAEL) and a no observed adverse effect level (NOAEL). Structure-activity data may also provide additional information for evaluating the dose-response curve.

Comparisons have been made regarding the relationships between LOAELS and NOAELS, based on rat toxicology studies involving 33 different chemicals (Dourson and Stara, 1983). A summary of the results is presented in Figure 16.5. The results show that all ratios were 10 or less, with most being 5 or less. The quality of these data are subject to the same caveats as those associated with the quality of the NOAEL data. Also, it is unknown to what extent these comparisons reflect true ratios of LOAELS to NOAELS and to what extent these findings are influenced by general dose selection procedures. Nonetheless, on the basis of this limited data base, it seems reasonable to apply a factor of 5 or 10, depending on the available dose-response information, when estimating a no effect level from a lowest observed effect level.

16.4.3 Databases Which Lack Chronic NOAELS or LOAELS.

A third concern relevant to the uncertainty factor approach involves the estimation of chronic action levels in the absence of adequately conducted chronic effects studies. In these cases, such estimates must rely on the results from studies using shorter exposure durations. It is possible that exposure limits designed to protect the public from subacute or subchronic effects of a particular chemical may not be adequate if populations are chronically exposed. Therefore, uncertainty factors should be applied to NOAELS or LOAELS identified in subacute or subchronic studies. These factors address the possibility that longer term exposures may produce adverse health effects at lower concentrations than those which produced subacute or subchronic effects.

The quantitative basis for this uncertainty factor comes from a studies which compared the lowest (signified in the study as "minimal") observed effect levels in rats exposed to the same chemical for varying exposure periods (Weil et al., 1969; Weil and McCollister 1963). Specifically, these studies compared the lowest observed effect levels in rats for 7-day (subacute), 90-day (subchronic), and 2-year exposure periods. The comparisons used data on 20 compounds, predominantly pesticides and surfactants. Weil et al. (1969) derived equations to predict the median and upper 95 percent confidence limits on the lowest observed effect levels to adjust from a 7-day exposure to a 90-day exposure and from either a 7-day or 90-day exposure to a 2-year exposure. Their findings are presented in Table 16.2. The findings of the Weil and McCollister (1963) study, as adapted by Dourson and Stara (1983), are presented in Figure 16.6.

These comparisons (Weil et al., 1969; Weil and McCollister 1963; Dourson and Stara, 1983) indicate that a factor of 10 should be adequate to adjust NOAELS or LOAELS from either a subacute to a subchronic exposure duration, or from a subchronic to a chronic exposure duration. According to these comparisons, a factor of 35 should be adequate when adjusting from a subacute to a chronic effect. The subacute to subchronic uncertainty factor should be considered with much caution, however, as it is possible that critical health effects associated with longer term exposures would not be adequately identified in subacute tests.

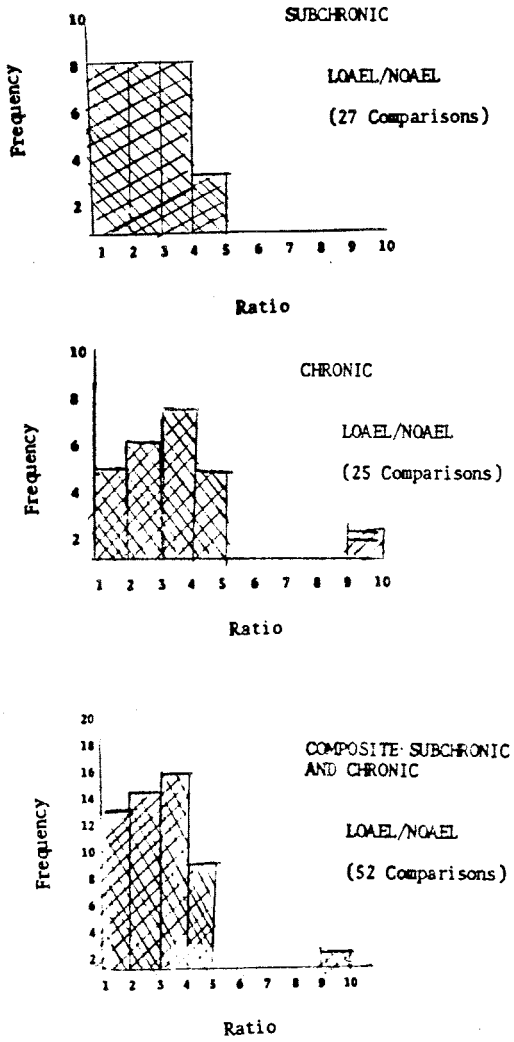


Figure 16.5. Frequency of Studies Versus the Ratio of the LOAEL to NOAEL After Either Subchronic, Chronic, of Composite Subchronic and Chronic Exposures.

Source: Dourson, M.L., and Stara, J.F., 1983, "Regulatory History and Experimental Support of Uncertainty (Safety) Factors," Reg. Toxicol. Pharmacol., Vol. 3, p. 233.

Table 16.2

Predicted Median and Upper Limit to Achieve
a Minimum Effect (M.E.) Level

Value Value	For 90-Day Feeding Study	For 2-Year Feeding Study
Median	$M.E._{.7}/3.0$	$M.E._{.90}/1.8$ or $M.E._{.7}/5.4$
95th Percentile	$M.E._{.7}/6.2$	$M.E._{.90}/5.7$ or $M.E._{.7},/35.3$

Source: Weil, C.S., et al., 1969, "Relationship between Single-Peroral, One-Week, and Ninety-Day Rat Feeding Studies," Toxicol. Appl. Pharmacol., Vol. 14, pp. 426-431.

16.4.4 Databases Without Adequate Information of Sensitive Effects.

When comparative information is available, the uncertainty factor approach should consider the severity and sensitivity of the effect. Some target organs and organ systems, such as the nervous and reproductive systems, may manifest a wide spectrum of toxic responses (NRC, 1986). Databases sufficient to characterize these responses are generally limited, however (NRC, 1986). It may be likely, therefore, that substances associated with neurotoxicity or reproductive/developmental toxicity through limited testing may also cause subtler effects on these systems which current testing procedures do not detect.

No formal guidance has been developed concerning the use of uncertainty factors to address effect severity and sensitivity. Until formal guidance is developed, the determination of whether or not an uncertainty factor should be applied, how large that uncertainty factor should be, is an issue specific to the individual risk assessments. Determinations regarding such uncertainty factors should rest largely on the understanding of the quantitative relationships between progressively more severe toxicological responses associated with a given organ or organ system. If there is no quantitative basis for the establishment of an uncertainty factor, this shortcoming should be discussed in the Risk Characterization (Section VI).

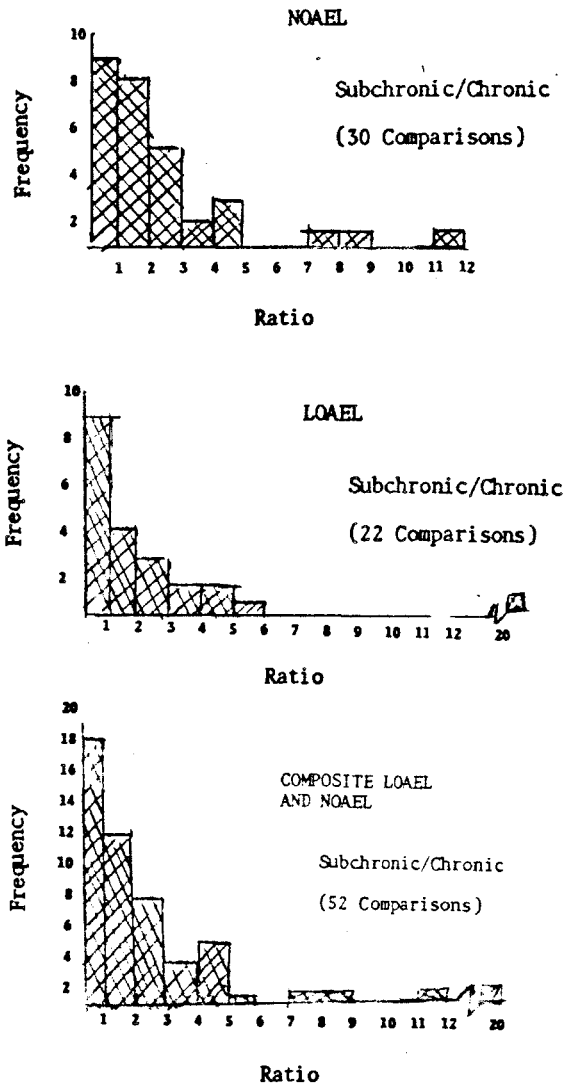


Figure 16.6. Frequency of Studies Versus the Ratio of Subchronic to Chronic Exposures for Either NOAELS, LOAELS, or Composite NOAEL-LOAEL Values.

Source: Dourson, M.L., and Stara, J.F., 1983, "Regulatory History and Experimental Support of Uncertainty (Safety) Factors," Reg. Toxicol. Pharmacol., Vol. 3, p. 231.

General guidance regarding the progression of concern regarding effect severity has been outlined by the Massachusetts Department of Environmental Quality Engineering (MDEQE, 1985). This guidance appears in MDEQE's paradigm acute/chronic effects, presented in Table 15.3. According to this outline, mild or transient effects are assigned a severity value of 1 (or representing the least severe effect). Irreversible effects are assigned the highest severity value of 3. A value of 2 is assigned to intermediate effects.

16.4.5 Default Procedures Using the Uncertainty Factor Approach.

There may be several situations in which no adequate quantitative data are available for quantitative risk extrapolation. If action levels are required of the assessment, the use of alternative approaches is warranted. These may involve the use of structure-activity data or generic default approaches. The State of Michigan (SAAC, 1981), for example, estimated an uncertainty factor of 0.00002 to be applied to LD50 or LC50 data in order to derive an acceptable exposure level for non-criteria air pollutants which fall into this category. This factor was based on the finding by MacNamara (1976) that the upper 95 percent confidence limit of the ratios between the no observed adverse effect level to the LD50 was 0.002. Further application of a hundred-fold uncertainty factor to this upper confidence limit results in the recommended default uncertainty factor.

The use of uncertainty factors is not generally recommended when deriving action levels for known or suspected carcinogens. There may be certain situations, however, when adequate information is unavailable for the quantitative assessment of carcinogenic effects. Factors of as high as 5,000 applied to the no observed adverse effect level (NOAEL) have been recommended to provide interim public health guidance in these situations (Hogan and Hoel, 1982). Such factors have been criticized for being arbitrary and not sufficiently conservative with regard to non-threshold effects. On the other hand, little guidance is available concerning the assessment of potential carcinogens with an inadequate quantitative data base. Until such guidance becomes available, therefore, an uncertainty factor of 5,000 may be used to set an upper bound on exposure when adequate quantitative data are lacking and interim guidance is needed. The lower bound on exposure to known or suspected carcinogens should be zero.

16.4.6 Presentation of the Results from the Uncertainty Factor Approach.

After the appropriate uncertainty factors have been selected and quantified, a summary of the risk derivation process should be presented. This summary should include the estimated no adverse effect levels (ENAELS) for each health effect and exposure duration of concern. The uncertainty factors and interspecies adjustments should be presented in such a way that the quantitative relationship between the ENAEL and the experimental findings is clear. The uncertainty factors factors established in this risk assessment policy are presented in Table 16.3.

The ENAEL may vary depending on the the reference human population which serves as the basis for that level. This is because air, water, and food intake requirements decrease per unit of body weight as body weight increases. Thus, for a given concentration of a chemical in these media, a small child would be expected to receive a greater intake on a mg/kg basis than an adult.

Intake differences are particularly relevant when acute, subacute, or subchronic effects are considered. Chronic effect levels are usually based on studies which are carried out over the lifetime, or a large proportion of the lifetime, of an experimental cohort. The reference human population for these effects, therefore, is usually the average adult weighing 70 kilograms (NRC, 1977). The reference population for acute, subacute, or subchronic effects is generally the 10 kilogram infant (NRC, 1977). When developmental effects are being specifically considered, the reference human population is the 60 kilogram female. The intake doses for these reference human populations should be derived in Chapter 16.2.

Table 16.3

Summary of the Uncertainty Factors Used in the Quantitative Assessment of Threshold Effects*

Parameter	Factor
Interspecies	10
Sensitive Populations	10
LOAEL to NOAEL	5, 10
Subacute to Subchronic Exposure	10
Subchronic to Chronic Exposure	10
Severity/Sensitivity of Observed Effect	variable

* See text for explanation.

SECTION VI: RISK CHARACTERIZATION

Risk characterization, according to the National Research Council (NRC, 1983a), is "the description of the nature and often the magnitude of human risk, including attendant uncertainty." Thus, in the risk characterization section, three areas of risk assessment are integrated: 1) the quantitative relationship between dose and effect, based on the existing data; 2) the quantitative estimates of population exposure to a substance; and 3) the assessment of the critical areas of uncertainty associated with the risk assessment. The characterization of health risks may be largely dependent on the way in which uncertainty is quantitatively integrated into the dose-response assessment and the exposure assessment. The worst case risk assessment reflects the maximum degree to which uncertainty is quantitatively considered. The risk assessment has the obligation to present the worst case assessment, as it represents the highest upper bound estimate of health risk, given the limitations imposed by scientific uncertainty and science policy choices made in deciding worst case assumptions. Characterizations based on assumptions less extreme than the worst case may be presented as part of a general description of the uncertainty surrounding the action levels. In presenting these alternative characterizations, however, the assessment should clearly state both the arguments and evidence which indicate why the alternatives may give a better estimate of the actual risk than the worst case assumption, and the arguments and evidence which support why the the worst case assumption should be favored.

The worst case assumptions which form the basis of the risk characterization may vary according to the policy choices made in deciding what these assumptions are. The risk characterization could therefore benefit from a comparison of the assessment's worst case characterization with those of other assessments. This comparison could provide risk managers with a perspective concerning the areas of scientific consensus and disagreement regarding the risk assessment findings, and the particular risk assessment issues that underlie these areas of agreement and disagreement. In addition, comparisons with existing standards and guidelines, as well as with the health concerns associated with alternatives to the substance of concern in the risk assessment, may provide the risk manager with additional perspectives from which the assessment's findings may be interpreted.

17. Risk Assessment Findings.

The risk assessment findings comprise a summary of what is known and what is not known about the substance's health risks. These findings include the conclusions that can be drawn from the exposure assessment, hazard identification, and hazard assessment sections. The findings are then compared to those of other assessments. Following this comparison is a qualitative appraisal of the existing data base. This appraisal is very important to the assessment, as it provides a context for evaluating the extent to which the derived action levelss reflect overall health concerns. It also provides the context for characterizing the specific risk assessment areas which, if addressed more extensively, would be most likely to lessen the uncertainty associated with the recommended action levels.

17.1 Characterization of Health Risks Associated with Exposure.

The characterization of health risks associated with exposure integrates the findings from the quantitative estimates of exposure (Chapter 9) and the quantitative health risk assessment (Chapter 16). Health risks should be presented for each health effect evaluated and for each exposure duration of concern. The criterion which connects a health endpoint, an exposure level, and an exposure duration is the action level. Depending on the particular health effect, the action level may be expressed as an estimated no adverse effect level (ENAEL) or as a cancer risk level. In either case, the action level, when compared to a particular exposure level and duration, provides the basis for determining whether a current or projected exposure warrants a public health concern.

The quantitative characterization of health risks should be accompanied by a discussion of the qualitative factors related to the derivation of the action levels. These factors include the toxicological importance of the health endpoints, the study populations used in the risk extrapolation (for example, laboratory animals, occupational cohorts, sensitive human populations), the population subgroups most susceptible to exposure-related effects, and the types of risk factors which could influence the substance's toxicity.

In addition to the characterization of human health risks, the environmental and ecological impacts associated with the substance could be characterized. Such characterizations are warranted for two reasons. Firstly, they serve to provide a more complete description of the adverse impacts resulting from the presence of a particular substance in the environment. Secondly, adverse impacts on the environment may indirectly affect human health by causing disruptions in the basis processes which sustain human life. Thus, the complete characterization of the human health impacts associated with exposure to a particular substance should include indirect, as well as direct, human health impacts.

17.2 Comparison with Findings of Previous Assessments.

The comparison of the risk assessment findings with the findings of previous assessments provides a context for evaluating the areas of scientific consensus and disagreement regarding the health risks of a particular substance. Differences may occur as a result of different risk assessment policy choices or because of new findings. The reasons for these differences should be discussed. When differences occur as a result of different policy choices, the discussion should address the scientific issues underlying these choices, and why the policy choice made in the assessment is preferable to the other choices. Specifically, the discussion should focus on comparing the assessments with respect to general policy objectives of considering all the relevant data, of having the assessment reflect the best scientific understanding of the issue, and of having the assessment reflect a conservative approach to the evaluation of scientific uncertainty.

17.3 Appraisal of Current Data Base.

The characterization of human health risks associated with exposure should be understood within the context of what is still not known about the substance's toxicity and exposure potential. Depending on the level of investigation given to a particular chemical, the amount of scientific ignorance may prove to be more significant than the actual data used in deriving the action levels. Every area of a risk assessment could probably be improved by more information. The main objective of the appraisal of uncertainty, however, is to present a summary of those key areas of the risk assessment for which adequate information was found to be lacking. The presentation of this information represents a key element of the risk assessment, for it stresses the importance that the process of evaluating health risks is continually open to improvement. A discussion of these areas of uncertainty associated with the risk assessment should help the risk manager to understand better the differences among risk assessments, and the degree to which future findings could change the current characterization of health risks.

Much of the toxicological data is derived from studies in which human beings or laboratory animals were exposed to relatively high doses of a substance. These studies, at least the human studies, may have direct relevance for individuals in occupational environments. Yet, the environmental situations of most concern are generally characterized by long-term, low level exposures to pollutants. Application of toxicological data to the assessment of ambient exposures, therefore, must consider toxicity differences resulting from different exposure levels, routes, and durations as well as possible interspecies and intraspecies variations in sensitivity. Such considerations are especially important in light of the fact that most risk assessments must rely on data generated from animal studies. The use of animal data introduces uncertainty into the estimation of human risk, as animals may vary widely in their responses to toxins. There is generally no a priori way of determining which animal will be the most predictive of a particular toxic response in human beings. Further uncertainty is introduced if the the animals are exposed to a chemical by a route different from normal human exposures.

There are approximately 70,000 synthetic chemicals in commerce, of which 25,000 are in common use (NRC, 1983a, 1983b). Approximately 500 to 1,000 new chemicals are introduced annually in the marketplace (OTA, 1981). Risk assessments are generally expected to derive acceptable exposure limits for the general population from a limited database. A recent study by the National Research Council (NRC, 1984), however, concluded that adequate information for complete health assessments exists for only 18 percent of 1,900 pharmaceutical products, 10 percent of 3,350 pesticide ingredients, 5 percent of 8,600 food additives, and 2 percent of 3,400 cosmetic ingredients. Only a very small percentage of chemicals in the environment have been tested for carcinogenicity (NAS, 1983), teratogenicity (Schardein, 1985), or neurotoxicity (Anger and Johnson, 1986). Investigations concerning the interactive effects of chemicals have been rare (USEPA, 1984). A summary of the NRC findings is presented in Figure 17.1 (NRC, 1984).

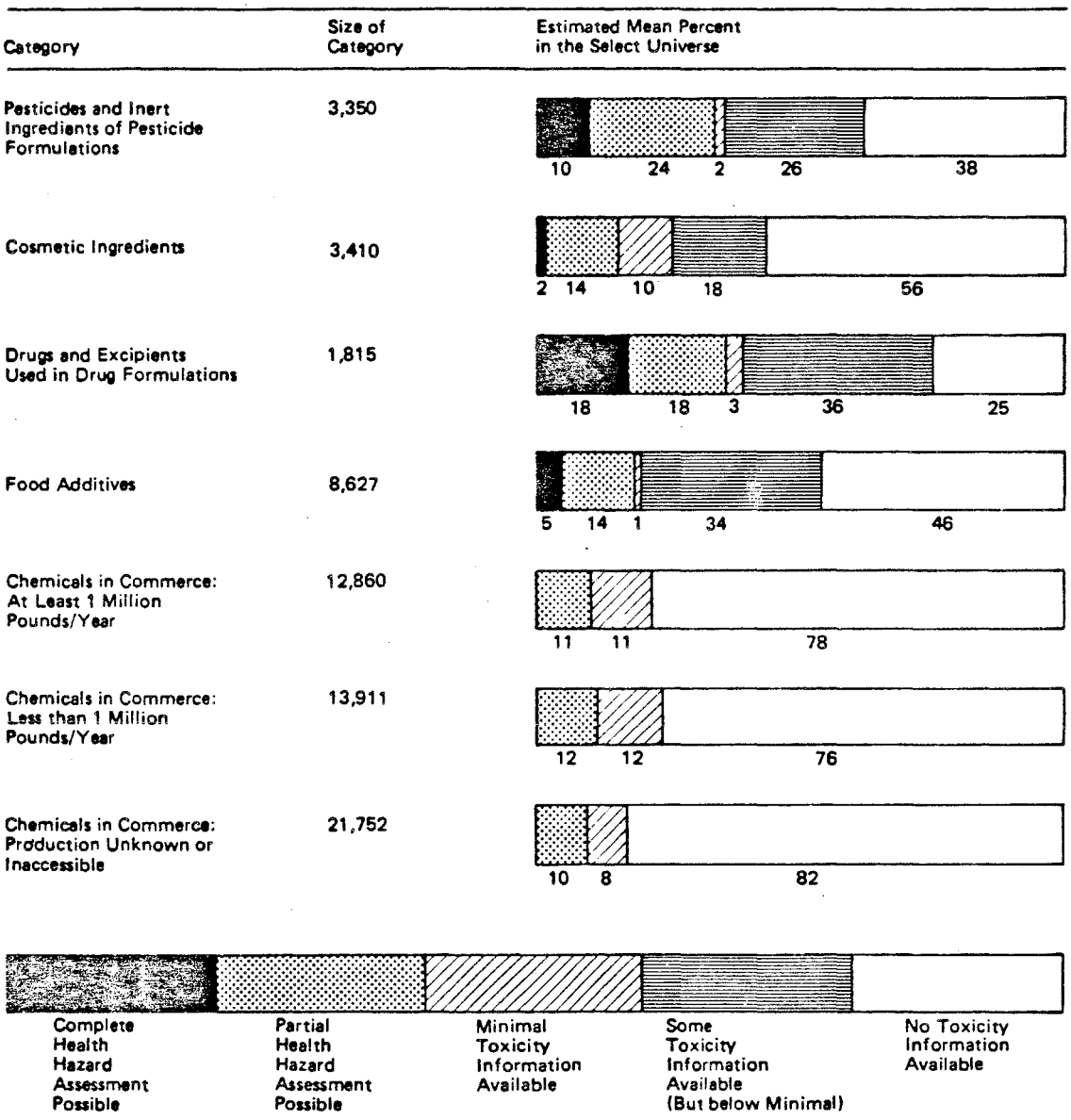


Figure 17.1. Ability to Conduct Health-Hazard Assessment of Selected Substances in Seven Categories.

Source: NRC (National Research Council), 1984, Toxicity Testing: Strategies to Determine Needs and Priorities, National Academy Press, Washington, D.C., p.118.

18. Comparison of Findings with Existing Standards and Guidelines.

Existing standards and guidelines reflect, to varying degrees, the findings of previous assessments. They may differ substantially from assessment findings, however, depending on how largely non-health related concerns were considered in their development. Such concerns may include risk management issues of technological feasibility, the degree of protection to be provided to the population, or the sampling and analytical limitations associated with substance identification and quantification. In addition, some guidelines may be based on default risk extrapolation procedures which are not directly comparable to the low-dose extrapolation procedures used in formal risk assessments. The ways in which the risk assessment findings differ from existing standards and guidelines may need to be explained. Specific attention should be focused on the extent to which the consideration of risk management criteria accounts for these differences.

19. Risks Associated with Alternative Substances.

A concern relevant to the entire process of risk assessment involves the degree of attention given to the substance undergoing investigation relative to other substances. It is possible, for example, that a risk management decision might encourage the use of an alternative substance with equal or even greater toxicity than the substance assessed, but without a corresponding level of analysis applied to it. This concern is especially relevant if the action levels indicate that severe use restrictions may be warranted for the substance undergoing assessment.

Deciding whether or not the health risks of alternative substances should be considered in the risk assessment depends on the assessment's stated purpose and scope. If such considerations are to be made, the assessment should define the level of analysis to be given to these substances. In general, the analysis need only be sufficient to alert the risk manager to the potential health concerns of the alternatives. Particularly, the analysis should emphasize the data availability and general findings for the following health effects categories: genetic toxicity, carcinogenicity, reproductive and developmental toxicity, and other acute/chronic toxicity. This task may be accomplished by a review of secondary sources and, possibly, a literature search using relevant computer databanks. In addition, efforts have been made to develop comparative databases for chemicals which produce similar health effects (see, for example, Gold et al. (1984) and Ames et al. (1987) for carcinogenic compounds). Reference to these databases may be useful in order to provide the risk manager with a broader perspective of the health implications associated with different public policy choices. Separate risk assessments on the alternatives may be required, however, before specific comparisons of health risks can be made between these substances and the substance undergoing assessment.

20. Conclusions.

The conclusions of the risk assessment should present the key findings of the investigation. The ultimate conclusion of the risk assessment should be whether or not an actual or projected exposure to a substance presents a public health concern. Regardless of the determination, a description of the evidence upon which this conclusion is based should be presented. This description should include a summary of the scientific uncertainties associated with this evidence. If the exposure is associated with a public health concern, the conclusions should also specify the health effects of concern, a description of the severity of the effects, and a description of the populations at risk.

Conclusions are relatively easy to make when the available information demonstrates that exposure of a specified level of a substance will or will not present a public health concern. These determinations may be applicable for certain well studied compounds. For the preponderance of chemicals, however, there is inadequate information on either exposure or toxicity. This shortcoming may place limitations of vary degrees on the conclusions that can be drawn from the data.

Depending on the level of scientific ignorance, risk assessment findings may or may not be useful in the development of an appropriate public health response to a chemical exposure situation. The measure of an effective public health response is its ability to maintain or reduce incidences of premature deaths and preventable adverse health effects. In order to best satisfy this criterion, the general public health policy should be that all avoidable exposures to toxic substances be prevented, except for those exposures in which a public health benefit can be realized. Risk assessment findings are thus most applicable to public health policy decisions either when an exposure cannot be prevented or when a preventable exposure has already occurred. The conclusions of the risks assessments should be interpreted within the context of this overall public health policy goal.

LITERATURE CITED

Adolph, E.F., 1949, "Quantitative Relations in the Physiological Constitutions of Mammals," Science, Vol. 109, pp. 579-585.

Ames, B.N., et al., 1987, "Ranking Possible Carcinogenic Hazards," Science, Vol. 236, pp. 271-280.

Anderson, N., 1986, Ranking Hazardous Air Pollutants for Formal Evaluation, First Revision, Maine Department of Human Services, Bureau of Health, Augusta, Maine, 13 pp.

Armitage, P., 1985, "Multistage Models of Carcinogenesis," Environ. Health Perspect., Vol. 63, pp. 195-201.

Balazs, T., et al., 1986, "Toxic Responses of the Cardiovascular System," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 387-411.

Berkow, R. (editor-in-chief), 1982, The Merck Manual, Fourteenth Edition, Merck Sharp and Dohme Research Laboratories, Rahway, New Jersey. 2578 pp.

Bickis, M., and Krewski, D., 1985, "Statistical Design and Analysis of the Long-Term Carcinogenicity Bioassay," in D.B. Clayson et al. (eds.), Toxicological Risk Assessment, Vol. 1, Biological and Statistical Criteria, CRC Press, Boca Raton, FL, pp. 125-147.

Bradley, M.O., et al., 1985, "Measurements of Chemical Interactions with DNA," in Flamm, W.G., and Lorentzen, R.J., (eds.), Mechanisms and Toxicity of Chemical Carcinogens and Mutagens, Princeton Scientific Publishing Co., Princeton, New Jersey, pp. 99-127.

Brown, N.A., and Fabro, S., 1983, "The Value of Animal Teratogenicity Testing for Predicting Human Risk," Clin. Obs. Gynec., Vol. 26, pp. 467-476.

Buben, J.A. and O'Flaherty, E.J., 1985, "Delineation of the Role of Metabolism in the Hepatotoxicity of Trichloroethylene and Perchloroethylene: A Dose-Effect Study," Toxicol. Appl. Pharmacol., Vol. 78, pp. 105-122.

Calabrese, E.J., 1983, Principles of Animal Extrapolation, John Wiley and Sons, New York, 603 pp.

Calabrese, E.J., 1985, "Uncertainty Factors and Interindividual Variation," Reg. Toxicol. Pharmacol., Vol. 5, pp. 190-196.

CDHS (California Department of Health Services), 1985, A Policy for Chemical Carcinogens: Guidelines for Risk Assessment and their Scientific Rationale, Department of Health Services, Sacramento, California.

Connelly, J.C., and Bridges, J.W., 1986, "Species Variation in Target Organ Toxicity," in G.M. Cohen (ed.), Target Organ Toxicity, Vol. I, CRC Press Boca Raton, FL, pp. 89-120.

Crawford, B.D., 1986, "Perspectives on the Somatic Mutation Model of Carcinogenesis," in Flamm, W.G., and Lorentzen, R.J., (eds.), Mechanisms and Toxicity of Chemical Carcinogens and Mutagens, Princeton Scientific Publishing Co., Princeton, New Jersey, pp. 13-59.

Crump, K.S., 1984, "A New Method for Determining Allowable Daily Intakes," Fund. Appl. Toxicol., Vol. 4, pp. 854-871.

Dixon, R.L., 1986 "Toxic Responses of the Reproductive System," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 432-473.

Doll, R., and Peto, R., 1981, The Causes of Cancer: Quantitative Estimates of Avoidable Risks of Cancer in the United States Today, Oxford University Press, New York, pp. 1197-1312.

Dourson, M.L. et al., 1986, Letter to the Editor, Fund. and Appl. Toxicol., Vol. 6, pp. 182-183.

Dourson, M.L., and Stara, J.F., 1983, "Regulatory History and Experimental Support of Uncertainty (Safety) Factors," Reg. Toxicol. Pharmacol., Vol. 3, pp. 224-238.

Dugard, P.H., 1987, "Skin Permeability Theory in Relation to Measurements of Percutaneous Absorption in Toxicology," in F. N. Marzulli and H.I. Maibach (eds.), Dermatoxicology, Third Edition, Washington, Hemisphere Publishing Co., pp. 95-120.

Flamm, W.G., and Lorentzen, R.J., 1985, "Editors' Introduction," in Flamm, W.G., and Lorentzen, R.J., (eds.), Mechanisms and Toxicity of Chemical Carcinogens and Mutagens, Princeton Scientific Publishing Co., Princeton, New Jersey, pp. vii-xii.

Fears, T.R., et al., 1977, "False Positive and False Negative Rates for Carcinogenicity Screens," Cancer Res., Vol. 37, pp. 1941-1945.

Freireich, E.J., et al., 1966, "Quantitative Comparison of Toxicity of Anticancer Agents in Mouse, Rat, Hamster, Dog, Monkey, and Man," Cancer Chemo. Rep., Vol. 50, pp. 219-244.

GAO (General Accounting Office), 1986, Pesticides: Need to Enhance FDA'S Ability to Protect the Public from Illegal Residues, Washington, D.C., 58 pp.

Gart, J.J., et al., 1979, "Statistical Issues in Interpretation of Chronic Bioassay Tests for Carcinogenicity," J. Nat'l Cancer Inst., Vol. 62, pp. 957-974.

Gillette, J.R., 1985, "Biological Variation: The Unsolvable Problem in Quantitative Extrapolations from Laboratory Animals and Other Surrogate Systems to Human Populations," in D.G. Hoel et al. (eds.), Risk Quantitation and Regulatory Policy, Cold Spring Harbor Laboratory, pp. 199-210.

Gold, L.S., et al., 1984, "A Carcinogenic Potency Database of the Standardized Results of Animal Bioassays," Environmental Health Perspect., Vol. 58, pp. 9-319.

Gold, L.S., et al., 1986, "Association Between Carcinogenic Potency and Tumor Pathology in Rodent Carcinogenisis Bioassays," Fund. Appl. Toxicol., Vol. 6, pp. 677-690.

Goldstein, B.D., 1983, "Toxic Substances in the Atmospheric Environment," J. Air Poll. Cont. Assoc., Vol. 33, pp. 454-467.

Halliwell, B., 1987, "Oxidants and Human Disease: Some New Concepts," FASEB J., Vol. 1, pp. 358-364.

Haseman, J.K., 1985, "Issues in Carcinogenicity Testing: Dose Selection," Fund. Appl. Toxicol., Vol. 5, pp. 66-78.

Higginson, J., 1980, "Importance of Environmental and Occupational Factors in Cancer," J. Toxicol. Environ. Health, Vol. 6, pp. 941-952.

Higginson, J. 1981, "Rethinking the Environmental Causation of Human Cancer," Food Cosmet. Toxicol, Vol. 19, pp. 539-548.

H.R. (U.S. House of Representatives), 1986, Occupational Health Hazard Surveillance: 72 Years Behind and Counting, Sixty-First Report by the Committee on Government Operations, House Report No. 99-979, U.S. Government Printing Office, Washington, D.C., 29 pp.

IARC (International Agency for Research on Cancer), 1980, Long-Term and Short-Term Screening Assays for Carcinogens: A Critical Appraisal, Supplement 2, IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Lyon, France. 426 pp.

IARC (International Agency for Research on Cancer), 1982, Chemicals, Industrial Processes and Industries Associated with Cancer in Humans, IARC Monographs Supplement 4, Lyon, France. 292 pp.

IARC (International Agency for Research on Cancer), 1983, Polynuclear Aromatic Compounds, 1: Chemical, Environmental, and Experimental Data, Vol. 32, IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Lyon, France. 477 pp.

IRLG (Interagency Regulatory Liaison Group), 1982, "Scientific bases for Identification of Potential Carcinogens and Estimation of Risks, J. Nat'l Cancer Inst., Vol. 63, pp. 242-267.

IRLG (Interagency Regulatory Liaison Group), 1986, "Interagency Regulatory Liaison Group Workshop on Reproductive Toxicity Risk Assessment," Environ. Health Perspect., Vol. 66, pp. 193-221.

Hogan, M.D., and Hoel, D.G., 1982, "Extrapolation to Man," in A.W. Hayes (ed.), Principles and Methods of Toxicology," Raven Press, New York, pp. 711-731.

Jackson, B.A., and Pertel, R., 1986, "Regulatory Implications of Ames' Mutagenicity Assay Using Salmonella typhimurium," Fund. Appl. Toxicol., Vol. 7, pp. 1-16.

Kattan, M., 1979, "Long-term Sequelae of Respiratory Illness in Infancy and Childhood," Pediatric Clinics of North America, Vol. 26, pp. 525-535.

Klaasen, C.D., 1986, "Distribution, Excretion, and Absorption of Toxicants," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 33-63.

Kleinbaum, D.G., et al., 1982, Epidemiological Research: Principles and Quantitative Methods, Lifetime Learning Publications, Belmont, CA, pp. 40-95.

Kuroki, T., and Heidelberger, C., 1971, "The Binding of Polycyclic Aromatic Hydrocarbons to the DNA, RNA, and Proteins of Transformable Cells in Culture," Cancer Res., Vol. 31, pp. 2168-2176.

Low, J.A., and Galbraith, R.S., 1974, "Pregnancy Characteristics of Intrauterine Growth Retardation," Obstetrics and Gynecology, Vol. 44, 1974, pp. 122-126.

Lu, F.C., 1985, "Safety Assessments of Chemicals with Thresholded Effects," Reg. Toxicol. Pharmacol., Vol. 5, pp. 460-464.

MacNamara, B.P., 1976, "Concepts in Health Evaluation of Commercial and Industrial Chemicals," in M.A. Mehlman et al. (eds.), New Concepts in Safety Evaluation, Hemisphere Publishing Corp., Washington, D.C., pp. 61-140.

Manson, J.M., 1986, "Teratogens," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 195-220.

MDEQE (Massachusetts Department of Environmental Quality Engineering), 1985, Chemical Health Effects Assessment Methodology, Massachusetts Department of Environmental Quality Engineering, Boston, Mass.

MacMahon, B., and Pugh, T.F, 1970, Epidemiology: Principles and Methods, Little, Brown and Company, Boston, pp. 1-46.

McQueen, C.A., and Williams, G.M., 1985, "Mammalian Cell DNA Repair Assays fof Carcinogens," in Flamm, W.G., and Lorentzen, R.J., (eds.), Mechanisms and Toxicity of Chemical Carcinogens and Mutagens, Princeton Scientific Publishing Co., Princeton, New Jersey, pp. 129-151.

Miller, E.C., et al., 1966, "Nucleic Acid Guanine: Reaction with the Carcinogen N-Acetyl-2-Acetylaminofluorene," Science, Vol. 153, pp. 1125-1127.

Miller, J.A., and Miller, E.C., 1971, "Chemical Carcinogenesis: Mechanism and Approaches to Its Control," J. Nat'l Cancer Inst., Vol. 47, pp. 5-14.

Munro, I.C., and Krewski, D.R., 1981, "Risk Assessment and Regulatory Decision Making," Food Cosmetic Toxicol., Vol. 19, pp. 549-560.

National Toxicology Program, 1985, Fourth Annual Report on Carcinogens, U.S. Department of Health and Human Services, Washington, D.C.

NIOSH (National Institute for Occupational Safety and Health), 1986, Registry of the Toxic Effects of Chemical Substances, U.S. Department of Health and Human Services, Cincinnati, Ohio.

NRC (National Research Council), 1977a, Drinking Water and Health, Vol. 1, National Academy Press, Washington, D.C., 939 pp.

NRC (National Research Council), 1977b, Drinking Water and Health, Vol. 3, National Academy Press, Washington, D.C., 415 pp.

NRC (National Research Council), 1986, Drinking Water and Health, Vol. 6, National Academy Press, Washington, D.C., pp. 1-293.

NRC (National Research Council), 1983a, Risk Assessment in the Federal Government: Managing the Process, National Academy Press, Washington, D.C., 191 pp.

NRC (National Research Council), 1983b, Identifying and Estimating the Genetic Impact of Chemical Mutagens, National Academy Press, Washington, D.C., 295 pp.

NRC (National Research Council), 1984, Toxicity Testing: Strategies to Determine Needs and Priorities, National Academy Press, Washington, D.C., 382 pp.

OSHA (Occupational Safety and Health Administration), 1980, "Identification, Classification and Regulation of Potential Occupational Carcinogens," Federal Register, Vol, 45, pp. 5002-5296.

OSTP (Office of Science and Technology Policy), 1985, "Chemical Carcinogens; A Review of the Science and Its Associated Principles, February, 1985," Federal Register, Vol. 50, pp. 10371-10442.

OTA (Office of Technology Assessment), 1981, Assessment of Technologies for Determining Cancer Risks from the Environment, Washington, D.C., 240 pp.

Page, H.S., and Asire, A.J., 1985, Cancer Rates and Risks, 3rd edition, U.S. Department of Health and Human Services, Washington, D.C., 136 pp.

Peto, R., 1977, "Epidemiology, Multi-Stage Models, and Short-term Mutagenicity Tests," in H.H. Hiatt et al. (eds), Origins of Human Cancer, Book C, Human Risk Assessment, Cold Spring Harbor Laboratory, pp. 1403-1428.

Peto, R., et al., 1984, "The TD50: A Proposed General Convention for the Numerical Description of the Carcinogenic Potency of Chemicals in Chronic-Exposure Animal Experiments," Environ. Health Perspect., Vol. 58, pp. 1-8.

Phillips, D.H., et al., 1978, "The Covalent Binding of Polycyclic Hydrocarbons to DNA in the Skin of Mice of Different Strains," Int. J. Cancer, Vol. 22, pp. 487-494.

Pinkel, D., 1958, "The Use of Body Surface Area as a Criterion of Drug Dosage in Cancer Chemotherapy," Cancer Res., Vol. 18, pp. 853-856.

Plaa, G.L., 1986, "Toxic Responses of the Liver," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York, pp. 286-309.

Purchase, I.F.H., 1980, "Interspecies Comparisons of Carcinogenicity," Br. J. Cancer, Vol. 41, pp. 454-468.

Purchase, I.F.H., 1985, "Carcinogenic Risk Assessment: A Toxicologist's View," in D.G. Hoel et al. (eds.), Risk Quantification and Regulatory Policy, Cold Spring Harbor Laboratory, pp. 175-186.

Rall, D.P., 1969, "Difficulties in Extrapolating the Results of Toxicity Studies in Laboratory Animals to Man," Environ. Res., Vol. 2, pp. 360-367.

Ramsey, J.C., and Andersen, M.E., 1984, "A Physiologically Based Description of the Inhalation Pharmacokinetics of Styrene in Rats and Humans," Toxicol. Appl. Pharmac., Vol. 73, pp. 159-175.

Reddy, J.K., and Lalwani, N.D., 1983, "Carcinogenesis by Hepatic Peroxisome Proliferators: Evaluation of the Risk of Hypolipidemic Drugs and Industrial Plasticizers to Humans," CRC Crit. Res. Toxicol., Vol. 12, pp. 1-59.

Renwick, A.G., 1982, "Pharmacokinetics in Toxicology," in A.W. Hayes (ed.), Principles and Methods of Toxicology, Raven Press, New York, pp. 659-709.

Russell, W.L., et al., 1982, "Dose-Response Curve for Ethylnitrosourea-Induced Specific Locus Mutations in Mouse Spermatogonia," Proc. Nat'l. Acad. Sci. USA, Vol. 79, pp. 3589-3591.

SAAC (Special Air Advisory Committee), 1981, A Proposed Framework for Processing Air Quality Permit Applications for New Emission Sources of Non-Criteria Pollutants, Final Report to the Michigan Air Pollution Control Commission, 27 pp.

Sato, A., and Nakajima, T., 1979, "A Structure-Activity Relationship of Some Chlorinated Hydrocarbons," Arch. Environ. Health, March/April, pp. 69-75.

Sato, A., et al., 1983, "Interaction Between Ethanol and Carbohydrate on the Metabolism in Rat Liver of Aromatic and Chlorinated Hydrocarbons," Tox. Appl. Pharmacol., Vol. 68, pp. 242-249.

Schardein, J.L., et al., 1985, "Species Sensitivities and Prediction of Teratogenic Potential," Environ. Health Perspect., Vol. 61, pp. 55-67.

Schardein, J.L., Chemically Induced Birth Defects, Marcel Dekker, Inc., New York, 1985.

Scott, K.E., and Usher, R., 1966, "Fetal Malnutrition: Its Incidence, Causes, and Effects," Amer. J. Obst. Gynecol., April 1, pp. 951-963.

Selkirk, J.K., and MacLeod, M.C., 1979, "Metabolism and Macromolecular Binding of Benzo (a) Pyrene and Its Noncarcinogenic Isomer Benzo (e) Pyrene in Cell Culture," in P.W. Jones and P. Leber (eds.), Polynuclear Aromatic Hydrocarbons: Third International Symposium on Chemistry and Biology - Carcinogenesis and Mutagenesis, Ann Arbor Science, Ann Arbor, Michigan, pp. 21-35.

Singer, B., 1975, "The Chemical Effects of Nucleic Acid Alkylation and Their Relation to Mutagenesis and Carcinogenesis," Proc. Nucl. Acid Res. Mol. Biol., Vol. 15, pp. 219-284.

Sipes, I.G., and Gandolfi, A.J., 1986, "Biotransformation of Toxicants," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 64-183.

Spiers, P.S., 1982, "Does Growth Retardation Predispose the Fetus to Congenital Malformation," The Lancet, February 6, pp. 312-314.

Squire, R.L., 1981, "Ranking Animal Carcinogens: A Proposed Regulatory Approach," Science, Vol. 214, pp. 877-880.

Stanier, R.Y., et al., 1976, The Microbial World, Prentice Hall, Englewood Cliffs, NJ, pp. 86-87.

Tennant, R.W., et al., 1987, "Prediction of Chemical Carcinogenicity in Rodents from in Vitro Genetic Toxicity Assays," Science, Vol. 236, pp. 933-941.

Thilly, W.G., and Call, K.M., 1986, "Genetic Toxicology," in C.D. Klaasen et al. (eds.), Casarett and Doull's Toxicology, MacMillan Publishing Company, New York. pp. 174-183.

Tichy, M., 1983, "Prediction of Adverse Activities from Physical and Chemical Properties of Vapors and Gases (QSAR Analysis)," in V. Fiserova-Bergerova (ed.), Modeling of Inhalation Exposure to Vapors: Uptake, Distribution and Elimination, Vol. II, CRC Press, Boca Raton, Florida, pp. 3-35.

Timbrell, J.A., 1986, "Liver as a Target Organ for Toxicity," in G.M. Cohen (ed.), Target Organ Toxicity, CRC Press, Boca Raton, FL, pp. 145-173.

Tomatis, L., et al., 1982, "Experimental Studies in the Assessment of Human Risk," in D. Shottenfeld and J.F. Fraumeni, Jr. (eds.), Cancer Epidemiology and Prevention, W.B. Saunders Company, Philadelphia, pp. 44-73.

USEPA (U.S. Environmental Protection Agency), 1987, Assessment of Health Risks to Garment Workers and Certain Home Residents from Exposure to Formaldehyde, Office of Pesticides and Toxic Substances, Washington, D.C., p. 1-15.

USEPA (U.S. Environmental Protection Agency), 1986a, "Guidelines for Carcinogen Risk Assessment," Federal Register, Vol. 51, pp. 33991-34003.

USEPA (U.S. Environmental Protection Agency), 1986b, "Guidelines for Mutagenicity Risk Assessment," Federal Register, Vol. 51, pp. 34005-34012.

USEPA (U.S. Environmental Protection Agency). 1986c, "Guidelines for the Health Assessment of Chemical Mixtures," Federal Register, Vol. 51, pp. 34013-34025.

USEPA (U.S. Environmental Protection Agency). 1986d, "Guidelines for the Health Assessment of Suspect Developmental Toxicants," Federal Register, Vol. 51, pp. 34027-34040.

USEPA (U.S. Environmental Protection Agency). 1986e, "Guidelines for Exposure Assessment," Federal Register, Vol. 51, pp. 34041-34054.

USEPA (U.S. Environmental Protection Agency), 1981, Chemical Substances Designation, Vol. I: Overview and Analysis, Washington, D.C., 154pp.

Vaino, H., 1987, Chief, Unit of Carcinogen Identification and Evaluation, International Agency for Research on Cancer, Letter to Elizabeth Bourque, Massachusetts Department of Public Health, January 20, 1987.

van den Berg, B.J., and Yerushalmy, J., 1966, "The Relationship of the Rate of Intrauterine Growth of Infants of Low Birth Weight to Mortality, Morbidity, and congenital Anomalies," J. Pediat, Vol. 69, pp. 531-545.

Vaughan, T.L., et al., 1986, "Formaldehyde and Cancers of the Pharynx, Sinus and Nasal Cavity: II. Residential Exposures," Int. J. Cancer, in press.

Wilson, J.G., 1973, Environment and Birth Defects, Academic Press, New York, pp. 137-171.

Wilson, J.G., 1980, "Environmental Effects on Intrauterine Death in Animals," in I.H. Porter and E.B. Hook (eds.), Human Embryonic and Fetal Death, Academic Press, New York, pp. 19-44.

WHO (World Health Organization), 1984, Guidelines for Drinking Water Quality, vol. 2: Health Criteria and Other Supporting Documentation, Geneva, pp. 170-176.

WHO (World Health Organization), 1985, Guide to Short-Term Tests for Detecting Mutagenic and Carcinogenic Chemicals, Environmental Health Criteria 51, World Health Organization, Geneva, pp. 24-125.

WHO (World Health Organization), 1986, Principles for Evaluating Health Risks from Chemicals during Infancy and Early Childhood: The Need for a Special Approach, Environmental Health Criteria 59, World Health Organization, Geneva, 73 pp.

Young, J.L., and Pollack, E.S. 1982, "The Incidence of Cancer in the United States," in D. Shottenfeld and J.F. Fraumeni, Jr. (eds.), Cancer Epidemiology and Prevention, W.B. Saunders Company, Philadelphia, pp. 138-165.

Yuspa, S., and Harris, C., 1982, "Molecular and Cellular Basis of Chemical Carcinogenesis," in D. Shottenfeld and J.F. Fraumeni, Jr. (eds.), Cancer Epidemiology and Prevention, W.B. Saunders Company, Philadelphia, pp. 23-43.